Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Who Is Represented in School Acceleration Evidence? Selection, Comparisons and Social–Emotional Follow-Up

Alexander Kuznetsov

alex@5steps.edu.sg

Center for Educational & Psychological Assessment, 5 Steps Academy Private Limited
https://orcid.org/0009-0002-6652-3544

Systematic Review

Keywords: academic acceleration, grade skipping, early college entry, gifted education, social–emotional adjustment, selection bias, stratified meta-analysis

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11806746/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Who Is Represented in School Acceleration Evidence? Selection, Comparisons and Social–Emotional Follow-Up

Alexander Kuznetsov, Ph.D.
Center for Educational & Psychological Assessment, 5 Steps Academy Private Limited, Singapore | alex@5steps.edu.sg | ORCID: 0009-0002-6652-3544

Abstract

Schools weighing academic acceleration confront two separate choices: whom to place ahead and how to track peer adjustment afterward. Why do prior reviews report near-zero psychosocial averages? Unstratified pooling mixes incompatible selection gates, follow-up lags, and composite scales that cancel transition strain against survivor gains. Following PRISMA 2020, six-API queries, citation snowballing, and thesis repository searches reconciled 252 baseline records into 2,593 unique works and 73 verified full texts across 50 reports and 30 study families. Stratified REML and Knapp–Hartung models synthesized $K = 20$ non-overlapping cohorts ($N = 7{,}112$) alongside national ($N = 34{,}685$) and clinical ($N = 1{,}182$) selection cohorts. Naive pooling reproduced the familiar null ($g = -0.052$, $I^2 = 67.1\%$) while hiding negative adjustment in school grade skipping ($k = 8$, $g = -0.147$, $p = .044$) and immediate transitions ($k = 14$, $g = -0.123$, $p_{KH} = .027$) that reversed in adult/post-program waves ($k = 6$, $g = +0.123$; slope $b = +0.240$, $p_{KH} = .048$). In early-college cohorts, uniform Year-1 strain ($k = 6$, $g = -0.180$, $I^2 = 0.0\%$) flipped among alumni after 15.3%–23.3% attrition ($k = 2$, $g = +0.305$, $I^2 = 0.0\%$; $Q_{between} = 18.26$, $p < .0001$). Across $k = 10$ cohorts, global, academic, and safety subscales beat peer and romantic integration by $\Delta g = +0.683$ ($p_{KH} < .0001$) amid scalar non-invariance. Survivor surveys cannot replace transition monitoring.

Keywords: academic acceleration; grade skipping; early college entry; gifted education; social–emotional adjustment; selection bias; stratified meta-analysis

1. Introduction

1.1. Two Distinct School Decisions: Placement Selection Versus Transition Monitoring

A school weighing grade skipping faces two distinct decisions. Should this pupil move ahead now, and what social support will the transition require next term? Data suited to one question rarely settle the other. Surveying fifty-year-old adults who skipped a grade in childhood describes midlife satisfaction among responders (Bernstein et al., 2021; Lubinski et al., 2014), yet it cannot tell principals which current applicants to admit, how excluded candidates from lower-income homes or twice-exceptional profiles would fare (LeBeau et al., 2025; Wells et al., 2009), or whether a ten-year-old felt isolated during autumn term among older classmates (Gronostaj et al., 2016; Hoogeveen et al., 2009; Kretschmann et al., 2016).

Consensus hardened quickly around that reassuring narrative. Policy monographs and secondary syntheses have long promoted one reassuring verdict: academic acceleration delivers large achievement gains alongside near-zero or mildly positive social–emotional effects (Colangelo et al., 2004; Ferrándiz et al., 2026; Kulik, 2004; Rogers, 2004, 2015; Schuur et al., 2021; Steenbergen-Hu & Moon, 2011; Steenbergen-Hu et al., 2016). In the benchmark random-effects synthesis of psychosocial outcomes, Steenbergen-Hu and Moon (2011) pooled 15 social–emotional study estimates inside a 38-study corpus to obtain g = +0.076 (95% CI [−0.025, +0.176]) despite wide dispersion (Q = 63.26, p < .001). Later second-order reviews (Steenbergen-Hu et al., 2016) and umbrella syntheses (Ferrándiz et al., 2026) treated that near-zero grand mean as settled proof of psychosocial safety.

1.2. Why Unstratified Averages Mask Structural Heterogeneity

What does a pooled mean near zero actually show? Whenever primary studies diverge in sample entry gates, educational exposures, follow-up lags, and subscale aggregation, two rival mechanisms can generate a zero mean. Acceleration might truly be neutral everywhere, or opposite structural forces might cancel each other in an unstratified average. Four features of the primary evidence point straight to structural cancellation.

First, **selection regimes** sort very different youths into acceleration cohorts. Parent-driven talent searches targeting the top 1% to 0.01% of quantitative or verbal reasoning (Bernstein et al., 2021; Gross, 2006; Kopelman, 2016; Park et al., 2013; Stanley, 1973) recruit families actively pursuing above-level challenge; for exceptionally gifted children with IQ scores at or above 160, retention in an age-graded classroom imposes severe social alienation (Gross, 2004, 2006). Regular primary and secondary school grade-skippers face a harsher peer ecology. They enter intact older classroom hierarchies alone without an institutional cohort of age-mates (Guignard et al., 2021; Hoogeveen et al., 2009, 2012; Kretschmann et al., 2016; Motamedi et al., 2022; Sayler & Brookshire, 1993). Competitive residential early-college academies (such as the Texas Academy of Mathematics and Science (TAMS; Boazman, 2010; Caplan et al., 2002; Hoggan, 2008; Stride, 1996), the University of Washington Early Entrance Program (UW EEP; Hertzog & Chung, 2015; Noble et al., 2007; Robinson & Janos, 1986), the Program for the Exceptionally Gifted (PEG; Cornell et al., 1991), and the University of Iowa National Academy of Arts, Sciences, and Engineering NAASE; Shepard et al., 2009; Wu, 2018) screen applicants on entrance exams, essays, interviews, and family readiness while staffing dedicated dormitory counselors.

Second, **measurement timing** mixes acute transition shock with post-attrition survivorship. Assessments collected during the opening semester or first year of grade skipping or early university entry capture immediate reference-group shifts: Big-Fish-Little-Pond drops in academic self-concept, homesickness, and peer status loss (Boazman, 2010; Hoggan, 2008; Kretschmann et al., 2016; Shepard et al., 2009; Singh et al., 2021). Surveys collected 5 to 35 years later sample a filtered survivor pool. They observe adults after early-program dropouts have left (Boazman, 2006; Caplan, 1996; Cornell et al., 1991) or restrict enrollment to adult educational winners, such as 25-year-olds admitted to elite STEM doctoral programs (Bernstein et al., 2021, Study 2; Wai et al., 2010).

Operational definitions diverge just as sharply. Third, **exposure operationalization** ranges from discrete whole-grade skipping in neighborhood schools, to full-time early university matriculation, to composite 'educational dose' tallies that sum Advanced Placement (AP) exams, dual-enrollment classes, summer talent camps, and grade skips into one continuous index (Bernstein et al., 2021; Kopelman, 2016; Wai et al., 2010). Taking an AP course beside same-age friends keeps peer networks intact, whereas skipping a grade or moving to a university residence hall severs them (Mokher et al., 2021; Smedsrud, 2018).

Fourth, **composite well-being and self-concept totals** routinely add satisfaction with campus safety, material living standards, and academic challenge directly onto peer popularity, same-sex friendships, and romantic dating (Boazman, 2006; Hoggan, 2008; Shepard et al., 2009). Suppose an early college entrant feels intellectually relieved and physically safe on campus while struggling to make friends or date older peers. Summing those items hides the social cost, and scalar non-invariance across groups distorts raw composite comparisons even further (Sayler et al., 2015).

1.3. Present Study Objectives and Methodological Guardrails

Selective narrative accounts of school acceleration often recycle a small canon of familiar papers without checking broader thematic registers or auditing sample-size conflicts between journal tables, online supplements, and underlying dissertations. We report a full PRISMA 2020 systematic review and stratified meta-analysis across six open bibliographic APIs, citation graphs, university thesis archives, and publisher full texts (N = 73 verified first-hand PDFs covering 50 substantive reports across 30 distinct study families). Four methodological guardrails anchor our design: (1) strict cohort-family linkage to block participant double-counting, (2) forensic preservation of primary-source discrepancies rather than silent smoothing, (3) intraclass cluster corrections paired with exact within-cohort subscale covariance equations, and (4) stratified random-effects synthesis across Selection Regime (Axis A), Measurement Timing (Axis B), and Acceleration Type (Axis C). Four research questions organize our inquiry:

- **RQ1 (Selection Funnel & Cohort Representation):** Who passes population and clinical gates into grade skipping or early college, how much attrition occurs before follow-up, and how do primary-source denominator or psychometric errors alter published claims?

- **RQ2 (Stratified Subgroup & Timing Moderation):** Does naive pooling conceal opposing psychosocial effects across selection regimes (talent-search top 1%, school/referral grade skipping, residential early college), measurement timing (immediate transition vs. adult/post-program follow-up), and acceleration operationalization?
- **RQ3 (Within-Cohort Multi-Domain Masking):** Inside the exact same accelerated cohorts, do global, academic, and safety subscales exceed peer, sociometric, and romantic adjustment once cross-domain covariance is modeled?
- **RQ4 (Decision-Rule Adequacy):** How well do published designs guide pre-placement candidate selection versus post-placement transition monitoring?

2. Methods

2.1. PRISMA 2020 Search Architecture and Resolution of the Baseline Register

Reporting follows the PRISMA 2020 statement for updated and expanded systematic reviews (Page et al., 2021). No prospective protocol registration is claimed. Figure 1 maps our two-track identification, eligibility checks, and cohort-family linkage workflow. Track 1 audited the baseline thematic register of $N = 252$ records (245 unique APA PsycINFO identifiers across four RIS files plus seven print and chapter items), which contained 18 core report versions and 114 on-topic candidates requiring full-text adjudication. Track 2 captured empirical studies through 8 November 2026 alongside constituent papers from earlier reviews (Kulik, 2004; Rogers, 2004, 2015; Schuur et al., 2021; Steenbergen-Hu & Moon, 2011) via programmatic queries across six scholarly APIs: OpenAlex, Crossref, Semantic Scholar, ERIC, Europe PMC/PubMed, and arXiv (Supplement S1, Section S1.1).

Boolean search strings crossed acceleration exposures (grade skipping, whole-grade acceleration, academic acceleration, early entrance to college, early university entry, radical acceleration, Texas Academy of Mathematics and Science) with psychosocial outcomes (social–emotional, socioemotional, self-concept, self-esteem, subjective well-being, peer connectedness, peer relations, sociometric, psychological adjustment, mental health). Backward reference and forward citation snowballing ran across all 18 baseline DOIs and prior syntheses, paired with direct harvests of five university thesis repositories: University of North Texas Digital Library, Iowa Research Online, McGill eScholarship, University of Waikato Research Commons, and Radboud Repository.

