Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/150611

Static and Dynamic Panel Data Models under a Complex Survey Design: A Statistical Assessment of Early Screen Exposure and Children's Learning at Ages 6–8 in the ECLS-K:2011

Jawharah Alfehaid[1], *Mona Gadan*[1,*], *Gayda Manssour*[1], *Imad Muhi*[2], *and Nurkasym Arkabaev*[3]

[1]Department of Early Childhood, College of Education, Qassim University, Buraydah, Saudi Arabia
[2]Department of Mathematics and Natural Sciences, College of Sciences and Human Studies, Prince Mohammad Bin Fahd University (PMU), Al Khobar, Saudi Arabia
[3]Institute of Mathematics, Physics, Engineering and Information Technologies, Osh State University, Osh 723500, Kyrgyzstan

Received: 25 Aug. 2026, Revised: 8 Oct. 2026, Accepted: 18 Oct. 2026
Published online: 1 Nov. 2026

Abstract: Concern that screen time harms young children's learning rests largely on observational associations whose size depends on how confounding is handled. We estimate the association between weekday television viewing and reading, mathematics, working memory and teacher-rated self-control among 7,486 children followed from the spring of kindergarten (age 6.1) to the spring of second grade (age 8.1) in the nationally representative Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011). We compare five static estimators on the same data (unadjusted and adjusted pooled regression, a value-added model, child fixed effects and a Mundlak decomposition) and three dynamic ones (Anderson–Hsiao first-difference instrumental variables, the cross-lagged panel model and its random-intercept extension), with all standard errors obtained by paired-jackknife replication (80 replicates) that repeats every estimation step. Each additional daily hour of television is associated with -0.115 SD in reading and -0.139 SD in mathematics without adjustment, -0.049 and -0.056 SD after adjustment, but only -0.008 and -0.010 SD within children; 82–95% of the adjusted associations with academic outcomes reflect stable differences between children. Within-child estimates are null for working memory and self-control, show no difference by family poverty, and pass a falsification test with future viewing. A small within-child penalty remains for heavy viewing above three hours a day (-0.05 to -0.06 SD). Dynamic models lead to the same conclusion: the instrumental-variable estimates are close to zero, and the significant cross-lagged effects of the classical cross-lagged model in both directions largely vanish once random intercepts separate stable between-child differences from within-child dynamics, except for a small lagged association with working memory. Computer use at age 6 predicts higher reading gains, and video-game time is unrelated to learning once children are compared with themselves. The results caution against causal readings of cross-sectional screen-time associations in early childhood.

Keywords: dynamic panel data, random-intercept cross-lagged panel model, Anderson–Hsiao estimator, fixed effects, Mundlak decomposition, jackknife replication, complex survey data, screen time, early childhood

1 Introduction

Few questions about early childhood attract as much public attention as the effect of screens on learning. Paediatric guidance recommends limiting screen time and favouring high-quality content [1], and the evidence base it draws on is large. Meta-analyses report that television viewing and video-game playing are inversely associated with academic performance in children and adolescents [2], and that greater quantity of screen use is associated with poorer language skills while educational content and co-viewing are associated with better ones [3]. A review of reviews found moderately strong evidence linking screen time to adiposity and depressive symptoms but weaker evidence for other outcomes [4].

Most of this evidence is observational, and the size of the reported associations depends heavily on how confounding is handled. Children who watch more television differ systematically from those who watch less in family resources,

* Corresponding author e-mail: m.gadan@qu.edu

© 2026 NSP
Natural Sciences Publishing Cor.

parental education, home learning activities and many characteristics that surveys do not record. When researchers exploit variation that is plausibly unrelated to such factors, the picture changes. Using the staggered introduction of television in the United States, Gentzkow and Shapiro [5] found no negative effect of preschool television exposure on adolescent test scores and could rule out negative effects larger than about 0.03 standard deviations per year of exposure. Kearney and Levine [6] found that access to an educational programme, *Sesame Street*, improved school readiness. In a different domain, a specification-curve analysis showed that digital technology use explains a very small share of the variance in adolescent well-being [7].

The present paper asks a simple statistical question with practical consequences: *how much of the association between screen time and early learning survives when each child is compared with himself or herself over time?* We use the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011), a nationally representative U.S. panel in which the same children were assessed directly every spring and their parents reported weekday television viewing [8]. We focus on ages 6 to 8, the first three years of formal schooling, when reading and mathematics skills develop quickly.

Our contribution is methodological as much as substantive. We estimate the same association with a ladder of five estimators on the same sample: unadjusted and adjusted pooled regressions, a value-added model with the lagged outcome, a child fixed-effects model and a Mundlak decomposition into within- and between-child components [9, 10]. All standard errors account for the complex sample design through paired-jackknife (JK2) replication, and every step of each estimator, including the within transformation and the child means, is repeated for each of the 80 replicate weights. This design makes transparent how much of the familiar negative association is due to stable differences between children. Because these static models cannot represent feedback over time, we add three dynamic specifications: the Anderson–Hsiao first-difference instrumental-variable estimator, which handles a lagged outcome together with child effects [32], and the cross-lagged panel model with and without random intercepts, which estimates effects in both directions, from viewing to learning and from learning to viewing [24, 35]. We complement the main analysis with tests of non-linearity, heterogeneity by family poverty, a falsification test based on future viewing, and analyses of computer use and video games.

Section 2 reviews related work, Section 3 presents the data and Section 4 the statistical methods. Section 5 reports the results, Section 6 discusses them and Section 7 concludes.

2 Related work

Screen time and early learning. Early longitudinal studies reported that television viewing before age 3 was associated with modestly lower reading and memory scores at ages 6–7, whereas viewing at ages 3–5 was associated with slightly higher reading recognition [11]. Television exposure in toddlerhood has been linked to later attentional problems [12] and to lower classroom engagement and mathematics achievement in middle childhood [13], and media exposure in infancy to lower cognitive and language development in toddlers [14]. More recently, a random-intercept cross-lagged analysis of a Canadian birth cohort found that higher screen time at 24 and 36 months was associated with poorer performance on a developmental screening test at later ages [15]. Reviews emphasise that content and context matter more than total time: educational programmes can benefit preschoolers, while background television and fast-paced entertainment are more problematic [16, 17]. Experiments support this view: nine minutes of a fast-paced cartoon impaired the executive function of four-year-olds immediately afterwards [18], and the effect of screen media on young children's executive functioning depends on the type of content [19].

Interactive media. Evidence on computers, apps and games is more favourable. Educational applications can support learning when they foster active, engaged, meaningful and socially interactive use [20]. Among U.S. children aged 9–10, gaming was associated with gains in intelligence over two years after controlling for genetic predispositions and socioeconomic status, whereas watching and socialising showed negative cross-sectional correlations [21]. Young children's media environments have also changed rapidly since the 2010s, with mobile devices now part of everyday life for children under 8 [22].