How did screening narrow those results? Across APIs and citation networks, 3,842 raw records were harvested, and normalized-DOI plus fuzzy title–year deduplication removed 1,249 duplicates to leave 2,593 unique works that absorbed and adjudicated all 114 baseline candidates. Title, abstract, and venue screening excluded 2,309 records (1,892 physics, vehicle kinematics, or biological aging homonyms plus 417 gifted-education papers without an acceleration exposure), which left 284 thematic acceleration reports for full-text audit. Open-access and thesis repository retrieval yielded 49 verified PDFs covering complete doctoral and master's dissertations for Caplan (1996), Hoggan (2008), Boazman (2006, 2010), Kopelman (2016), Wu (2018), Glazer (1982), Jett (2018), Easter (2019), and Dare (2013), plus Volume 2 of *A Nation Deceived* (Colangelo et al., 2004). Targeted subscription acquisition added 24 publisher full texts, online supplements, and book chapters that captured 100% of critical priority papers (Caplan et al., 2002; Guignard et al., 2021; Kretschmann et al., 2016; Hoogeveen et al., 2009; Lubinski et al., 2014 with its Online Supplemental Material and 2015 Corrigendum; Schuur et al., 2021; Steenbergen-Hu & Moon, 2011) and 100% of controlled historical benchmarks (Cornell et al., 1991 plus ERIC ED313821; Robinson & Janos, 1986; Sayler & Brookshire, 1993). Eight secondary or retrospective subscription items lacked publisher PDFs (PRISMA 2020 Item 16b; three were already superseded by full underlying dissertations or updated chapters in our archive; Section 5), two postwar matched school benchmarks came from Kulik's (2004, Table 3) quantitative tables (dropped in Sensitivity Model S6), and 207 thematic papers lacking empirical psychosocial data or duplicating included cohorts were excluded to yield 50 substantive reports across 30 study families.

PRISMA 2020 Stage	Two-Track Identification, Screening & Full-Text Verification	Deduplication, Exclusions & Synthesis Streams
1. Identification	Track 1 — Baseline Thematic Register: N = 252 records (APA PsycINFO RIS exports across 4 batches: n = 245; print/chapter/policy records: n = 7), comprising n = 18 substantive baseline core reports and n = 114 candidate records adjudicated at full text. Track 2 — Multi-API & Repository Harvest: N = 3,842 raw records retrieved across OpenAlex, Crossref, Semantic Scholar, ERIC, Europe PMC/PubMed, and arXiv, plus bidirectional snowballing on 18 seed DOIs and university thesis repositories.	Records Removed Before Screening: Cross-API normalized-DOI and fuzzy title–year duplicates removed: n = 1,249. Deduplicated scholarly corpus retained for screening: n = 2,593 unique records.
2. Title & Abstract Screening	Records Screened (Title, Abstract & Venue): n = 2,593 deduplicated scholarly records (absorbing and adjudicating 100% of the n = 114 candidate topic records from the baseline register).	Records Excluded at Screening: n = 2,309 total excluded, comprising off-domain homonyms in physics, vehicle kinematics, and biological aging (n = 1,892) and non-acceleration gifted-education papers (n = 417).
3. Full-Text Eligibility Audit	Thematic Acceleration Reports Sought for Full-Text Audit: n = 284 strictly on-topic records. Verified first-hand full-text documents audited: N = 73 total (n = 49 open-access and university thesis PDFs + n = 24 publisher full texts, online supplements, and book chapters; 100% of critical priority and benchmark targets verified).	Full-Text Retrieval & Exclusions: Unacquired secondary subscription items: n = 8 (three superseded by complete underlying theses/chapters in archive); historical matched school benchmarks verified via Kulik (2004, Table 3): n = 2 (excluded in Sensitivity Model S6); off-outcome, commentary, or redundant secondary summaries excluded: n = 207.
4. Included in Synthesis	Total Substantive Evidence Base: 50 substantive reports, dissertations, and supplements linked across 30 distinct study families / institutional programs under strict cohort-family unit separation (18 reconciled baseline core reports + 32 newly integrated full-text articles, dissertations, and national/clinical cohorts).	(1) Primary Stratified Meta-Analysis: K = 20 strictly non-overlapping cohorts (N = 7,112; Subgroups A1 k = 4, A2 k = 8, A3 k = 8; Timing B1 k = 14 vs. B2 k = 6; Type C1 k = 9, C2 k = 8, C3 k = 3; Multi-domain contrast k = 10). (2) Isolated Sensitivity Pool (k = 5): overlapping sibling waves (k = 4, Model S5), age-unmatched outlier (k = 1, Model S2), and 100% first-hand PDF pool (K = 18, Model S6). (3) Selection-Funnel & Qualitative Context: national NCES (N = 34,685), clinical referral (N = 1,182), internal predictors, alumni surveys, and qualitative/review audits.

Figure 1. PRISMA 2020 two-track flow diagram showing baseline register reconciliation (*N* = 252 records; *n* = 114 baseline candidates adjudicated), six-API and repository harvesting (*N* = 3,842 raw records; *n* = 2,593 deduplicated), first-hand full-text verification (*N* = 73 primary PDFs), and unit-of-analysis linkage across 50 substantive reports in 30 distinct study families.

2.2. Eligibility, Strict Separation of Units, and Source-Discrepancy Protocol

Eligibility rules were explicit. Two criteria governed inclusion in the systematic review: (a) empirical data on deliberate school grade skipping, early university entrance, or quantified multi-year academic acceleration among high-ability or gifted learners, paired with (b) quantitative or qualitative measurement of social–emotional adjustment, self-concept, peer connectedness, behavioral difficulties, subjective well-being, or selection and attrition gates. Fifty substantive reports, theses, and supplements met eligibility and mapped onto 30 distinct study families and institutional programs (Supplement S1, Table S1).

Guardrail 1 enforced strict unit separation. Citations, thesis chapters, longitudinal cohorts, and subscale denominators were never conflated, and at most one primary estimate per independent participant cohort entered our main quantitative pool (K = 20 independent cohorts; N = 7,112 analytic participants; Table 1). Four multi-paper linkages mattered most: (a) inside the Study of Mathematically Precocious Youth (SMPY Cohorts 1–3), Bernstein et al.'s (2021, Study 1; N = 1,506) 35-year midlife follow-up served as the primary estimate, earlier overlapping SMPY waves (Swiatek & Benbow, 1991, N = 214; Richardson & Benbow, 1990, N = 1,247; Pollins, 1983, N = 42) were held in sibling-wave Sensitivity Model S5, and Lubinski et al. (2014; sex-only tables on the 2012–2013 wave) plus Park et al. (2013; STEM publications only) supplied cohort context; (b) Bernstein et al.'s (2021, Study 2; N = 450) SMPY Cohort 5, recruited separately at age 25 from elite STEM doctoral programs in 1992, entered as an independent cohort; (c) inside Radboud University's Center for the Study of Giftedness (CBO), Hoogeveen et al. (2012; N = 203) entered the primary pool while Hoogeveen et al. (2009; N = 309 sociometric / N = 216 self-concept across 18 Dutch secondary schools) moved to Sensitivity Model S5 to avoid school-network overlap; and (d) inside TAMS and University of Iowa NAASE, non-overlapping entry classes (TAMS Fall 2006 in Hoggan, 2008; TAMS Fall 2008 in Boazman, 2010; TAMS 2001–2005 alumni in Boazman, 2006) stayed distinct in the primary pool and were collapsed into institutional clusters in Sensitivity Model S4 (K = 17).

Age confounding required a dedicated holdout check. One open-access survey by Abdulsalam (2026; N = 400 middle-basic pupils in Kogi State, Nigeria) rated younger grade-skipped children directly against older classmates on unadjusted teacher maturity items (g = −1.304). Comparing raw developmental maturity across unequal chronological ages confounds placement with biological growth, so Abdulsalam (2026) was held out of the primary pool a priori and tested in Sensitivity Model S2. Two postwar matched school studies extracted via Kulik (2004, Table 3; Justman, 1953; Pevec, 1965) entered the K = 20 primary pool for historical grade-skipping coverage and were removed in Sensitivity Model S6 (K = 18 strictly first-hand PDFs).

Guardrail 2 governed source anomalies alongside JBI result-level critical appraisal for cohort, cross-sectional, and qualitative studies (JBI, 2024a, 2024b, 2025). We rejected composite quality scores and never smoothed over printed contradictions among article prose, tables, degrees of freedom, and dissertations. Every discrepancy was logged forensically (Section 3.2; Table 2), analyzed at its verified table-fitted denominator in the primary specification, and shifted simultaneously to its competing value in Sensitivity Model S3.

2.3. Effect-Size Computation, Cluster and Covariance Adjustments, and Stratified Synthesis

Metric orientation stayed uniform throughout. Every quantitative psychosocial endpoint was placed onto small-sample bias-corrected Hedges' g (Hedges, 1981). Positive values (+g) always mark healthier social–emotional adjustment, self-concept, peer connectedness, or subjective well-being among accelerated learners (or higher acceleration dose), whereas negative values (−g) mark transition strain, distress, or weaker interpersonal integration (equations in Supplement S1, Section S1.3). Two-group post-test Cohen's d values were scaled by the exact gamma factor $J(df) \approx 1 - 3/(4 \cdot df - 1)$. Controlled longitudinal cohorts reporting baseline and follow-up means (such as Boazman, 2010) yielded baseline-adjusted difference-in-differences (gDiD) using change-score standard deviations $SDchange = \sqrt{(SDpre^2 + SDpost^2 - 2 \cdot rpre,post \cdot SDpre \cdot SDpost)}$. Whenever paired t ratios accompanied means and SDs (Shepard et al., 2009), we solved algebraically for exact rpre,post (r = .81); elsewhere we set r = .50 and checked bounds across $r \in \{.20, .35, .50\}$. Correlations (r) mapped onto $d = 2r/\sqrt{(1 - r^2)}$ with delta-method variance $Vd = 4Vr/(1 - r^2)^3$. In Robinson and Janos's (1986) four-group comparison of 24 UW EEP entrants against 23 National Merit

Scholars, 24 regular UW freshmen, and 27 equally qualified non-accelerating age-mates, shared-treatment covariance weights prevented triple-counting the focal group.

Guardrail 3 implemented two variance corrections in run_stratified_meta_analysis.py. First, intact-classroom or school-clustered studies lacking multilevel standard errors (Guignard et al., 2021; Hoogeveen et al., 2009, 2012; Kretschmann et al., 2016; Motamedi et al., 2022; Sayler & Brookshire, 1993) had sampling variances multiplied by DEFF = 1 + (m̄ − 1)ρ at school-level ICC ρ = .05 with bounds ρ ∈ {.00, .10} (Hedges, 2007). Second, multi-subscale cohorts ($m > 1$) were composited using within-cohort correlation $r_{cov} = .35$, anchored to Guignard et al.'s (2021, Table 4) path β = .35 between peer connectedness and school satisfaction and tested across $r_{cov} \in \{.20, .35, .50\}$. How did we test multi-domain masking across the $k = 10$ cohorts reporting both global/academic/safety subscales (Domain A) and peer/sociometric/romantic subscales (Domain B)? Separate subscale *p*-values were never compared; each cohort contributed its paired contrast $\Delta g_i = g_{i,A} - g_{i,B}$ with exact covariance-adjusted variance:

$$\mathrm{Var}(\Delta g_i) = \mathrm{Var}(g_i,\mathrm{A}) + \mathrm{Var}(g_i,\mathrm{B}) - 2 \cdot r_\mathrm{cross} \cdot \sqrt{[\mathrm{Var}(g_i,\mathrm{A}) \cdot \mathrm{Var}(g_i,\mathrm{B})]}$$

evaluated across $r_\mathrm{cross} \in \{.20, .35, .50\}$.

Naive pooling was never treated as substantive. Guardrail 4 replaced naive grand-mean pooling with stratified REML random-effects models (paired with DerSimonian–Laird [DL] checks), Hartung–Knapp–Sidik–Jonkman (Knapp–Hartung, KH; Knapp & Hartung, 2003) t-intervals, Cochran's Q, I^2, 95% prediction intervals (PI), and Egger's weighted asymmetry regression. Cohorts were partitioned across Subgroup A (Selection Regime: A1 Talent-Search Top 1%–0.01%, A2 Specialist Referral / School Grade Skipping, A3 Competitive Residential / Supported Early College), Subgroup B (Measurement Timing: B1 Immediate Transition / School-Age Trajectory vs. B2 Long-Term Adult / Post-Program Follow-Up), the A3 × B Early-College Phase Interaction, and Subgroup C (Operationalization: C1 School Grade Skipping, C2 Early College Entry, C3 Composite Acceleration Index). Mixed-effects REML meta-regressions (Models MR-1 to MR-3) and six structural sensitivity suites (Models S1 to S6) tested robustness across covariance/ICC grids, age confounding, disputed denominators, institutional clustering, earlier overlapping waves, and 100% first-hand PDF restriction.

Table 1. Primary Strictly Non-Overlapping Quantitative Synthesis Pool ($K = 20$ Cohorts, $N = 7{,}112$) and Isolated Sensitivity Cohorts ($k = 5$)

ID	Study & Country	Study Family	Analytic *N* (*n*_acc / *n*_ctrl)	Axes (A | B | C)	Comparator & Social–Emotional Outcome	Hedges' *g* [95% CI]
C01	Bernstein et al. (2021, Study 1), USA	SMPY_Cohorts_1_3	1,506 (1,506 / —)	A1 | B2 | C3	Within-SMPY dose; age-50 composite well-being (SWLS, Ryff, PANAS, CSE)	−0.120 [−0.221, −0.019]
C02	Gross (2006 / 2004), Australia	Australia_Gross_IQ160	50 (17 / 33)	A1 | B2 | C1	IQ ≥ 160 non-accelerants; Coopersmith SEI & young-adult social/life adjustment	+0.581 [−0.006, +1.168]
C03	Kopelman (2016), USA	Davidson_Young_Scholars	59 (59 / —)	A1 | B1 | C3	≥ 99.9th-percentile DYS dose; SLSS global + BMLSS school/domain satisfaction	+0.261 [−0.162, +0.684]
C04	Bernstein et al. (2021, Study 2), USA	SMPY_GradSTEM_C5	450 (450 / —)	A1 | B2 | C3	Elite STEM doctoral cohort (recruited age 25); age-50 well-being composite	+0.221 [+0.035, +0.407]
C05	Hoogeveen et al. (2012), Netherlands	Dutch_CBO_Radboud	203 (148 / 55)	A2 | B1 | C1	Non-accelerated gifted peers; self-concept, social diary & parent CBCL (DEFF)	+0.060 [−0.271, +0.391]
C06	Guignard et al. (2021), France	French_Secondary	388 (26 / 362)	A2 | B1 | C1	Unconfounded GS-only vs. Grades 6 & 10 peers; MSLSS & peer connectedness (DEFF)	−0.124 [−0.481, +0.234]
C07	Kretschmann et al. (2016), Germany	German_ELEMENT	1,971 (96 / 1,875)	A2 | B1 | C1	Full-PSM Grades 4–6 peers; school satisfaction, solidarity, peer relations & anxiety	−0.082 [−0.242, +0.078]

ID	Study & Country	Study Family	Analytic *N* (*n*_acc / *n*_ctrl)	Axes (A | B | C)	Comparator & Social–Emotional Outcome	Hedges' *g* [95% CI]
C08	Motamedi et al. (2022), Iran	Iran_Tehran_Motamedi	549 (165 / 384)	A2 | B1 | C1	Clustered elementary classmates; SDQ behavioral & peer adjustment (DEFF)	−0.455 [−0.706, −0.205]
C09	Glazer (1982), Canada	Canada_Montreal_HS	49 (25 / 24)	A2 | B2 | C1	1:1 matched graduates (≥ 2-yr skip); activities, friendships & social rating	−0.474 [−0.901, −0.047]
C10	Sayler & Brookshire (1993), USA	US_NELS88_National	708 (365 / 343)	A2 | B1 | C1	NELS:88 gifted-class 8th graders; locus, self-concept, peer popularity & behavior	+0.006 [−0.093, +0.105]
C11	Justman (1953), USA	US_NYC_Justman	150 (75 / 75)	A2 | B1 | C1	IQ/age-matched junior-high peers; California Test of Personality (excl. in S6)	−0.020 [−0.338, +0.299]
C12	Pevec (1965), USA	US_HS_Pevec	80 (40 / 40)	A2 | B1 | C1	Matched older high-school classmates; CTP personal/social adjustment (excl. in S6)	−0.357 [−0.794, +0.081]
C13	Hoggan (2008), USA	TAMS_UNT_Fall2006	77 (77 / —)	A3 | B1 | C2	Semester-1 pre–post PWI-A well-being (masks Relationships d = −0.34)	−0.196 [−0.452, +0.060]
C14	Boazman (2010) / Sayler et al. (2015), USA	TAMS_UNT_Fall2008	213 (122 / 91)	A3 | B1 | C2	Regular-age UNT Honors freshmen; 1-yr baseline-adjusted PWI-A DiD	−0.224 [−0.554, +0.105]
C15	Boazman (2006), USA	TAMS_Alumni_01_05	380 (78 / 302)	A3 | B2 | C2	Age-matched normative young adults; post-TAMS PWI-A (1–5 yrs post-exit)	+0.242 [−0.007, +0.491]
C16	Shepard et al. (2009), USA	Iowa_NAASE_2009	21 (21 / —)	A3 | B1 | C2	Year-1 pre–post Piers–Harris 2 Total (r = .81; Popularity g = −0.30)	−0.126 [−0.385, +0.133]
C17	Wu (2018) / Wu et al. (2022), USA	Iowa_NAASE_Alumni	76 (76 / —)	A3 | B2 | C2	Upper-year/alumni NAASE survey; perceived social, friendship & romantic impact	+0.385 [+0.107, +0.663]
C18	Singh et al. (2021), USA	US_EECP_Singh	33 (33 / —)	A3 | B1 | C2	Prospective 3-mo Semester-1 DASS-21 distress trajectory (reversed)	−0.375 [−0.706, −0.043]
C19	Robinson & Janos (1986), USA	UW_Robinson_EEP	98 (24 / 74)	A3 | B1 | C2	Covariance-combined Merit (n=23), UW (n=24) & qualified age-mates (n=27); CPI	−0.237 [−0.541, +0.066]
C20	Cornell et al. (1991), USA	MaryBaldwin_PEG	51 (33 / 18)	A3 | B1 | C2	Capable non-accelerated honors girls (Table 2 N=18); Year-1 CPI adjustment	+0.150 [−0.234, +0.535]
SIB-1	Swiatek & Benbow (1991), USA [Sens. S5]	SMPY_Cohorts_1_3	214 (107 / 107)	A1 | B2 | C2	Sex/ability-matched SMPY non-accelerants at age 23 (isolated from C01)	+0.049 [−0.218, +0.316]
SIB-2	Richardson & Benbow (1990), USA [Sens.]	SMPY_Cohorts_1_3	1,247 (1,247 / —)	A1 | B2 | C3	SMPY Cohort 1 acceleration dose at age 23 (isolated from C01)	+0.060 [−0.051, +0.171]
SIB-3	Pollins (1983), USA [Sens.]	SMPY_Cohorts_1_3	42 (21 / 21)	A1 | B1 | C1	Matched unaccelerated SMPY boys (5-yr CPI follow-up; isolated from C01)	+0.177 [−0.418, +0.771]
SIB-4	Hoogeveen et al. (2009), NLD [Sens. S5]	Dutch_CBO_Radboud	309 (44 / 265)	A2 | B1 | C1	18 Dutch secondary schools; SDQ-II (N=216) & sociometrics (N=309; DEFF)	−0.549 [−0.833, −0.264]
OUT-1	Abdulsalam (2026), Nigeria [Sens. S2]	Nigeria_KogiState	400 (200 / 200)	A2 | B1 | C1	Unmatched older middle-basic classmates (severe chronological-age confounding)	−1.304 [−1.551, −1.056]

Note. C01–C20 form the K = 20 non-overlapping primary pool (N = 7,112). SIB-1 to SIB-4 and OUT-1 are held out of primary pooling to prevent sample overlap or chronological-age confounding and are evaluated in Sensitivity Models S2 and S5. A1 = Talent-search top 1%–0.01%; A2 = Specialist referral / school cohorts; A3 = Residential / supported early college; B1 = Immediate / school-age trajectory; B2 = Adult / post-program follow-up; C1 = School grade skipping; C2 = Early college entry; C3 = Composite acceleration index.

3. Results

3.1. The Two-Stage Selection Funnel: Pre-Entry Filtering and Post-Entry Survivorship

Who actually enters acceleration studies, and who stays until follow-up? RQ1 must be answered before any group difference can be interpreted. Linking national probability surveys, clinical referral clinics, and residential early-college rosters exposes a steep two-stage selection funnel.

Stage 1a governs population access. Wells et al. (2009) examined N = 34,685 pupils across two nationally representative US longitudinal surveys: the National Education Longitudinal Study of 1988 (NELS:88; N = 21,475) and the Education Longitudinal Study of 2002 (ELS:2002; N = 13,210). Whole-grade skipping was uncommon and dropped from 1.4% in NELS:88 to 0.8% in ELS:2002. Net of standardized test scores (OR = 1.54–1.61 per SD), grade skipping favored higher family socioeconomic status (OR = 1.25–1.30 per SD, p < .01), female pupils (OR = 1.38–1.85), and private or Catholic schools (OR = 2.00–2.34). Census age-for-grade proxies require extra caution: Long (1974) showed that mobility-by-grade links flipped sign among children whose fathers lacked a high-school diploma because modal grade age mixes entry cutoffs and grade retention with deliberate skipping.