Statistical issues. Three methodological strands are relevant. First, the estimation of cognitive production functions requires strong assumptions about the role of past inputs and unobserved endowments; value-added models with lagged achievement and child fixed-effects models rest on different and testable restrictions [23]. Second, panel models that do not separate within-person from between-person variance can confound stable traits with dynamic effects, a criticism directed at the cross-lagged panel model [24, 25]. The random-intercept cross-lagged panel model (RI-CLPM) addresses this problem by separating stable between-person differences from within-person fluctuations [24, 35], and it has been argued that the classical model is almost never the right choice for causal questions [36]. In econometrics, the corresponding issue arises when a lagged dependent variable is combined with individual effects: the within estimator is then biased in short panels [34], and consistent estimators difference out the individual effect and instrument the lagged difference with earlier levels [32, 33]. Fixed-effects regression removes all time-invariant confounders [26, 27], and the Mundlak specification recovers the within estimates while keeping the between-person component [9, 10]. Third, analyses of complex survey data require sampling weights and design-based variance estimation; replication methods

© 2026 NSP
Natural Sciences Publishing Cor.

such as the jackknife are general and easy to apply to any estimator [28,29,30]. Whether to weight regression estimates at all is itself a substantive choice [31], which motivates our unweighted robustness check.

3 Data

3.1 The ECLS-K:2011

The ECLS-K:2011 follows a nationally representative sample of children who attended kindergarten in the United States in 2010–11 [8]. The sample was selected in three stages (geographic primary sampling units, schools, and children within schools), and children were followed through fifth grade. We use the public-use kindergarten–fifth grade data file and three spring rounds: kindergarten (2011), first grade (2012) and second grade (2013), referred to as waves 1, 2 and 3. Children's average age at assessment was 6.12, 7.12 and 8.13 years.

Our analytic sample comprises the 7,486 children with a positive value of the longitudinal weight W6C6P60, which adjusts for non-response to the child assessments and parent interviews in all relevant rounds up to second grade. These children attended 898 schools in the spring of kindergarten. Table 1 reports weighted descriptive statistics.

3.2 Measures

Outcomes. Reading and mathematics were assessed with adaptive two-stage tests and scored with item response theory; we use the theta scores. Working memory was measured with the Numbers Reversed task (W-ability score). Self-control is the teacher-reported scale from the Social Rating Scale. Each outcome was standardised within wave to a weighted mean of 0 and standard deviation of 1, so coefficients are in standard-deviation (SD) units.

Television. Parents reported how many hours and minutes the child watched television on a typical weekday. In kindergarten the question was asked separately for three periods of the day (before 8 a.m., 8 a.m.–6 p.m., after 6 p.m.), which we sum; in first and second grades it was asked as a single weekday total. Values above 16 hours were treated as missing. Children watched 2.03 hours per weekday on average in kindergarten, 1.63 in first grade and 1.58 in second grade; the share watching more than three hours fell from 15.2% to 5.0%.

Computer use and video games. In kindergarten parents reported whether there was a computer at home that the child used and how often (never, once or twice a week, 3–6 times a week, every day). In first and second grades parents reported weekday hours of video-game play (0.58 and 0.65 hours on average).

Covariates. Time-invariant covariates are sex, race/ethnicity (White, Black, Hispanic, Asian, other), non-English primary home language and the continuous socioeconomic status composite of the first year, standardised. Age at assessment and wave indicators are time-varying. Family poverty (household income below the federal poverty threshold, 25.6% of children) is used for heterogeneity analysis.

Missing data. Outcomes are missing for less than 1% of the analytic sample and television for 1.8–3.1% depending on the wave. Analyses use all available child-waves for each outcome.

Table 1: Weighted descriptive statistics of the analytic sample (7,486 children)

Variable	Mean (SE)	≤1 h (%)	>3 h (%)	Missing (%)
TV hours, age 6	2.03 (0.03)	30.7	15.2	1.8
TV hours, age 7	1.63 (0.03)	48.8	6.1	2.3
TV hours, age 8	1.58 (0.02)	50.0	5.0	3.1
Video-game hours, age 7	0.58 (0.02)	91.3	1.0	2.3
Video-game hours, age 8	0.65 (0.01)	88.9	0.8	3.2
Computer use, age 6: never (%)	28.9 (0.7)			0.7
Computer use, age 6: 1-2/week (%)	37.2 (0.7)			0.7
Computer use, age 6: 3-6/week (%)	22.8 (0.5)			0.7
Computer use, age 6: daily (%)	11.1 (0.4)			0.7
Girl (%)	48.5 (0.4)			0.0
Below poverty (%)	25.6 (1.0)			0.0
Non-English home (%)	15.1 (0.8)			0.0
Age at assessment, wave 1 (years)	6.12 (0.01)			0.2
Age at assessment, wave 2 (years)	7.12 (0.01)			0.2
Age at assessment, wave 3 (years)	8.13 (0.01)			0.3

Note. Estimates weighted with the longitudinal weight W6C6P60; standard errors from 80 paired-jackknife (JK2) replicates. Television and video-game time are parent-reported hours on a typical weekday; values above 16 hours were treated as missing. Computer use at age 6 combines households without a computer the child uses ("never") with the reported frequency of use.

© 2026 NSP
Natural Sciences Publishing Cor.

4 Statistical methods

4.1 Five estimators

Let y_{it} denote an outcome of child i at wave $t \in \{1,2,3\}$, x_{it} weekday television hours, $\mathbf{z}_i$ time-invariant covariates and $\mathbf{d}_{it}$ wave indicators and age at assessment. We write the outcome as

$$y_{it} = \beta x_{it} + \boldsymbol{\delta}^{\top}\mathbf{d}_{it} + \boldsymbol{\gamma}^{\top}\mathbf{z}_i + \alpha_i + \varepsilon_{it}, \tag{1}$$

where α_i collects all stable unobserved characteristics of the child and family. The question is how β is estimated when α_i is correlated with television viewing.

E1 (pooled, unadjusted). Weighted least squares of y_{it} on x_{it} and $\mathbf{d}_{it}$ across all child-waves, ignoring $\mathbf{z}_i$ and α_i.

E2 (pooled, adjusted). E1 plus the observed covariates $\mathbf{z}_i$; α_i is still absorbed in the error term.

E3 (value-added). For waves 2 and 3, the outcome is regressed on lagged television and the lagged outcome,

$$y_{it} = \beta_{VA}\, x_{i,t-1} + \lambda\, y_{i,t-1} + \boldsymbol{\delta}^{\top}\mathbf{d}_{it} + \boldsymbol{\gamma}^{\top}\mathbf{z}_i + u_{it}. \tag{2}$$

The lagged outcome is a summary of past inputs and endowments; the model is consistent if it captures all relevant history [23].

E4 (child fixed effects). Subtracting child means from (1) eliminates α_i and $\mathbf{z}_i$:

$$y_{it} - \bar{y}_i = \beta_W\,(x_{it} - \bar{x}_i) + \boldsymbol{\delta}^{\top}(\mathbf{d}_{it} - \bar{\mathbf{d}}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i). \tag{3}$$

β_W is identified only from changes in a child's own viewing between waves and is free of any time-invariant confounder [26,27]. Children observed in only one wave do not contribute.

E5 (Mundlak decomposition). Adding the child mean $\bar{x}_i$ to the pooled model,

$$y_{it} = \beta_W\, x_{it} + \theta\, \bar{x}_i + \boldsymbol{\delta}^{\top}\mathbf{d}_{it} + \boldsymbol{\gamma}^{\top}\mathbf{z}_i + e_{it}, \tag{4}$$

yields a within coefficient β_W equal to that of (3) when the child means of all time-varying regressors are included, and nearly identical to it when, as here, only the mean of the exposure is added [9,10]. The between-child association is $\beta_B = \beta_W + \theta$: the expected difference between two children whose average viewing differs by one hour. Comparing β_B with β_W shows how much of the pooled association reflects stable differences between children.

4.2 Dynamic specifications

The static models assume that current viewing affects current outcomes only. To allow for persistence and feedback, we estimate three dynamic models on children observed in all three waves.

Anderson–Hsiao estimator. Adding the lagged outcome to (1) gives $y_{it} = \rho\, y_{i,t-1} + \beta x_{it} + \boldsymbol{\delta}^{\top}\mathbf{d}_{it} + \alpha_i + \varepsilon_{it}$. A within transformation would correlate the demeaned lagged outcome with the demeaned error and bias ρ and β in a three-wave panel [34]. Differencing removes α_i,

$$\Delta y_{i3} = \rho\, \Delta y_{i2} + \beta\, \Delta x_{i3} + \Delta\delta + \Delta\varepsilon_{i3}, \tag{5}$$

and Δy_{i2} is instrumented with y_{i1}, which is uncorrelated with $\Delta\varepsilon_{i3} = \varepsilon_{i3} - \varepsilon_{i2}$ if the errors are serially uncorrelated [32,33]. With three waves, (5) is the only differenced equation. We consider two assumptions about television: *strict exogeneity*, under which Δx_{i3} is its own instrument, and *predetermination*, which allows current viewing to respond to past outcome shocks and uses x_{i1} and x_{i2} as instruments for Δx_{i3}. Estimation is by weighted two-stage least squares, and instrument strength is assessed with weighted first-stage F statistics.

Cross-lagged panel models. Let $\mathbf{v}_{it} = (x_{it}, y_{it})^{\top}$. The RI-CLPM decomposes

$$\mathbf{v}_{it} = \boldsymbol{\mu}_t + \boldsymbol{\kappa}_i + \mathbf{w}_{it}, \qquad \mathbf{w}_{it} = \mathbf{B}\,\mathbf{w}_{i,t-1} + \boldsymbol{\zeta}_{it}, \qquad \mathbf{B} = \begin{pmatrix} a_x & c_{yx} \\ c_{xy} & a_y \end{pmatrix}, \tag{6}$$

where $\boldsymbol{\kappa}_i$ are correlated random intercepts with covariance matrix $\boldsymbol{\Phi}$, $\mathbf{w}_{it}$ are within-child deviations, a_x and a_y are autoregressive parameters and c_{xy} ($x_{t-1} \to y_t$) and c_{yx} ($y_{t-1} \to x_t$) are cross-lagged effects [24,35]. The classical CLPM is the special case $\boldsymbol{\Phi} = \mathbf{0}$, in which cross-lagged paths mix within- and between-child variation. We constrain $\mathbf{B}$ to be

© 2026 NSP
Natural Sciences Publishing Cor.

equal over time and leave the wave-1 covariance of $\mathbf{w}_{i1}$ and the innovation covariances of $\boldsymbol{\zeta}_{i2}$ and $\boldsymbol{\zeta}_{i3}$ free. The implied covariance matrix of the six repeated measures is

$$\operatorname{Cov}(\mathbf{v}_{it},\mathbf{v}_{is}) = \boldsymbol{\Phi} + \mathbf{B}^{t-s}\mathbf{C}_s, \qquad t \geq s, \qquad \mathbf{C}_t = \mathbf{B}\,\mathbf{C}_{t-1}\mathbf{B}^{\top} + \boldsymbol{\Psi}_t, \tag{7}$$

with $\mathbf{C}_1$ the wave-1 within covariance and $\boldsymbol{\Psi}_t$ the innovation covariance. With 16 free parameters for 21 distinct moments, the RI-CLPM has 5 degrees of freedom and the CLPM 8. Both are estimated by minimising the maximum-likelihood discrepancy $F = \log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \operatorname{tr}\{\mathbf{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\} - \log|\mathbf{S}| - 6$, where $\mathbf{S}$ is the weighted covariance matrix; this pseudo-maximum-likelihood estimator is consistent for the population parameters under the sampling design, and its sampling variance is obtained by replication.

4.3 Additional analyses

Non-linearity: x_{it} is replaced by indicators for $\leq$1, 2–3 and $>$3 hours (reference: 1–2 hours) in E2 and E4. *Heterogeneity*: an interaction between television and family poverty is added to E4. *Falsification*: in a fixed-effects model for waves 1–2, next-wave television $x_{i,t+1}$ is added; if within-child changes in viewing caused changes in outcomes, future viewing should not predict current outcomes once current viewing is controlled. *Other media*: the frequency of computer use in kindergarten (3 or more times a week versus less) is related to second-grade outcomes in adjusted and value-added models, and video-game hours at waves 2–3 enter pooled and fixed-effects models together with television. *Unweighted estimation*: E2 and E4 are re-estimated without weights, with standard errors clustered by child, following the recommendation to compare weighted and unweighted estimates [31].

4.4 Design-based variance estimation

The ECLS-K:2011 provides 80 replicate weights $w_i^{(r)}$ for the paired jackknife (JK2) [8]. For any estimator $\hat{\beta}$ computed with the full-sample weight w_i, the variance is

$$\widehat{V}(\hat{\beta}) = \sum_{r=1}^{80} \left(\hat{\beta}^{(r)} - \hat{\beta}\right)^2, \tag{8}$$

where $\hat{\beta}^{(r)}$ is computed with replicate weight r [28,29]. Because the weights are constant within child, the within transformation in (3) is the same for all replicates, but the child means in (4), the lagged variables in (2) and the regression coefficients are recomputed for each replicate. For the dynamic models, the instrumental-variable regressions and the maximisation of the likelihood are likewise repeated for each replicate weight. The variance of the difference between two estimators, such as $\hat{\beta}_{E2} - \hat{\beta}_{E4}$, is obtained in the same way from the replicate differences, which automatically accounts for the covariance between the two estimates. Tests use $|t| > 1.96$ for $p < 0.05$.

4.5 Software

All computations were programmed in Python (`numpy`, `pandas`, `statsmodels`). The public-use data file was read directly from its fixed-width ASCII layout using the official SPSS syntax. The fixed-effects implementation was checked against a regression on child dummy variables (identical coefficients to six decimals). The cross-lagged estimator was checked by Monte Carlo simulation: in 20 simulated three-wave panels of 7,000 children generated from an RI-CLPM with $a_x = 0.2$, $a_y = 0.4$, $c_{xy} = -0.04$ and $c_{yx} = -0.05$, the mean estimates were 0.196, 0.406, -0.042 and -0.049. The code is provided as supplementary material.

5 Results

5.1 From pooled to within-child estimates

Table 2 and Figure 1 report the five estimators for the four outcomes. Without adjustment (E1), each additional weekday hour of television is associated with 0.115 SD lower reading, 0.139 SD lower mathematics, 0.086 SD lower working

© 2026 NSP
Natural Sciences Publishing Cor.

memory and 0.040 SD lower self-control, all highly significant. Adjusting for sex, race/ethnicity, home language and socioeconomic status (E2) more than halves these associations, to −0.049, −0.056, −0.031 and −0.007 SD.