Stage 1b filters clinical referrals even more sharply. LeBeau et al. (2025) tracked placement among N = 1,182 K–12 pupils assessed at the Belin-Blank Center giftedness clinic, where 8.0% (n = 94) skipped a grade and 22.0% received subject acceleration. In multivariable Bayesian logistic regression (Model 3) adjusted for age and achievement, each Full-Scale IQ point raised grade-skipping odds by 9% (OR = 1.09, 95% credible interval [1.04, 1.15], or 3.64-fold per 15 IQ points), whereas modest family resources and neurodevelopmental diagnoses slashed placement odds: medium versus high SES gave OR = 0.21 (95% CrI [0.05, 0.97], −79%), low SES gave OR = 0.37 (−63%), ADHD gave OR = 0.30 (95% CrI [0.10, 0.87], −70%), autism spectrum disorder (ASD) gave OR = 0.55 (−45%), and specific learning disorder (SLD) gave OR = 0.60 (−40%). Twice-exceptional and working-class candidates are largely screened out before transition research even begins. Stanley's (1973) early SMPY reports likewise stressed above-level testing, academic guidance, and pupil eagerness; his upper-quartile College Board rule of thumb was a pragmatic heuristic rather than a validated psychosocial threshold, and early Saturday math counts (30 invited, 21 enrolled, 19 finishing nine weeks) reflected voluntary persistence.

Stage 2 reshapes long-term cohorts through post-entry attrition and collider conditioning. Look at the two waves inside Bernstein et al. (2021). Study 1 tracked youths recruited around age 13 via parent-initiated SMPY talent searches (N = 1,506 analyzed at age 50) and found a statistically significant negative correlation between pre-college acceleration dose and age-50 psychological well-being (r = −.06, p = .02; g = −0.120, 95% CI [−0.221, −0.019]). Study 2 sampled N = 478 adults at age 25 who had already cleared undergraduate hurdles into top-15 US STEM doctoral programs (N = 450 analyzed at age 50) and found an opposite-sign positive correlation (r = +.11, p = .02; g = +0.221, 95% CI [+0.035, +0.407]). Although framed narratively as a constructive replication dismissing 'fruitless concerns', formal contrast testing reveals sharp opposite-sign heterogeneity (Q(1) = 9.98, p = .0016, I^2 = 90.0%) driven by conditioning on an age-25 attainment collider (see also Noreen et al., 2025, on retrospective narratives; Park et al., 2013, on STEM publication timing across n = 363 matched SMPY skippers vs. 657 controls; and Wai et al., 2010, on composite STEM dose).

Residential early-college programs show the same survivor filter inside a single academic year. Who disappears before spring post-testing? In Boazman's (2010) controlled evaluation of the Fall 2008 TAMS class, 22 of 144 entrants (15.3%) left during Year 1 (10 for personal or emotional distress and 12 for academic dismissal) and were dropped listwise (n = 122), whereas 22 non-responding UNT Honors controls (n = 91) remained enrolled and were retained via multiple imputation. In Cornell et al.'s (1991) study of Mary Baldwin College's female PEG cohort, 10 of 43 entrants (23.3%: five dropouts and five incomplete follow-ups) were excluded from longitudinal CPI tables (n = 33), while the companion ERIC technical report (Cornell et al., 1989, ED313821) logged first-year depression, homesickness, and suicidal ideation among entrants. Caplan's (1996) underlying dissertation for Caplan et al. (2002) similarly recorded six first-semester exits among 180 TAMS freshmen plus 12 lost SACQ forms (N = 162), and multi-decade alumni surveys routinely rest on

33%–60% response rates (Hertzog & Chung, 2015; Noble et al., 2007; Richardson & Benbow, 1990; Swiatek & Benbow, 1991).

3.2. Forensic Audit of Primary-Source Discrepancies and Review Error Propagation

Primary texts contradicted secondary summaries repeatedly. Checking primary tables and dissertations side by side settled four major reporting conflicts across core cohorts: (1) In Guignard et al. (2021), the clash between n = 57 grade-skippers in the Abstract and n = 56 in Methods was resolved by Table 1's mutually exclusive Venn cells: 26 GS-only + 7 (GL+GS+D48) + 15 (GL+GS) + 8 (GS+D48) = 56 skippers (11.38% of N = 492) and confirmed 57 was an abstract typo. (2) Pairing Caplan et al. (2002, Gifted Child Quarterly) with Caplan's (1996) UNT dissertation reconstructed the exact sample cascade (N = 180 entrants → 6 first-semester exits → N = 174 GPA roster → F(5, 168) = 3.36 after two outlier drops in 1996 vs. F(5, 165) = 3.21 [N = 171] after three unnoted drops in 2002 → N = 162 SACQ roster after 12 lost forms). Page 129 of the 2002 article also copied verbatim the 1996 dissertation's suppressor error: Family Environment Scale Cohesion carried a negative partial weight (β = −.19, t = −1.91, p = .058) yet was described as a positive adjustment predictor from its zero-order r = +.21. (3) Lubinski et al.'s (2014) Online Supplemental Material (SOM-U) reported midlife SWLS, Ryff, PANAS, and core self-evaluation means solely by sex without acceleration splits (sharing Bernstein et al.'s [2021, Study 1] wave), while a 2015 Psychological Science Corrigendum revised the US patent base rate from '< 1%' to '1.1%'. (4) UNT's official deposit of Hoggan (2008) omitted cited Appendix B (page 166 jumps straight to References on page 167), printed 0–10 item-scale SDs (1.36, 1.12, 1.39) beside 0–100 percentage means (73.8, 76.0, 73.5) in Table 4, and listed +t = 4.245 next to d = −0.34 for Personal Relationships in Table 5.

Secondary reviews missed equally consequential table errors. In Sayler and Brookshire (1993), Methods (p. 152) listed *n* = 334 gifted-class controls (*N* = 1,022), Table 1 (p. 153) transposed digits to *n* = 343 and *n* = 344 (*N* = 1,031–1,032) while dropping the minus sign on minimum locus of control (2.450 above maximum 1.520), and ANOVAs used *df* = 2, 1009 (*N* = 1,012). Kulik (2004, Table 3) extracted Table 1's global self-concept while skipping Table 2 (p. 154), where accelerated eighth-graders reported lower perceived popularity ($\chi^2 = 34.4$, $p < .001$) and lower peer importance ($\chi^2 = 12.3$, $p = .015$). In Hoogeveen et al. (2009), Methods (p. 54) claimed *N* = 357 three-wave completers (53 + 303 = 356), whereas Table 2 (SDQ-II) analyzed *N* = 216 (31 vs. 185), Table 5 (sociometrics) analyzed *N* = 309 (44 vs. 265), and Table 6 misprinted partial $\eta^2 = .50$ for .050 (*F* = 15.98). In Cornell et al. (1991), Methods (p. 136) stated '13 ninth and tenth graders' as controls, yet Table 2 (p. 139) and *t*(49) / *F*(10, 40) proved $N_{ctrl} = 18$. In Sayler et al. (2015; *N* = 533), multigroup CFA proved that the Personal Wellbeing Index–Adult (PWI-A) failed scalar invariance between TAMS freshmen and UNT Honors peers (ΔCFI = −.097, RMSEA = .167) on four of seven items (Standard of Living, Health, Personal Safety, Community Connectedness). Audit of Schuur et al.'s (2021) review uncovered direct error propagation: Table 6 (p. 43) copied a prose sign typo from page 88 of Boazman (2010) ('bad mood, *r* = .38' rather than Table 4.2's *r* = −.38, *p* < .01), and Table 5 (p. 41) coded Hoggan (2008) as 'No change' across PWI domains despite a drop in Personal Relationships (*d* = −0.34, *p* = .004).

Table 2. Summary Forensic Audit of 15 Primary-Source and Secondary-Review Discrepancies Explicitly Preserved in Synthesis (Guardrail 2)

#	Source Report(s)	Verified Primary Discrepancy or Statistical Anomaly	Handling in Synthesis & Sens. S3
1	Caplan (1996) vs. Caplan et al. (2002)	Tables *N* = 162 vs. GPA *F*(5, 168) in 1996 (*N*=172) and *F*(5, 165) in 2002 (*N*=171); 6 dropouts + 12 lost SACQ forms; negative suppressor FES Cohesion (β = −.19) misdescribed as positive in both texts.	Excluded from *K*=20 (no external control); internal predictor role documented.
2	Guignard et al. (2021)	Abstract *n* = 57 vs. Methods *n* = 56 skippers; Table 1 Venn sum (26+7+15+8 = 56) proves 57 is a typo; p. 26 prints Kretschmann's DOI on Košir et al.	Primary uses unconfounded GS-only (*n*=26 vs. 362); D48 gifted control (*n*=30) tested in S3.