Within children (E4), the associations shrink further, to −0.008 SD for reading (SE 0.004, $p \approx 0.06$), −0.010 SD for mathematics (SE 0.004, $p < 0.05$), −0.001 SD for working memory and +0.005 SD for self-control. The difference between the adjusted pooled and within-child estimates is significant for the three cognitive outcomes (bottom panel of Table 2): 85% of the adjusted reading association, 82% of the mathematics association and 95% of the working-memory association disappear when stable child characteristics are removed. The Mundlak model makes the same point from the other side: the between-child association is −0.079 SD for reading and −0.084 SD for mathematics, about ten times the within-child association.

The value-added model (E3) lies close to the fixed-effects results for reading and mathematics (−0.006 and +0.001 SD) but not for working memory, where lagged television retains a negative coefficient (−0.027 SD, SE 0.007). Since working memory has the lowest wave-to-wave stability of the three cognitive outcomes (coefficient on the lagged outcome about 0.4–0.5, against 0.65–0.84 for reading and mathematics), the lagged score is a weaker control for stable differences, and the fixed-effects estimate is the more credible one.

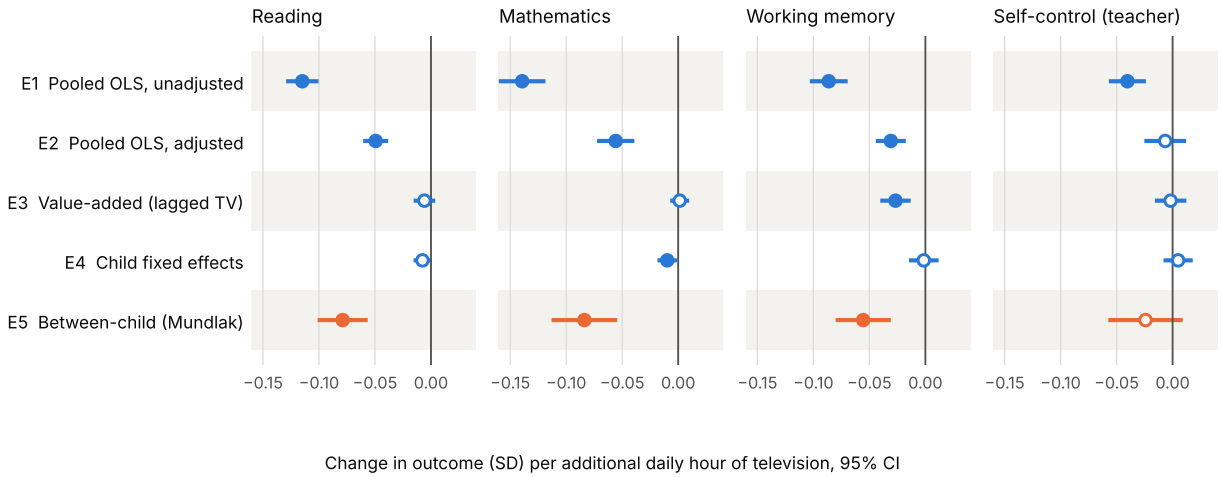

Fig. 1: Association between daily television hours and four outcomes under five estimators (SD units per hour, 95% confidence intervals from 80 JK2 replicates). Filled markers: $p < 0.05$; open markers: not significant. E5 shows the between-child component of the Mundlak decomposition; its within component equals E4.

Table 2: Association between daily television hours and outcomes under five estimators (SD units per hour)

Estimator	Reading	Mathematics	Working memory	Self-control
E1 Pooled OLS, unadjusted	−0.115*** (0.007)	−0.139*** (0.011)	−0.086*** (0.009)	−0.040*** (0.008)
E2 Pooled OLS, adjusted	−0.049*** (0.006)	−0.056*** (0.009)	−0.031*** (0.007)	−0.007 (0.010)
E3 Value-added (lagged TV)	−0.006 (0.005)	0.001 (0.004)	−0.027*** (0.007)	−0.002 (0.007)
E4 Child fixed effects (within)	−0.008 (0.004)	−0.010* (0.004)	−0.001 (0.007)	0.005 (0.007)
E5 Mundlak: within	−0.007 (0.004)	−0.009* (0.005)	−0.001 (0.007)	0.006 (0.007)
E5 Mundlak: between	−0.079*** (0.011)	−0.084*** (0.015)	−0.056*** (0.013)	−0.024 (0.017)
Difference E2 − E4	−0.042*** (0.006)	−0.046*** (0.008)	−0.029*** (0.007)	−0.012 (0.009)
Children (E4)	7 430	7 426	7 431	6 980
Child-waves (E4)	21 811	21 793	21 809	19 660

Note. Outcomes standardised within wave (weighted mean 0, SD 1). All models include wave indicators and age at assessment; E2, E3 and E5 add sex, race/ethnicity, non-English home language and socioeconomic status; E3 also includes the lagged outcome. E4 and E5 use children observed in at least two waves. Standard errors (in parentheses) from 80 JK2 replicates, re-estimating every step of each estimator. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-sided).

© 2026 NSP
Natural Sciences Publishing Cor.

5.2 Heavy viewing, heterogeneity and falsification

Pooled models suggest a non-linear pattern (Table 3, panel a, and Figure 2): compared with 1–2 hours, children watching at most one hour score about 0.06 SD higher in reading and mathematics, and those watching more than three hours about 0.19–0.21 SD lower. Within children, the advantage of light viewing vanishes, but a smaller penalty for heavy viewing remains: when a child's viewing rises above three hours, reading and mathematics scores are 0.051 and 0.059 SD lower than in waves in which the same child watched 1–2 hours. No such pattern appears for working memory or self-control.

The within-child association does not differ by family poverty (panel b): the interaction terms are small and non-significant for all four outcomes. In the falsification test (panel c), next-wave television is not significantly associated with current outcomes (−0.009 SD for reading, −0.007 for mathematics). For reading, however, its point estimate equals that of current viewing, so the small within-child reading association cannot be distinguished from a common trend in viewing and achievement.

Unweighted estimates with child-clustered standard errors (panel d) are very close to the weighted ones: −0.052 and −0.061 SD for reading and mathematics in the adjusted pooled model, and −0.004 and −0.007 SD within children. The main conclusions therefore do not depend on the weighting.

Table 3: Non-linearity, heterogeneity, falsification and unweighted robustness checks (SD units)

	Reading	Mathematics	Working memory	Self-control
(a) Television categories (reference: 1–2 hours)				
Pooled OLS, ≤1 h	0.064*** (0.016)	0.065*** (0.018)	0.049** (0.016)	0.016 (0.022)
Pooled OLS, 2–3 h	−0.012 (0.025)	−0.052* (0.025)	−0.024 (0.029)	−0.003 (0.031)
Pooled OLS, >3 h	−0.188*** (0.033)	−0.206*** (0.039)	−0.100** (0.038)	−0.030 (0.039)
Child FE, ≤1 h	−0.004 (0.010)	0.006 (0.009)	0.025 (0.023)	−0.024 (0.021)
Child FE, 2–3 h	−0.022 (0.015)	−0.028* (0.013)	−0.012 (0.028)	−0.009 (0.028)
Child FE, >3 h	−0.051* (0.022)	−0.059** (0.019)	0.031 (0.034)	−0.020 (0.027)
(b) Heterogeneity by family poverty (child FE)				
TV hours	−0.009 (0.005)	−0.016** (0.006)	−0.002 (0.008)	0.001 (0.008)
TV hours × below poverty	0.004 (0.011)	0.019 (0.011)	0.002 (0.016)	0.014 (0.017)
(c) Falsification: future television (child FE, waves 1–2)				
Current TV hours	−0.009 (0.008)	−0.017* (0.008)	0.013 (0.010)	−0.015 (0.013)
Next-wave TV hours	−0.009 (0.008)	−0.007 (0.007)	0.018 (0.014)	−0.014 (0.013)
(d) Unweighted estimates, child-clustered standard errors				
Pooled OLS, adjusted (E2)	−0.052*** (0.007)	−0.061*** (0.007)	−0.032*** (0.006)	−0.007 (0.007)
Child fixed effects (E4)	−0.004 (0.004)	−0.007 (0.003)	−0.002 (0.006)	0.003 (0.006)

Note. Panels (a)–(c) are weighted with JK2 standard errors; covariates as in Table 2. In panel (c), a non-zero coefficient on next-wave television would indicate that within-child changes in viewing anticipate, rather than follow, changes in outcomes. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-sided).

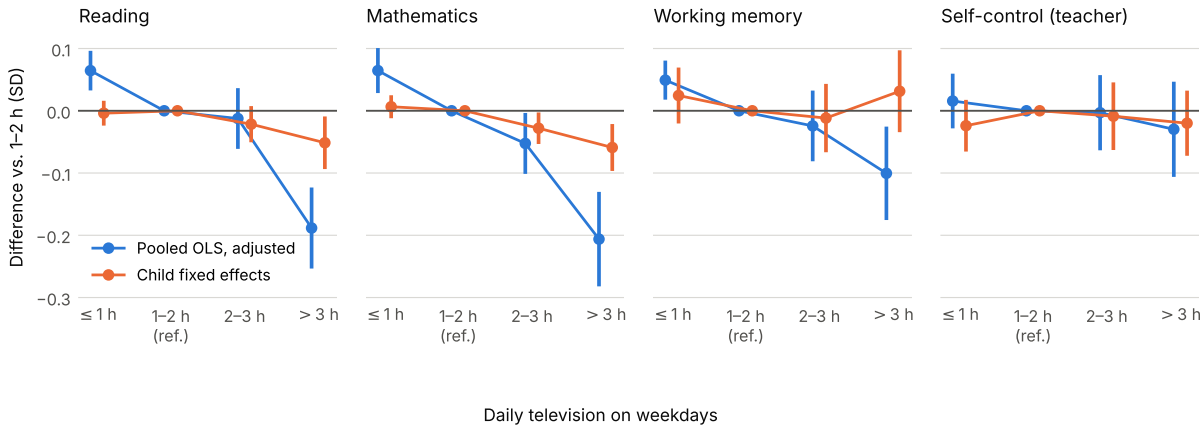

Fig. 2: Television categories and outcomes: adjusted pooled regression versus child fixed effects (difference from 1–2 hours per weekday, SD units, 95% confidence intervals).

© 2026 NSP
Natural Sciences Publishing Cor.

5.3 Dynamic models

Table 4 and Figure 3 report the dynamic specifications, estimated on the 6,951 children (5,705 for self-control) observed in all three waves.

Anderson–Hsiao. The instruments are strong: weighted first-stage F statistics exceed 240 for the lagged outcome difference and 1,600 for the television difference. Treating television as strictly exogenous, an additional daily hour is associated with −0.004 SD in reading (SE 0.007), −0.005 SD in mathematics (SE 0.007) and +0.002 SD in working memory (SE 0.012), all far from significance. Allowing television to respond to past outcome shocks (predetermined television) gives the same picture with wider intervals. The only significant coefficient is a small positive association with teacher-rated self-control (+0.024 SD, SE 0.010) under strict exogeneity, which is not robust to the predetermined specification. The lagged-outcome coefficients ($\hat{\rho}$ between 0.06 and 0.25) are much smaller than the value-added coefficients of Section 5.1, as expected once stable differences between children are removed.

Classical CLPM. Without random intercepts, the cross-lagged model finds significant paths in both directions for all four outcomes. One hour of television predicts 0.017–0.048 SD lower outcomes a year later, and one SD higher achievement predicts 0.065–0.105 fewer hours of television a year later: the path from learning to viewing is larger than the path from viewing to learning. The model fits poorly (RMSEA 0.12–0.14; as for all fit indices reported here, these values are indicative only under the complex sample design).

RI-CLPM. Adding correlated random intercepts improves the fit markedly (RMSEA 0.025–0.067, χ^2 reduced by a factor of 5 to 43 for 3 degrees of freedom) and changes the conclusions. These fit statistics are computed as if the weighted covariance matrix came from a simple random sample of 5,705–6,951 children; under the complex design their reference distribution is unknown, so they should be read as descriptive comparisons between the CLPM and the RI-CLPM fitted to the same matrix, not as formal tests. Inference on the parameters themselves relies on the jackknife standard errors, which are valid under the design. The random intercepts of television and the outcomes are negatively correlated (−0.16 to −0.34): children who watch more television than average over the three years also score lower over the three years. Within children, however, the cross-lagged effects of television on reading (−0.005, SE 0.010), mathematics (+0.012, SE 0.009) and self-control (+0.005, SE 0.017) are not significant, and neither are the reverse paths from outcomes to television. The one exception is working memory: an hour of television above a child's own average predicts 0.032 SD lower working memory a year later (SE 0.014, $p < 0.05$). Within-child deviations in television are not persistent ($\hat{a}_x \approx 0$), which is consistent with the measurement-related wave-to-wave variation discussed in Section 6.

Table 4: Dynamic panel specifications: first-difference instrumental variables and cross-lagged panel models

	Reading	Mathematics	Working memory	Self-control
(a) Anderson–Hsiao first-difference IV, Δy_{i3} on Δy_{i2} and Δx_{i3}				
TV strictly exogenous: TV (β)	−0.004 (0.007)	−0.005 (0.007)	0.002 (0.012)	0.024* (0.010)
TV strictly exogenous: lagged outcome (ρ)	0.245*** (0.054)	0.198** (0.063)	0.147*** (0.029)	0.057 (0.039)
TV predetermined: TV (β)	0.003 (0.010)	−0.005 (0.012)	0.010 (0.018)	0.023 (0.018)
TV predetermined: lagged outcome (ρ)	0.254*** (0.054)	0.212*** (0.063)	0.130*** (0.027)	0.058 (0.040)
First-stage F, Δy_{i2} (predetermined)	318	247	786	659
First-stage F, Δx_{i3} (predetermined)	2004	2006	1945	1694
(b) Cross-lagged panel model (CLPM)				
$\text{TV}_{t-1} \rightarrow \text{outcome}_t$ (SD per hour)	−0.021*** (0.005)	−0.017*** (0.005)	−0.048*** (0.007)	−0.023** (0.008)
$\text{Outcome}_{t-1} \rightarrow \text{TV}_t$ (hours per SD)	−0.086*** (0.011)	−0.105*** (0.014)	−0.065*** (0.014)	−0.065*** (0.017)
Autoregression, outcome	0.825*** (0.013)	0.846*** (0.008)	0.508*** (0.013)	0.478*** (0.013)
Autoregression, TV	0.310*** (0.022)	0.305*** (0.021)	0.316*** (0.021)	0.319*** (0.026)
χ^2 (df), RMSEA	800 (8), 0.119	1038 (8), 0.136	956 (8), 0.131	950 (8), 0.144
(c) Random-intercept cross-lagged panel model (RI-CLPM)				
$\text{TV}_{t-1} \rightarrow \text{outcome}_t$ (SD per hour)	−0.005 (0.010)	0.012 (0.009)	−0.032* (0.014)	0.005 (0.017)
$\text{Outcome}_{t-1} \rightarrow \text{TV}_t$ (hours per SD)	−0.004 (0.037)	0.077 (0.042)	0.006 (0.030)	0.001 (0.029)
Autoregression, outcome	0.665*** (0.046)	0.431*** (0.094)	0.166*** (0.023)	0.092** (0.034)
Autoregression, TV	−0.003 (0.021)	−0.004 (0.020)	−0.004 (0.021)	−0.013 (0.022)
Correlation of random intercepts	−0.30	−0.34	−0.24	−0.16
χ^2 (df), RMSEA	162 (5), 0.067	99 (5), 0.052	33 (5), 0.029	22 (5), 0.025
Children	6 951	6 941	6 947	5 705