#	Source Report(s)	Verified Primary Discrepancy or Statistical Anomaly	Handling in Synthesis & Sens. S3
3	Hoggan (2008)	Cited Appendix B absent in UNT deposit; Table 4 mixes 0–10 *SD*s (1.12, 1.39) with 0–100 *M*s; Table 5 prints +*t* = 4.245 for *d* = −0.34; five non-nested *n*s (189 to 60).	Rescaled to 0–100 metric (*n*=77 paired); domain split (*n*=60) modeled in RQ3.
4	Boazman (2010) & Sayler et al. (2015)	22/144 (15.3%) TAMS Year-1 dropouts deleted listwise vs. 22/91 Honors imputed; Table 3.8 *M* = 18.41 mathematically impossible; p. 88 drops minus on *r* = −.38; PWI-A fails scalar invariance (ΔCFI = −.097).	Baseline-adjusted DiD (*g* = −0.224) in primary; raw Year-1 endpoint tested in S3.
5	Boazman (2006)	Table 5 prints unsigned *d* = .12 beside negative *t* = −0.969 for Personal Relationships (true *d* = −0.12).	Signed *d* = −0.12 restored in primary composite and domain contrast.
6	Hoogeveen et al. (2009, 2012)	2009 Methods p. 54 claims N = 357 (53+303 = 356), Table 2 N = 216, Table 5 N = 309, Table 6 misprints η_p² = .50 for .050; 2012 table ns vary across N = 203.	2012 (*N*=203) in primary; 2009 (*g* = −0.549) isolated in S5 & RQ3 (*k*=10).
7	Sayler & Brookshire (1993) vs. Kulik (2004)	Methods *n* = 334 (*N*=1,022) vs. Table 1 *n* = 343/344 vs. ANOVA df = 2, 1009 (*N*=1,012); dropped minus on min (2.450); Kulik omitted Table 2 lower peer popularity (*p* < .001).	All 4 domains pooled with DEFF (*n*=365 vs. 343); *n*=334 tested in S3; domain split in RQ3.
8	Cornell et al. (1991) & ERIC ED313821	Methods p. 136 prints '13' controls vs. Table 2 & *t*(49)/*F*(10,40) proving *N*_ctrl = 18; 10/43 (23.3%) PEG entrants omitted (5 dropouts); acute Year-1 depression in ERIC report.	Verified Table 2 *N* = 51 (33 vs. 18) in primary; text *n* = 13 DiD tested in S3.
9	Kretschmann et al. (2016)	Page 199 Likert anchor typo ('1 = strongly disagree to 4 = strongly disagree'); MICE + Full-PSM control *n* varies (1,709–2,349).	Pooled across all 4 socioemotional scales (*n*=96 vs. 1,875, DEFF); domain split in RQ3.
10	Singh et al. (2021)	Retrospective alumni arm (*n*=66) misapplies independent one-way ANOVA (df = 2, 195) to repeated recall; p. 5 prints both 9.8 and 10.5 for stress.	Only prospective 3-month DASS-21 cohort (*n*=33) included in primary pool.
11	Motamedi et al. (2022)	Abstract/grade sum *n* = 165 vs. Results sex sum 73+90 = 163; Table 1 prints error df = 6; Table 2 prints 0/003.	Table *n* = 165 (DEFF = 1.85) in primary; *n* = 163 perturbed in S3.
12	Glazer (1982)	Abstract claims *n* = 24 pairs 'differ little in general', whereas Table 12 sums to 25 vs. 24 with $\chi^2(3) = 22.02$, *p* < .001 worse social adjustment (*g* = −1.175).	Table 12 *n* = 25 vs. 24 in primary; *n* = 24 perturbed in S3; domain split in RQ3.
13	Hertzog & Chung (2015); Mammadov et al. (2018)	Hertzog Table 1 prints 80.0% four times for *n* = 1/119 (0.84%) and Table 4 *n* = 130 for 103; Mammadov Table 1 swaps Phase-2 male/female percentages.	Corrected denominators reported in alumni triangulation (43.3% romantic cost).
14	Schuur et al. (2021)	Table 6 copies Boazman's p. 88 sign typo ('bad mood, *r* = .38' instead of *r* = −.38, *p* < .01); Table 5 codes Hoggan (2008) as 'No change' despite Relationships *d* = −0.34 (*p* = .004).	Corrected primary statistics extracted directly from original UNT dissertations.
15	Wells et al. (2009); Rogers (2015); Lubinski (2014)	Wells Tables 3–4 print identical Step-1 and achievement-adjusted Step-2 binary logits; Rogers counts multi-comparator cohorts as independent studies; Lubinski 2015 Corrigendum fixes patent rate.	Each cohort linked once per study family; verified adjusted ORs & SOM-U used.

Note. Sensitivity Model S3 perturbs all disputed primary-source denominators and comparator specifications simultaneously (Section 3.5).

3.3. Stratified Subgroup Meta-Analysis and Mixed-Effects Meta-Regression (RQ2)

Table 3 and Figure 2 display primary REML and Knapp–Hartung estimates across $K = 20$ non-overlapping cohorts ($N = 7{,}112$). Naive pooling reproduced the classic null benchmark. Combining all 20 cohorts without regard to entry regime, follow-up lag, or subscale domain yielded $g = -0.052$ (SE = 0.055, 95% CI [−0.160, +0.057], KH 95% CI [−0.173, +0.069], $p_z = .350$, $p_{KH} = .383$; DL $g = -0.051$), near Steenbergen-Hu and Moon's (2011) $g = +0.076$. Homogeneity was soundly rejected ($Q(19) = 57.75$, $p = 8.75 \times 10^{-6}$, $I^2 = 67.1\%$, $\tau^2_{REML} = 0.0383$, $\tau = 0.196$), with a 95% prediction interval from −0.479 to +0.375. Egger's intercept was flat (−0.038, SE = 0.825, $t(18) = -0.05$, $p = .964$), which ruled out small-study asymmetry and pointed to structural moderation.

Subgroup A (Selection Regime). Partitioning cohorts by entry gate exposed sharp divergence. Regime A1 (Talent-Search / Extreme Top 1%–0.01%, $k = 4$) averaged $g = +0.157$ (SE = 0.138, 95% CI [−0.114, +0.427], $p_z = .256$, $I^2 = 81.2\%$), split internally between childhood talent-search entry in Bernstein et al. (2021, Study 1; $g = -0.120$), adult STEM-doctoral recruitment in Study 2 ($g = +0.221$), and strong gains among exceptionally gifted learners in Gross (2006; IQ ≥ 160, $g = +0.581$) and Kopelman (2016; ≥ 99.9th percentile Davidson Young Scholars, $g = +0.261$). What about ordinary schools? Regime A2 (Specialist Referral & School Grade-Skipping Cohorts, $k = 8$) showed a significant negative pooled effect: $g = -0.147$ (SE = 0.073, 95% CI [−0.291, −0.004], $p_z = .044$; KH 95% CI [−0.321, +0.026], $p_{KH} = .084$; $I^2 = 59.3\%$, $\tau^2 = 0.0224$). Across regular primary and secondary classrooms in France ($g = -0.124$), Germany ($g = -0.082$), Iran ($g = -0.455$), Canada ($g = -0.474$), the Netherlands ($g = +0.060$ in 2012; $g = -0.549$ in 2009), and the United States ($g = +0.006$ in NELS:88; −0.020 in Justman; −0.357 in Pevec), pupils entering older classroom peer groups faced modest net psychosocial strain consistent with qualitative reports of post-skip support needs (Gronostaj et al., 2016), student concept maps of peer loss and bullying (Aboud et al., 2026; Dare, 2013; Dare et al., 2019), and teacher worries over isolation (Hoogeveen et al., 2005). Regime A3 (Competitive Residential & Supported Early-College Programs, $k = 8$) looked flat when averaged across waves ($g = -0.044$, SE = 0.096, 95% CI [−0.233, +0.145], $p_z = .650$, $I^2 = 70.1\%$) only because Year-1 drops and alumni rebounds canceled out.

Subgroup B (Measurement Timing) and the Early-College Reversal. Stratifying all 20 cohorts by follow-up horizon accounted for a major share of dispersion. Across B1: Immediate Transition and School-Age Trajectories ($k = 14$), accelerated pupils showed significant negative adjustment under both normal-z and Knapp–Hartung tests: $g = -0.123$ (SE = 0.049, 95% CI [−0.218, −0.027], $p_z = .0116$; KH 95% CI [−0.229, −0.016], $p_{KH} = .0271$; $I^2 = 44.6\%$, $\tau^2 = 0.0123$). Across B2: Long-Term Adult and Post-Program Follow-Ups ($k = 6$), the mean flipped positive to $g = +0.123$ (SE = 0.132, 95% CI [−0.135, +0.382], $p_z = .349$, $I^2 = 83.4\%$). In mixed-effects meta-regression Model MR-1, follow-up horizon explained $R^2_{analog} = 27.2\%$ of between-study variance (intercept $\beta_0 = -0.126$, $p_z = .0359$; adult slope $\beta_1 = +0.240$, SE = 0.108, 95% CI [+0.028, +0.451], $p_z = .0261$; KH 95% CI [+0.002, +0.477], $p_{KH} = .0481$) and left residual heterogeneity non-significant ($Q_{resid}(18) = 19.81$, $p = .344$).

Inside Regime A3 (Early-College Programs), that sign flip was complete. Every cohort tested during the first semester or first year of early university entry (A3 × B1, $k = 6$) shared a homogeneous negative effect: $g = -0.180$ (SE = 0.063, 95% CI [−0.303, −0.057], $p_z = .0041$; KH 95% CI [−0.341, −0.019], $p_{KH} = .0350$; $Q(5) = 4.55$, $p = .473$, $I^2 = 0.0\%$), right alongside Rogers's (2015, Table 2, p. 26) −0.27 benchmark for residential campus high schools. Alumni survivor surveys flipped the sign completely. Both post-program early-college studies (A3 × B2, $k = 2$: Boazman, 2006; Wu, 2018) shared a homogeneous positive mean: $g = +0.305$ (SE = 0.095, 95% CI [+0.120, +0.491], $p_z = .0013$; $Q(1) = 0.57$, $p = .452$, $I^2 = 0.0\%$). Between-phase divergence hit $Q_{between}(1) = 18.26$, $p = .000019$ and explained 100% of variance inside Regime A3.

Subgroup C (Acceleration Operationalization). Discrete school grade skipping (C1, $k = 9$: $g = -0.115$, 95% CI [−0.269, +0.039]) and early university matriculation (C2, $k = 8$: $g = -0.044$, 95% CI [−0.233, +0.145]) both carried negative point estimates. Composite Acceleration Dose Indices (C3, $k = 3$: $g = +0.085$, 95% CI [−0.177, +0.346]) sat $\Delta g = +0.200$ higher than school grade skipping because dose tallies mix AP/IB classes and summer camps (where age-peer ties stay intact) with grade jumps. Joint meta-regressions pairing regime with timing (MR-2: Regime A2 immediate-transition intercept $\beta_0 = -0.170$, $p_z = .0498$) and operationalization

with timing (MR-3: grade-skipping immediate-transition intercept $\beta_0 = -0.146$, $p_z = .0826$) documented collinearity between entry gate and follow-up lag in published studies (Supplement S1, Table S2).

Table 3. Stratified Random-Effects Meta-Analysis (REML and Knapp–Hartung) Across Structural Axes A, B, A3 × B, and C ($K = 20$ Independent Cohorts, $N = 7{,}112$)

Analytic Stratum / Subgroup	*k*	REML *g*	*SE*	95% CI (Normal *z*)	95% CI (Knapp–Hartung)	*p_z* (*p*_KH)	I^2 (τ^2)
Naive Unstratified Pool (Masking Benchmark)	20	−0.052	0.055	[−0.160, +0.057]	[−0.173, +0.069]	.350 (.383)	67.1% (.038)
Axis A: Selection Regime (Q_B(2) = 3.89, p = .143; S5 p = .0005)							
A1: Talent-Search / Extreme Top 1%–0.01%	4	+0.157	0.138	[−0.114, +0.427]	[−0.282, +0.596]	.256 (.338)	81.2% (.051)
A2: Specialist Referral & School Cohorts	8	**−0.147**	0.073	**[−0.291, −0.004]**	[−0.321, +0.026]	**.044** (.084)	59.3% (.022)
A3: Competitive Residential / Early College	8	−0.044	0.096	[−0.233, +0.145]	[−0.272, +0.184]	.650 (.663)	70.1% (.051)
Axis B: Measurement Timing (MR-1 slope $\beta_1 = +0.240$, p_z = .026, p_KH = .048)							
B1: Immediate / School-Age Trajectory	14	**−0.123**	0.049	**[−0.218, −0.027]**	**[−0.229, −0.016]**	**.012 (.027)**	44.6% (.012)
B2: Adult / Post-Program Follow-Up	6	+0.123	0.132	[−0.135, +0.382]	[−0.238, +0.485]	.349 (.420)	83.4% (.080)
Interaction A3 × B: Early College by Phase (Q_B(1) = 18.26, p = .000019)							
A3 × B1: Semester-1 / Year-1 Transition	6	**−0.180**	0.063	**[−0.303, −0.057]**	**[−0.341, −0.019]**	**.004 (.035)**	**0.0%** (.000)
A3 × B2: Post-Program Alumni Survivors	2	**+0.305**	0.095	**[+0.120, +0.491]**	[−0.898, +1.509]	**.001** (.191)	**0.0%** (.000)
Axis C: Acceleration Operationalization (Q_B(2) = 1.69, p = .429)							
C1: School Grade Skipping	9	−0.115	0.079	[−0.269, +0.039]	[−0.318, +0.088]	.143 (.227)	63.5% (.031)
C2: Early College Entry	8	−0.044	0.096	[−0.233, +0.145]	[−0.272, +0.184]	.650 (.663)	70.1% (.051)
C3: Composite Acceleration Dose Index	3	+0.085	0.133	[−0.177, +0.346]	[−0.489, +0.658]	.526 (.591)	83.2% (.039)

Note. Primary models use $r_{cov} = .35$ and school cluster ICC ρ = .05. Naive unstratified 95% prediction interval = [−0.479, +0.375]; Egger intercept = −0.038 (SE = 0.825, $t(18) = -0.05$, $p = .964$). Boldface marks statistically significant strata ($p < .05$).