Note. Children observed in all three waves. Panel (a): weighted two-stage least squares on the wave 2–3 difference; instruments are y_{i1} and Δx_{i3} (strictly exogenous TV) or y_{i1}, x_{i1} and x_{i2} (predetermined TV). Panels (b)–(c): maximum likelihood on the weighted covariance matrix of the six repeated measures, with autoregressive and cross-lagged paths constrained equal over time and wave-specific innovation (co)variances; the RI-CLPM adds correlated random intercepts for TV and the outcome, so that its lagged paths describe within-child dynamics. Fit statistics treat the weighted covariance matrix as if from a simple random sample of the unweighted size and are indicative only. All standard errors from 80 JK2 replicates re-estimating each model. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-sided).

© 2026 NSP
Natural Sciences Publishing Cor.

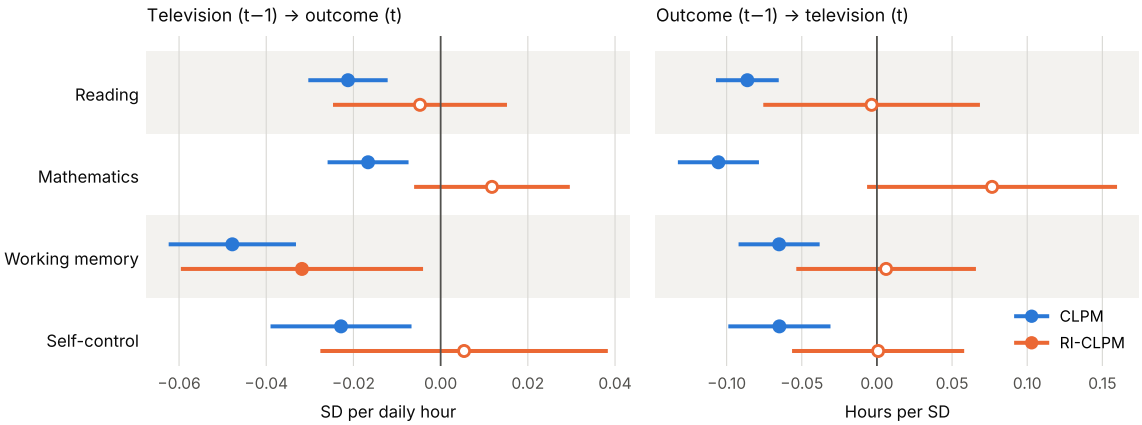

Fig. 3: Cross-lagged effects in both directions under the classical cross-lagged panel model (CLPM) and the random-intercept CLPM (RI-CLPM), with 95% confidence intervals from 80 JK2 replicates. Filled markers: $p < 0.05$; open markers: not significant.

5.4 Computer use and video games

Table 5 reports the other media. In kindergarten, 33.9% of children used a computer at least three times a week. Frequent computer use is associated with 0.159 SD higher second-grade reading and 0.132 SD higher mathematics after adjustment, and the association remains after controlling for kindergarten achievement: +0.083 SD (SE 0.019) for reading and +0.040 SD (SE 0.018) for mathematics, with no association for working memory. Because the computer item was asked in a single wave, these value-added estimates cannot be checked against fixed effects.

Video-game time at ages 7–8 is positively associated with reading and mathematics in pooled models (+0.044 and +0.033 SD per hour), but the associations disappear within children (+0.008 and −0.001 SD). As for television, the pooled associations of video games reflect differences between children, not changes within them.

Table 5: Computer use and video games: associations with learning outcomes (SD units)

	Reading	Mathematics	Working memory
(a) Computer use at age 6 (3+ times a week vs. less), outcome at age 8			
Adjusted	0.159*** (0.027)	0.132*** (0.025)	0.051 (0.031)
Value-added (age-6 outcome)	0.083*** (0.019)	0.040* (0.018)	0.011 (0.030)
(b) Video-game hours, ages 7–8 (television held constant)			
Pooled OLS, adjusted	0.044*** (0.012)	0.033* (0.015)	0.007 (0.013)
Child fixed effects	0.008 (0.008)	−0.001 (0.006)	−0.001 (0.015)

Note. Panel (a): single-wave exposure (the computer item changes wording between waves), so only adjusted and value-added estimates are possible; models also control for television hours at age 6. Panel (b): video games were recorded at ages 7 and 8 only, so the fixed-effects estimates rely on two waves. Weighted; JK2 standard errors. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-sided).

6 Discussion

The central finding is that the negative association between television viewing and early learning is mostly an association between children, not within children. Moving from an unadjusted pooled regression to child fixed effects reduces the estimated association with reading and mathematics by more than 90%, from more than 0.1 SD to about 0.01 SD per hour, and eliminates the associations with working memory and self-control. The magnitude that remains is small: even a child whose viewing increased by two hours a day would be expected to score about 0.02 SD lower in mathematics, a very small effect by conventional standards. This pattern is consistent with natural-experiment evidence that preschool television has no large negative effect on later test scores [5] and with the conclusion that average digital-media associations are small once analytic choices are made explicit [7].

© 2026 NSP
Natural Sciences Publishing Cor.

Two qualifications prevent a conclusion that television is harmless. First, heavy viewing is associated with a within-child penalty of 0.05–0.06 SD in reading and mathematics. This threshold pattern is compatible with displacement of reading, play and conversation when viewing becomes the dominant activity, and with the importance of content [16, 18]. Second, fixed effects remove only time-invariant confounders. Time-varying shocks, such as a family disruption that increases viewing and lowers achievement in the same year, would still bias the within estimates, and the falsification test cannot rule out a common trend for reading. Conversely, measurement error in parent-reported viewing biases the within estimates towards zero, and more so than the pooled estimates [37].

Measurement error. Suppose that reported viewing is $x_{it} = x^*_{it} + u_{it}$, with classical errors u_{it} of variance σ^2_u, uncorrelated with true viewing x^*_{it}, with the outcome and across waves. The pooled estimator is attenuated by the reliability $r = 1 - \sigma^2_u/\sigma^2_x$ of the reports, whereas the fixed-effects estimator converges to $\lambda_W \beta$, with

$$\lambda_W = 1 - \frac{\sigma^2_u \,(T_i - 1)/T_i}{\sigma^2_{x,W}}, \tag{9}$$

where $\sigma^2_{x,W}$ is the variance of observed viewing around each child's mean [37,27]. Because the within transformation removes the stable part of true viewing but not the error, $\lambda_W < r$: only 42% of the variance of observed viewing (net of wave means) lies within children in our sample. For a reliability of parent reports of $r = 0.8$, 0.7 or 0.6, equation (9) gives $\lambda_W = 0.69$, 0.53 or 0.38. Correcting the fixed-effects estimates accordingly yields at most -0.012, -0.015 or -0.021 SD per hour for reading and -0.015, -0.019 or -0.026 SD for mathematics. Even with the most pessimistic reliability, the corrected within-child associations remain less than half of the adjusted pooled associations (-0.049 and -0.056 SD), which would themselves increase if corrected for the same error. The change in question wording between waves 1 and 2 adds a non-classical component to u_{it} that is partly absorbed by the wave indicators; restricting the analysis to waves 2 and 3, which used the same item, gave similar results (Section 6, Limitations). The fixed-effects estimates are therefore best read as attenuated by measurement error but not by enough to change the conclusions, and the pooled estimates as upper bounds contaminated by selection.

The contrast with interactive media is instructive. Frequent computer use in kindergarten predicts higher reading gains to second grade, in line with arguments that active and meaningful use supports learning [20], but this association rests on a value-added model, which performed poorly for working memory, and could still reflect selection. Video games show the same pattern as television: a positive pooled association with reading and mathematics that vanishes within children. Screen time is not one exposure, but for none of the three media do the data support large effects in either direction.

Dynamics and reverse causation. The dynamic models add two results. First, the classical cross-lagged model would have led to a misleading conclusion: significant effects in both directions, with achievement predicting later viewing more strongly than viewing predicts later achievement. Once stable differences are separated out, both directions vanish for reading, mathematics and self-control, mirroring the static comparison between pooled and fixed-effects estimates and confirming the critique of the classical model [24,36]. Second, the Anderson–Hsiao estimates, which rely on different assumptions (serially uncorrelated errors and valid lagged instruments rather than a joint normal covariance structure), agree with the RI-CLPM for the academic outcomes. Working memory is the one outcome for which the evidence is mixed: the RI-CLPM and the value-added model find a small lagged association (-0.03 SD per hour), whereas the fixed-effects and Anderson–Hsiao estimates are close to zero. A delayed effect on working memory, which develops quickly at these ages, cannot be ruled out and deserves study with more waves. Three features of this outcome may explain why it behaves differently from reading and mathematics. First, it is measured with a single short task and is much less stable from year to year (lagged-outcome coefficients of 0.4–0.5 against 0.65–0.84 in the value-added models), so controls for past scores remove less of the stable differences between children and estimators that rely on them, such as the value-added model, are more exposed to residual confounding. Second, reading and mathematics are taught every day at school, and instruction may compensate for small home influences, whereas working memory is not an explicit object of teaching. Third, experimental studies show that the pacing and content of television affect young children's executive functions in the short run [18,19], a mechanism that could accumulate over a year for executive functions but would not directly affect academic knowledge. These explanations are hypotheses that the present data cannot test.

Methodological implications. The estimator ladder shows how strongly conclusions about screen time depend on modelling choices. Cross-sectional or pooled analyses, even when adjusted for socioeconomic status, attribute to screens differences that belong to families. Value-added models are an improvement for stable outcomes but not for noisier ones such as working memory. Fixed-effects and Mundlak specifications, which are routine in econometrics [27,10], and random-intercept cross-lagged models, increasingly advocated in developmental science [24,25,35], should become the default in longitudinal studies of children's media use; the classical cross-lagged model should not be used to assess the direction of effects. With complex survey data, replication methods make it straightforward to obtain valid standard errors for such estimators and for the differences between them.

Limitations. The television question changed between kindergarten (three periods of the day) and later waves (a single total), so part of the within-child variation between waves 1 and 2 may reflect question wording; restricting the

© 2026 NSP
Natural Sciences Publishing Cor.

fixed-effects analysis to waves 2 and 3, which used the same item, gives similarly small associations for reading and mathematics (−0.005 SD, SE 0.006, for both). With only three waves, the dynamic models are just identified or nearly so: the Anderson–Hsiao estimator rests on a single differenced equation, and the RI-CLPM requires time-invariant lagged effects. The change in the television item between waves 1 and 2 may also inflate within-child variation in viewing and attenuate dynamic estimates. Television, computer and video-game use were reported by parents and measure time, not content. The data were collected in 2011–2013, before tablets and smartphones became the dominant screens of young children [22]; the results speak to the methodological question and to television, not to today's mobile media. Finally, the sample is from one country.

7 Conclusion

Using a nationally representative panel of U.S. children aged 6 to 8 and design-based inference for five static and three dynamic estimators, we find that the negative association between television viewing and learning is almost entirely explained by stable differences between children. Within children, an extra hour of weekday television is associated with at most 0.01 SD lower reading and mathematics and with no change in working memory or self-control, although heavy viewing above three hours a day carries a small penalty. Dynamic models that allow for persistence and feedback reach the same conclusion and show that the apparent reciprocal effects of the classical cross-lagged model are an artefact of stable between-child differences, with the possible exception of a small delayed association with working memory. These findings do not show that screens are harmless, and they concern mostly television in the early 2010s. They do show that cross-sectional associations overstate the role of screen time by an order of magnitude, and that parents and policy makers should give at least as much weight to the content of children's media and to the family context in which it is used as to the number of hours. Confirming these conclusions for mobile devices requires recent panel data with repeated measures of both media use and learning, analysed with within-child estimators.

Acknowledgement

This study uses the public-use data file of the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011), produced by the National Center for Education Statistics (NCES), U.S. Department of Education. NCES bears no responsibility for the analyses and interpretations presented here.

Data availability

The ECLS-K:2011 kindergarten–fifth grade public-use data file is freely available from the NCES website [8]. The analysis code (Python) is provided as supplementary material.

Conflict of interest

The authors declare that there is no conflict regarding the publication of this paper.

References

[1] Y. Reid Chassiakos, J. Radesky, D. Christakis, M. A. Moreno, C. Cross and Council on Communications and Media, Children and adolescents and digital media, Pediatrics **138**(5), e20162593 (2016). https://doi.org/10.1542/peds.2016-2593

[2] M. Adelantado-Renau, D. Moliner-Urdiales, I. Cavero-Redondo, M. R. Beltran-Valls, V. Martínez-Vizcaíno and C. Álvarez-Bueno, Association between screen media use and academic performance among children and adolescents: A systematic review and meta-analysis, JAMA Pediatrics **173**(11), 1058 (2019). https://doi.org/10.1001/jamapediatrics.2019.3176

[3] S. Madigan, B. A. McArthur, C. Anhorn, R. Eirich and D. A. Christakis, Associations between screen use and child language skills: A systematic review and meta-analysis, JAMA Pediatrics **174**(7), 665 (2020). https://doi.org/10.1001/jamapediatrics.2020.0327

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

[5] M. Gentzkow and J. M. Shapiro, Preschool television viewing and adolescent test scores: Historical evidence from the Coleman study, Quarterly Journal of Economics **123**(1), 279–323 (2008). https://doi.org/10.1162/qjec.2008.123.1.279

© 2026 NSP
Natural Sciences Publishing Cor.