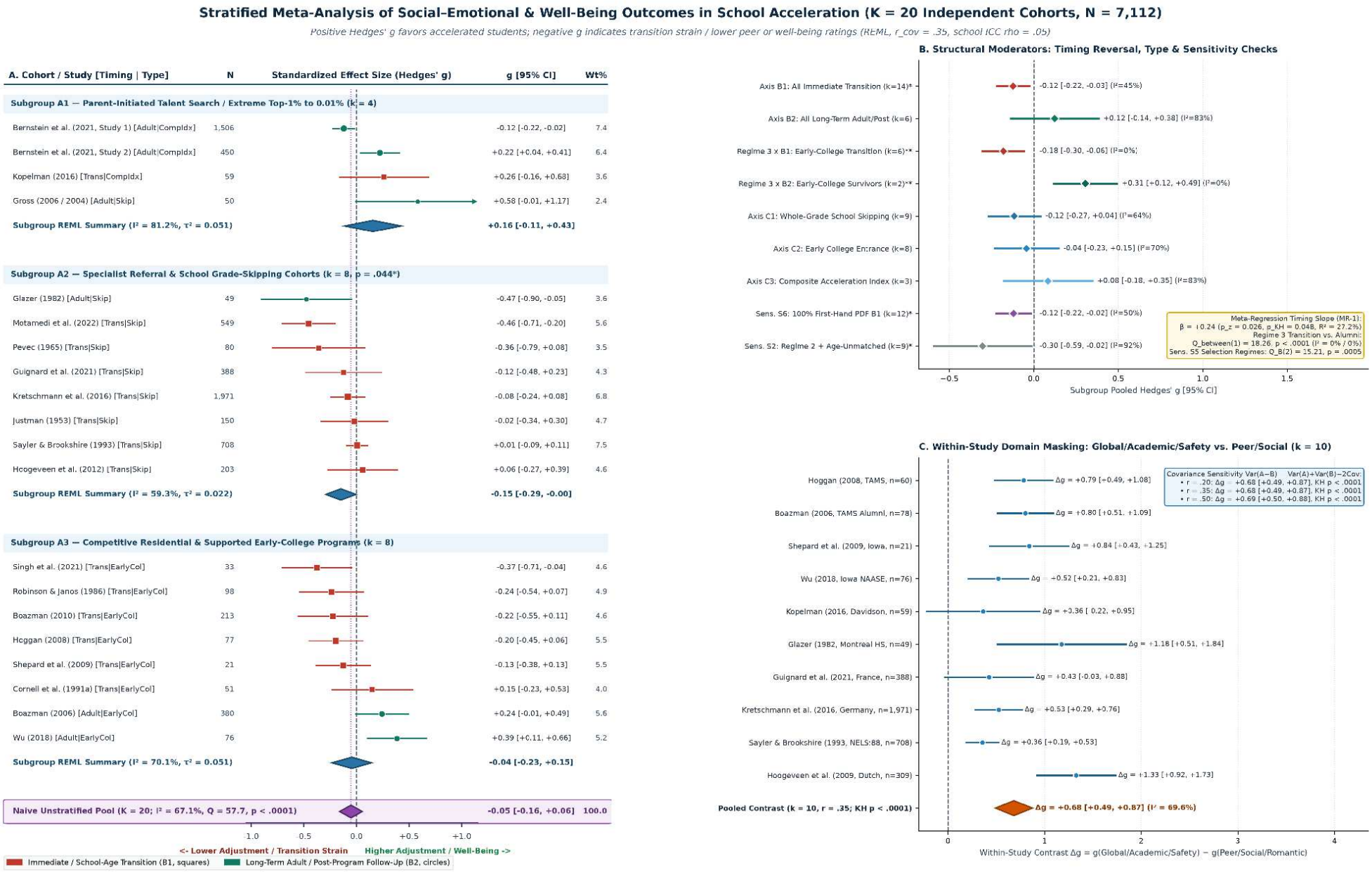


Figure 2. Stratified forest plot of social–emotional and well-being outcomes in academic acceleration (K = 20 independent cohorts, N = 7,112). Panel A: Cohort Hedges' g and REML subgroup summaries across Selection Regime (Subgroups A1–A3) alongside the naive masking benchmark. Panel B: Pooled estimates across Measurement Timing (B1 vs. B2), the Early-College Transition-to-Survivor reversal (A3 × B1 vs. A3 × B2), Operationalization (C1–C3), and Sensitivity Models S6 and S2. Panel C: Within-cohort covariance-adjusted multi-domain contrasts (Δg = Global/Academic/Safety minus Peer/Sociometric/Romantic) across k = 10 cohorts in five countries.

3.4. Within-Cohort Multi-Domain Masking Across Five Countries (RQ3)

Does composite scoring hide peer strain inside the exact same students? Table 4 tests RQ3 free of between-study confounding across all $k = 10$ independent first-hand cohorts (United States, Canada, France, Germany, and the Netherlands) that administered disaggregated subscales to the same participants. Ten out of ten cohorts showed a positive gap (exact binomial $p = .00098$). At empirical $r_{cross} = .35$, REML pooling yielded $\Delta g = +0.683$ (SE = 0.098, 95% CI [+0.491, +0.875], $p_z = 3.09 \times 10^{-12}$; KH 95% CI [+0.457, +0.909], $p_{KH} = .000077$; $Q(9) = 29.59$, $p = .0005$, $I^2 = 69.6\%$), with a 95% prediction interval [+0.073, +1.293] wholly above zero. Bounds at $r = .20$ ($\Delta g = +0.676$, KH 95% CI [+0.453, +0.899], $p_{KH} = .000075$) and $r = .50$ ($\Delta g = +0.690$, KH 95% CI [+0.460, +0.919], $p_{KH} = .000079$) were virtually identical.

Cohort-level contrasts show why summed scale totals misread peer adaptation. In German primary-to-secondary grade skipping (Kretschmann et al., 2016; $N = 1{,}971$), propensity-score matching uncovered weaker peer relations across Grades 4–6 ($g = -0.337$; girls $d = -0.35$, $p < .05$; Grade-6 boys $d = -0.47$, $p < .01$) offset by lower school anxiety among boys ($g = +0.188$; $\Delta g = +0.525$). In Dutch secondary schools (Hoogeveen et al., 2009; $N = 309$), higher academic and math self-concept ($g = +0.546$) sat beside steep deficits in sociometric status and same-sex social self-concept ($g = -0.781$; $\Delta g = +1.327$). In US eighth-grade data (Sayler & Brookshire, 1993; $N = 708$), positive behavior and locus-of-control scores ($g = +0.096$) masked lower perceived popularity and peer importance ($g = -0.261$; $\Delta g = +0.357$). Residential early-college freshmen gained campus safety and academic stimulation while losing ground in personal and romantic ties, consistent with Hertzog and Chung's (2015, Table 4) 43.3% rate of detrimental dating impact among UW EEP alumni, Wu's (2018) 28.9% detrimental romantic rate at Iowa NAASE, Dai et al.'s (2015) BFLPE trajectories at China's USTC Special Class for the Gifted Young, Australasian university reports (Easter, 2019; Kronborg, 2018), and Jett's (2018; Jett & Rinn, 2020) accounts of age stigma and 'passing as older' among lone radical accelerants.

Table 4. Within-Cohort Covariance-Adjusted Multi-Domain Contrasts (Δg = Global/Academic/Safety Minus Peer/Sociometric/Romantic) Across $k = 10$ Independent Cohorts ($r_{cross} = .35$)

Cohort (*N* Analytic)	Domain A: Global / Academic / Safety (*g*)	Domain B: Peer / Sociometric / Romantic (*g*)	Contrast Δ*g* [95% CI]
Hoggan (2008, TAMS, *n* = 60)	Personal Safety & Achievement (+0.445)	PWI-A Personal Relationships (−0.340)	+0.785 [+0.485, +1.085]
Boazman (2006, TAMS Alumni, *N* = 380)	Safety & Standard of Living (+0.680)	PWI-A Personal Relationships (−0.120)	+0.800 [+0.513, +1.087]
Shepard et al. (2009, Iowa, *n* = 21)	PH-2 Happiness & Satisfaction (+0.540)	PH-2 Popularity / Peer Status (−0.300)	+0.840 [+0.432, +1.248]
Wu (2018, Iowa NAASE, *n* = 76)	Perceived Intellectual Growth (+1.180)	Romantic & Peer Relationships (+0.660)	+0.520 [+0.209, +0.831]
Kopelman (2016, Davidson, *n* = 59)	BMLSS School Satisfaction (+0.443)	Global / Interpersonal Satisfaction (+0.079)	+0.364 [−0.223, +0.951]
Glazer (1982, Montreal HS, *N* = 49)	Extracurricular & Sports Participation (0.000)	Overall Peer Social Adjustment (−1.175)	+1.175 [+0.509, +1.841]
Guignard et al. (2021, France, *N* = 388)	MSLSS School Life Satisfaction (+0.089)	Peer Connectedness (−0.336)	+0.425 [−0.029, +0.879]
Kretschmann et al. (2016, Germany, *N* = 1,971)	Academic SC & Low Anxiety (+0.188)	Peer Relations Grades 4–6 (−0.337)	+0.525 [+0.286, +0.764]
Sayler & Brookshire (1993, US, *N* = 708)	Locus of Control & Behavior (+0.096)	Perceived Peer Popularity/Status (−0.261)	+0.357 [+0.189, +0.525]
Hoogeveen et al. (2009, Dutch, *N* = 309)	SDQ-II Math & School SC (+0.546)	Sociometric Preference & Social SC (−0.781)	+1.327 [+0.920, +1.734]

Cohort (*N* Analytic)	Domain A: Global / Academic / Safety (*g*)	Domain B: Peer / Sociometric / Romantic (*g*)	Contrast Δ*g* [95% CI]
REML Pooled Contrast (*k* = 10, *r* = .35)	**KH 95% CI [+0.457, +0.909], *p*_KH = .000077**	**95% Prediction Interval [+0.073, +1.293]**	**+0.683 [+0.491, +0.875]**

Note. Paired variance uses Var(A − B) = Var(A) + Var(B) − 2*r*√[Var(A)Var(B)]. Sensitivity at *r* = .20 (Δ*g* = +0.676, p_{KH} = .000075) and *r* = .50 (Δ*g* = +0.690, p_{KH} = .000079) confirms invariance.