[6] M. S. Kearney and P. B. Levine, Early childhood education by television: Lessons from Sesame Street, American Economic Journal: Applied Economics **11**(1), 318–350 (2019). https://doi.org/10.1257/app.20170300
[7] A. Orben and A. K. Przybylski, The association between adolescent well-being and digital technology use, Nature Human Behaviour **3**(2), 173–182 (2019). https://doi.org/10.1038/s41562-018-0506-1
[8] K. Tourangeau, C. Nord, T. Lê, K. Wallner-Allen, N. Vaden-Kiernan, L. Blaker and M. Najarian, Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011): User's Manual for the ECLS-K:2011 Kindergarten–Fifth Grade Data File and Electronic Codebook, Public Version (NCES 2019-051), U.S. Department of Education, National Center for Education Statistics, Washington, DC, 2019. https://nces.ed.gov/pubs2019/2019051.pdf
[9] Y. Mundlak, On the pooling of time series and cross section data, Econometrica **46**(1), 69–85 (1978). https://doi.org/10.2307/1913646
[10] A. Bell, M. Fairbrother and K. Jones, Fixed and random effects models: making an informed choice, Quality & Quantity **53**(2), 1051–1074 (2019). https://doi.org/10.1007/s11135-018-0802-x
[11] F. J. Zimmerman and D. A. Christakis, Children's television viewing and cognitive outcomes: A longitudinal analysis of national data, Archives of Pediatrics & Adolescent Medicine **159**(7), 619 (2005). https://doi.org/10.1001/archpedi.159.7.619
[12] D. A. Christakis, F. J. Zimmerman, D. L. DiGiuseppe and C. A. McCarty, Early television exposure and subsequent attentional problems in children, Pediatrics **113**(4), 708–713 (2004). https://doi.org/10.1542/peds.113.4.708
[13] L. S. Pagani, C. Fitzpatrick, T. A. Barnett and E. Dubow, Prospective associations between early childhood television exposure and academic, psychosocial, and physical well-being by middle childhood, Archives of Pediatrics & Adolescent Medicine **164**(5) (2010). https://doi.org/10.1001/archpediatrics.2010.50
[14] S. Tomopoulos, B. P. Dreyer, S. Berkule, A. H. Fierman, C. Brockmeyer and A. L. Mendelsohn, Infant media exposure and toddler development, Archives of Pediatrics & Adolescent Medicine **164**(12) (2010). https://doi.org/10.1001/archpediatrics.2010.235
[15] S. Madigan, D. Browne, N. Racine, C. Mori and S. Tough, Association between screen time and children's performance on a developmental screening test, JAMA Pediatrics **173**(3), 244 (2019). https://doi.org/10.1001/jamapediatrics.2018.5056
[16] D. R. Anderson and K. Subrahmanyam, Digital screen media and cognitive development, Pediatrics **140**(Supplement 2), S57–S61 (2017). https://doi.org/10.1542/peds.2016-1758C
[17] K. Kostyrka-Allchorne, N. R. Cooper and A. Simpson, The relationship between television exposure and children's cognition and behaviour: A systematic review, Developmental Review **44**, 19–58 (2017). https://doi.org/10.1016/j.dr.2016.12.002
[18] A. S. Lillard and J. Peterson, The immediate impact of different types of television on young children's executive function, Pediatrics **128**(4), 644–649 (2011). https://doi.org/10.1542/peds.2010-1919
[19] B. Huber, M. Yeates, D. Meyer, L. Fleckhammer and J. Kaufman, The effects of screen media content on young children's executive functioning, Journal of Experimental Child Psychology **170**, 72–85 (2018). https://doi.org/10.1016/j.jecp.2018.01.006
[20] K. Hirsh-Pasek, J. M. Zosh, R. M. Golinkoff, J. H. Gray, M. B. Robb and J. Kaufman, Putting education in "educational" apps: Lessons from the science of learning, Psychological Science in the Public Interest **16**(1), 3–34 (2015). https://doi.org/10.1177/1529100615569721
[21] B. Sauce, M. Liebherr, N. Judd and T. Klingberg, The impact of digital media on children's intelligence while controlling for genetic differences in cognition and socioeconomic background, Scientific Reports **12**(1), 7720 (2022). https://doi.org/10.1038/s41598-022-11341-2
[22] S. Mann, A. Calvin, A. Lenhart and M. Robb, The 2025 Common Sense Census: Media Use by Kids Zero to Eight, Common Sense Media, San Francisco, 2025. https://doi.org/10.65077/augg2895
[23] P. E. Todd and K. I. Wolpin, On the specification and estimation of the production function for cognitive achievement, The Economic Journal **113**(485), F3–F33 (2003). https://doi.org/10.1111/1468-0297.00097
[24] E. L. Hamaker, R. M. Kuiper and R. P. P. P. Grasman, A critique of the cross-lagged panel model, Psychological Methods **20**(1), 102–116 (2015). https://doi.org/10.1037/a0038889
[25] D. Berry and M. T. Willoughby, On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse, Child Development **88**(4), 1186–1206 (2017). https://doi.org/10.1111/cdev.12660
[26] P. D. Allison, Fixed Effects Regression Models, SAGE, Thousand Oaks, CA, 2009. https://doi.org/10.4135/9781412993869
[27] J. M. Wooldridge, Econometric Analysis of Cross Section and Panel Data, 2nd ed., MIT Press, Cambridge, MA, 2010.
[28] K. F. Rust and J. N. K. Rao, Variance estimation for complex surveys using replication techniques, Statistical Methods in Medical Research **5**(3), 283–310 (1996). https://doi.org/10.1177/096228029600500305
[29] K. M. Wolter, Introduction to Variance Estimation, 2nd ed., Springer, New York, 2007. https://doi.org/10.1007/978-0-387-35099-8
[30] T. Lumley, Analysis of complex survey samples, Journal of Statistical Software **9**(8) (2004). https://doi.org/10.18637/jss.v009.i08
[31] G. Solon, S. J. Haider and J. M. Wooldridge, What are we weighting for?, Journal of Human Resources **50**(2), 301–316 (2015). https://doi.org/10.3368/jhr.50.2.301
[32] T. W. Anderson and C. Hsiao, Estimation of dynamic models with error components, Journal of the American Statistical Association **76**(375), 598–606 (1981). https://doi.org/10.1080/01621459.1981.10477691
[33] M. Arellano and S. Bond, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, The Review of Economic Studies **58**(2), 277 (1991). https://doi.org/10.2307/2297968
[34] S. Nickell, Biases in dynamic models with fixed effects, Econometrica **49**(6), 1417 (1981). https://doi.org/10.2307/1911408
[35] J. D. Mulder and E. L. Hamaker, Three extensions of the random intercept cross-lagged panel model, Structural Equation Modeling: A Multidisciplinary Journal **28**(4), 638–648 (2021). https://doi.org/10.1080/10705511.2020.1784738

© 2026 NSP
Natural Sciences Publishing Cor.

[36] R. E. Lucas, Why the cross-lagged panel model is almost never the right choice, Advances in Methods and Practices in Psychological Science **6**(1) (2023). https://doi.org/10.1177/25152459231158378

[37] Z. Griliches and J. A. Hausman, Errors in variables in panel data, Journal of Econometrics **31**(1), 93–118 (1986). https://doi.org/10.1016/0304-4076(86)90058-8

© 2026 NSP
Natural Sciences Publishing Cor.