3.5. Structural Sensitivity Suite (Models S1 to S6)

Six pre-specified sensitivity suites tested whether any conclusion rested on variance priors, disputed denominators, or secondary benchmarks (Supplement S1, Table S3):

- **Model S1 (Nine-Cell Covariance × School ICC Grid):** Across every pairing of $r_{cov} \in \{.20, .35, .50\}$ with school ICC ρ ∈ {.00, .05, .10}, immediate transition B1 stayed negative (*g* = −0.107 to −0.138), adult follow-up B2 stayed positive (*g* = +0.119 to +0.128), and regime rank A1 (+0.156 to +0.158) > A3 (−0.038 to −0.050) > A2 (−0.134 to −0.159) never changed.
- **Model S2 (Age-Confounded Comparison Check):** Adding Abdulsalam's (2026; *N* = 400, *g* = −1.304) age-unmatched Nigerian classroom ratings doubled Regime A2 strain from *g* = −0.147 (I^2 = 59.3%) to *g* = −0.304 (I^2 = 92.3%, *Q*(8) = 104.31), which proves why unstandardized maturity comparisons across unequal ages must stay quarantined.
- **Model S3 (Simultaneous Discrepancy Perturbation):** Swapping every disputed source specification at once (Guignard et al.'s high-ability D48 control [*n* = 30], Motamedi et al.'s sex-summed *n* = 163, Glazer's abstract *n* = 24, Sayler & Brookshire's prose *n* = 334, Cornell et al.'s prose *n* = 13 change-score DiD, and Boazman's [2010] raw Year-1 post-test) kept Regime A2 significantly negative (*g* = −0.159, 95% CI [−0.309, −0.009], p_z = .0383) and preserved the B1 (−0.106) versus B2 (+0.124) split.
- **Model S4 (Institutional-Family Clustering, *K* = 17):** Collapsing the three TAMS classes into one TAMS cluster and the two Iowa NAASE cohorts into one Iowa cluster shifted Regime A3 to *g* = −0.110 (I^2 = 31.9%), close to school grade skipping (−0.147).
- **Model S5 (Sibling-Wave Substitution):** Swapping Bernstein et al. (2021, Study 1) for Swiatek and Benbow's (1991) matched SMPY pairs and Hoogeveen et al. (2012) for Hoogeveen et al.'s (2009) longitudinal sociometric cohort sharpened Subgroup A divergence ($Q_{between}(2)$ = 15.21, *p* = .00050): Regime A1 = +0.199 (p_z = .0052, I^2 = 0.0%), Regime A2 = −0.229 (p_z = .0072, p_{KH} = .0312), and immediate transition B1 = −0.168 (p_z = .0028, p_{KH} = .0112; $Q_{between}(1)$ = 6.58, *p* = .0103).
- **Model S6 (100% First-Hand PDF Restriction, *K* = 18):** Dropping the two postwar benchmarks taken from Kulik (2004; Justman, 1953; Pevec, 1965) left every finding intact: overall *g* = −0.042 (I^2 = 69.5%), Regime A1 = +0.157, Regime A2 = −0.150, Regime A3 = −0.044, and immediate transition B1 = −0.120 (95% CI [−0.224, −0.017], p_z = .0226; KH 95% CI [−0.240, −0.001], p_{KH} = .0483).

4. Discussion

4.1. Reconciling Four Decades of Literature: Why 'No Average Harm' Misstates What Was Observed

Gifted-education policy has leaned for twenty years on near-zero meta-analytic means to declare psychosocial worries 'fruitless' (Bernstein et al., 2021; Colangelo et al., 2004; Steenbergen-Hu & Moon, 2011; Steenbergen-Hu et al., 2016). Our unstratified benchmark (*g* = −0.052, 95% CI [−0.160, +0.057]) replicates that grand mean almost to the second decimal place. Yet treating −0.052 as a universal safety constant misreads the data (*Q*(19) = 57.75, *p* < .0001, I^2 = 67.1%, 95% PI [−0.479, +0.375]). Three structural mechanisms create that mirage of neutrality.

Transition strain and survivor recovery carry opposite signs. Across all 14 cohorts measured during school-age skipping or Year-1 early college transition, psychosocial adjustment was significantly negative (*g* = −0.123, p_z = .012, p_{KH} = .027) and reached zero heterogeneity among first-year early-college entrants (*g* = −0.180, I^2 = 0.0%). Where did positive pooled means appear? Only in adult and post-program waves (*g* = +0.123 overall; *g* = +0.305, I^2 = 0.0% among early-college alumni), after 15.3% to 23.3% of early-college entrants had dropped out or been omitted (Boazman, 2010; Cornell et al., 1991) or after sampling conditioned

on age-25 admission to top-15 STEM doctoral programs (Bernstein et al., 2021, Study 2 vs. Study 1, $Q(1) = 9.98$, $p = .0016$). Thriving at age 50 among program completers does not erase peer strain at age 14.

Selection gates also invert baseline person–environment fit. For exceptionally gifted children in the top 0.01% (Gross, 2006; Kopelman, 2016; Thomas, 1987, 1993), forced age-grade retention inflicts acute boredom and peer rejection, so radical acceleration lifts social self-esteem substantially ($g = +0.581$ in Gross, 2006). Routine school grade skipping operates in a different social world. In specialist-referred and school-skipping cohorts (Subgroup A2: $g = −0.147$, $p = .044$; Model S5 $g = −0.229$, $p_{KH} = .031$), moderately to highly gifted pupils step into older adolescent hierarchies where pubertal timing, same-sex social self-concept, and sociometric standing put younger skippers at a disadvantage (Glazer, 1982; Hoogeveen et al., 2009; Kretschmann et al., 2016; Motamedi et al., 2022; Sayler & Brookshire, 1993). Guignard et al. (2021) add an important wrinkle: formal gifted labeling independently lowers adolescent peer connectedness ($\beta = −.13$, $p < .05$) beyond grade skipping itself.

Composite well-being instruments conceal interpersonal costs a second time. Across $k = 10$ cohorts in five countries, global, academic, and campus-safety subscales outscored peer, sociometric, and romantic adjustment by $\Delta g = +0.683$ ($p_{KH} < .0001$; 95% PI [+0.073, +1.293]). Worse still, Sayler et al. (2015) proved that PWI-A violates scalar invariance between early-college freshmen and regular-age honors peers ($\Delta CFI = −.097$), which invalidates raw composite mean comparisons.

4.2. Selection Equity and Post-Placement Monitoring: What Schools Can and Cannot Infer (RQ4)

How should school psychologists and administrators act on these boundaries? Table 5 maps our quantitative strata (RQ1–RQ3) directly onto the school's two practical choices: pre-placement selection and post-placement monitoring (RQ4).

Table 5. Decision-Relevant Questions for Interpreting and Applying School Acceleration Evidence

Decision Question	What the Synthesized Evidence Establishes (N = 50 Reports / 30 Families)	Information Still Required for a Validated Individual Rule
1. Who entered the evidence base?	National access is rare (0.8%–1.4%) and stratified by high SES and private schooling (Wells et al., 2009; N=34,685); clinical gatekeeping excludes 63%–79% of lower/medium-SES and 40%–70% of ADHD/ASD/SLD pupils (LeBeau et al., 2025; N=1,182).	Prospective tracking of applicants, offers, decliners, and underserved/2e candidates to estimate placement effects beyond privileged survivors.
2. What educational exposure differed?	Discrete school grade skipping (C1: $g = −0.115$), supported early university entry (C2: $g = −0.044$), and course-blended composites (C3: $g = +0.085$; Wai et al., 2010; Mokher et al., 2021) represent non-interchangeable interventions.	Explicit specification of a single placement contrast against a defined counterfactual (e.g., whole-grade skip vs. subject acceleration with age peers).
3. When and across which domains was adjustment measured?	Immediate transitions show consistent strain (B1: $g = −0.123$, $p = .012$; early-college Year 1: $g = −0.180$, $p = .004$, $I^2 = 0\%$), concentrated in peer popularity, sociometrics, and dating ($\Delta g = +0.683$, $p < .0001$ across $k = 10$ cohorts).	Disaggregated pre-placement baseline and scheduled multi-wave post-placement monitoring using scalar-invariant peer and affective instruments.
4. Does an internal correlation validate an admission cutoff?	First-week self-concept ($r = +.45$), low family conflict ($r = −.35$; Caplan, 1996/2002), and autonomous motivation (Mammadov et al., 2018) predict adjustment *among admitted entrants*, while 15%–23% exit in Year 1.	External out-of-sample validation in an unselected applicant pool demonstrating calibration, classification accuracy, and net benefit of an admission cutoff.

Decision Question	What the Synthesized Evidence Establishes (N = 50 Reports / 30 Families)	Information Still Required for a Validated Individual Rule
5. Was a post-placement monitoring & response rule tested?	Six of seven German interviewees needed post-skip support (Gronostaj et al., 2016); residential programs provide counseling (Shepard et al., 2009), whereas individual skippers lack institutional buffers (Jett & Rinn, 2020).	A pre-specified monitoring schedule, explicit clinical/peer-support trigger threshold, documented intervention, and comparative evaluation after action.

Note. Decision matrix linking stratified meta-analytic estimates (RQ1–RQ3) to result-level methodological appraisal (RQ4).

Two concrete policy lessons follow from Table 5. First, observational strain estimates neither justify barring twice-exceptional and lower-SES pupils nor justify holding back exceptionally gifted youth. Clinical and school gates systematically screen out children with ADHD, ASD, SLD, and modest family income (LeBeau et al., 2025; Wells et al., 2009), so published cohorts say almost nothing about how underrepresented learners fare when acceleration is paired with accommodations, whereas withholding acceleration from extreme high-IQ learners (Gross, 2006) inflicts documented social harm. Second, schools and early-college academies should retire blanket promises of 'zero social risk' and institute scheduled 6- to 12-month peer-connectedness checks. Regression weights inside admitted TAMS or UW classes (Caplan et al., 2002; Mammadov et al., 2018) describe variance among selected entrants rather than validating applicant cutoffs. Repeated group means in Hoggan (2008), Shepard et al. (2009), and Singh et al. (2021) pinpoint when transition stress peaks (months 3 to 6), yet no trial has validated an individual trigger threshold for mentoring or counseling, particularly for lone school skippers and solitary early university entrants lacking an age-peer cohort (Hoogeveen et al., 2009; Jett, 2018; Jett & Rinn, 2020; Kretschmann et al., 2016).

4.3. Methodological Standards for Future Acceleration Research and Synthesis

Four reporting reforms are overdue in educational placement research. Primary investigators should (a) publish full CONSORT/JBI participant flows from eligibility and offer through acceptance, Year-1 exit, and wave-specific denominators rather than deleting dropouts listwise (Boazman, 2010; Cornell et al., 1991); (b) test longitudinal and cross-group scalar invariance before summing composite scales (Sayler et al., 2015); (c) archive subscale means, SDs, and inter-scale correlation matrices for exact covariance synthesis; and (d) model classroom clustering via multilevel estimators or explicit ICC reporting. Meta-analysts must likewise audit primary tables directly instead of copying prose correlations or secondary tables that flipped signs in Schuur et al. (2021) and omitted peer-popularity deficits in Kulik (2004).

5. Limitations and Conclusion

Four constraints bound this synthesis. First, neither schools nor families randomize pupils to whole-grade skipping or early college; even propensity-score matching (Kretschmann et al., 2016; Park et al., 2013) and ability-matched controls (Robinson & Janos, 1986; Swiatek & Benbow, 1991) leave unmeasured baseline volition and maturity differences open. Second, while 73 first-hand PDFs (covering 100% of priority papers and controlled benchmarks) were audited directly, eight peripheral items from our 28-item subscription list lacked publisher PDFs; three of those eight were superseded by full dissertations or updated chapters in our archive (Boazman, 2006, for Boazman & Sayler, 2011; Wu, 2018, for Wu et al., 2022; Rogers, 2015, for Rogers, 2007), and Sensitivity Model S6 confirmed that restricting K = 18 strictly to first-hand PDFs preserved every key result (g_{B1} = −0.120, p_{KH} = .0483). Third, cohort-family linkage prevented multi-paper duplication across SMPY, Radboud CBO, TAMS, and Iowa NAASE, though historical publications rarely archive individual ID crosswalks. Fourth, small cell counts (k = 2 to 4 in A1, C3, and A3 × B2) yield wide Knapp–Hartung t-intervals with k − 1 degrees of freedom, which calls for larger multi-site prospective cohorts.

Who enters school acceleration studies, together with when and how adjustment is measured, governs what the evidence shows. Naive pooling manufactures a near-zero mean (g = −0.052) by averaging immediate transition strain in regular schools (g = −0.147) and Year-1 early-college programs (g = −0.180) against positive outcomes among exceptionally gifted youth and post-attrition adult survivors (g = +0.123 to +0.305), while composite well-being totals hide a two-thirds-SD gap between academic/safety satisfaction and peer or

romantic integration ($\Delta g = +0.683$). Honoring those structural boundaries lets schools expand equitable access for underrepresented and twice-exceptional pupils while delivering targeted peer support across the first year of accelerated placement.

Availability of Data, Code, and Materials

Online Supplement S1 provides exact multi-API search strings, our 50-report / 30-study-family master register (Table S1), effect-size and covariance equations, complete meta-regression and sensitivity outputs (Tables S2–S3), and forensic source notes. Machine-readable effect-size tables (T05_EFFECT_SIZES_TABLE.csv), fitted model outputs (T05_META_ANALYSIS_RESULTS.json), and Python replication scripts (run_api_harvester.py, run_stratified_meta_analysis.py, generate_q1_figures.py) are archived with the submission materials. Copyrighted publisher PDFs are not redistributed.

Funding and Competing Interests

This study received no external grant funding. The author declares zero competing financial or non-financial interests.

Ethics Approval and Consent to Participate

Not applicable. This study is a PRISMA 2020 systematic review and secondary meta-analysis of published aggregate literature and involved no direct recruitment of human participants.

Copyright Notice

© 2026 Alexander Kuznetsov. Personal reading, archival download for private study, and scholarly citation are permitted. Redistribution, commercial reuse, and circulation of adapted versions require written permission from the copyright holder except where applicable preprint or journal publishing licences or law permit otherwise.

References

Abdulsalam, A. A. (2026). Holistic development of grade-skipped and non-grade-skipped pupils in Kogi State, Nigeria. *Kakanchi Multidisciplinary Journal of COEASU*, *1*(1), 1–17. https://doi.org/10.68071/e784k99g

Aboud, Y. Z., Helali, M. M., & AlAli, R. M. (2026). How Saudi students perceive academic acceleration: A group concept mapping study in inclusive classrooms. *International Journal of Learning, Teaching and Educational Research*, *25*(8), 95–123. https://doi.org/10.26803/ijlter.25.8.4

Bernstein, B. O., Lubinski, D., & Benbow, C. P. (2021). Academic acceleration in gifted youth and fruitless concerns regarding psychological well-being: A 35-year longitudinal study. *Journal of Educational Psychology*, *113*(4), 830–845. https://doi.org/10.1037/edu0000500

Boazman, J. K. (2006). *Well-being of gifted students following participation in an early-college-entrance program* [Master's thesis, University of North Texas]. UNT Digital Library. https://digital.library.unt.edu/ark:/67531/metadc5505/

Boazman, J. K. (2010). *Well-being and academic success in gifted college students: Early-college entrants and honors college students* [Doctoral dissertation, University of North Texas]. UNT Digital Library. https://doi.org/10.12794/metadc31526

Boazman, J. K., & Sayler, M. F. (2011). Personal well-being of gifted students following participation in an early college-entrance program. *Roeper Review*, *33*(2), 76–85. https://doi.org/10.1080/02783193.2011.554153

Caplan, S. M. (1996). *Family and self-concept factors contributing to the adjustment and achievement of early entrants* [Doctoral dissertation, University of North Texas]. UNT Digital Library. https://digital.library.unt.edu/ark:/67531/metadc279103/

Caplan, S. M., Henderson, C. E., Henderson, J., & Fleming, D. L. (2002). Socioemotional factors contributing to adjustment among early-entrance college students. *Gifted Child Quarterly*, *46*(2), 124–134. https://doi.org/10.1177/001698620204600205

Colangelo, N., Assouline, S. G., & Gross, M. U. M. (Eds.). (2004). *A nation deceived: How schools hold back America's brightest students* (Vols. 1–2). Connie Belin & Jacqueline N. Blank International Center for Gifted Education and Talent Development, University of Iowa.

Cornell, D. G., Callahan, C. M., & Loyd, B. H. (1989). *Adjustment of academically talented females in a secondary school acceleration program* (AEL Minigrant Report No. 11; ERIC Document Reproduction Service No. ED313821). Appalachia Educational Laboratory.

Cornell, D. G., Callahan, C. M., & Loyd, B. H. (1991). Personality growth of female early college entrants: A controlled, prospective study. *Gifted Child Quarterly*, *35*(3), 135–143. https://doi.org/10.1177/001698629103500305

Dai, D. Y., Steenbergen-Hu, S., & Zhou, Y. (2015). Cope and grow: A model of affective-social development of gifted students in an early college entrance program. *Gifted Child Quarterly*, *59*(2), 75–90. https://doi.org/10.1177/0016986214568719

Dare, L. (2013). *A structured conceptualization of reasons to accelerate: Words from the WISE* [Master's thesis, University of Western Ontario]. Scholarship@Western.

Dare, L., Nowicki, E. A., & Smith, S. (2019). On deciding to accelerate: High-ability students identify key considerations. *Gifted Child Quarterly*, *63*(3), 159–171. https://doi.org/10.1177/0016986219828073

Easter, A. M. D. (2019). *'Time on my side': Experiences of accelerated students who entered university early in Aotearoa New Zealand* [Doctoral dissertation, University of Waikato]. Waikato Research Commons.

Ferrándiz, C., García-Perales, R., Pons, R. M., Fernández-Vidal, M. C., Rojo-Guillamón, M. I., & Ferrando-Prieto, M. (2026). What works for inclusive gifted education? An umbrella review of acceleration, enrichment, and ability grouping. *Frontiers in Education, 11, Article 1925181.* https://doi.org/10.3389/feduc.2026.1925181

Glazer, S. P. (1982). *Peer relationships and major acceleration in high school* [Master's thesis, McGill University]. eScholarship@McGill.

Gronostaj, A., Werner, E., Bochow, E., & Vock, M. (2016). How to learn things at school you don't already know: Experiences of gifted grade-skippers in Germany. *Gifted Child Quarterly*, *60*(1), 31–46. https://doi.org/10.1177/0016986215609999

Gross, M. U. M. (2004). Radical acceleration. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived: How schools hold back America's brightest students* (Vol. 2, pp. 87–96). University of Iowa.

Gross, M. U. M. (2006). Exceptionally gifted children: Long-term outcomes of academic acceleration and nonacceleration. *Journal for the Education of the Gifted*, *29*(4), 404–429. https://doi.org/10.4219/jeg-2006-247

Guignard, J.-H., Bacro, F., & Guimard, P. (2021). School life satisfaction and peer connectedness of intellectually gifted adolescents in France: Is there a labeling effect? *New Directions for Child and Adolescent Development*, *2021*(179), 127–142. https://doi.org/10.1002/cad.20448

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*(4), 341–370. https://doi.org/10.3102/1076998606298043

Hertzog, N. B., & Chung, R. U. (2015). Outcomes for students on a fast track to college: Early college entrance programs at the University of Washington. *Roeper Review*, *37*(1), 39–49. https://doi.org/10.1080/02783193.2014.976324

Hoggan, B. (2008). *Success factors among early college entrants* [Doctoral dissertation, University of North Texas]. UNT Digital Library. https://doi.org/10.12794/metadc9091

Hoogeveen, L., van Hell, J. G., & Verhoeven, L. (2005). Teacher attitudes toward academic acceleration and accelerated students in the Netherlands. *Journal for the Education of the Gifted*, *29*(1), 30–59. https://doi.org/10.1177/016235320502900103

Hoogeveen, L., van Hell, J. G., & Verhoeven, L. (2009). Self-concept and social status of accelerated and nonaccelerated students in the first 2 years of secondary school in the Netherlands. *Gifted Child Quarterly*, *53*(1), 50–67. https://doi.org/10.1177/0016986208326556

Hoogeveen, L., van Hell, J. G., & Verhoeven, L. (2012). Social–emotional characteristics of gifted accelerated and non-accelerated students in the Netherlands. *British Journal of Educational Psychology*, *82*(4), 585–605. https://doi.org/10.1111/j.2044-8279.2011.02047.x

JBI. (2024a). Appendix 7.1: Critical appraisal checklist for cohort studies. *JBI Manual for Evidence Synthesis* (26 March version). https://jbi-global.atlassian.net/wiki/spaces/MANUAL/pages/355599324/

JBI. (2024b). Appendix 7.5: Critical appraisal checklist for analytical cross sectional studies. *JBI Manual for Evidence Synthesis* (26 March version). https://jbi-global.atlassian.net/wiki/spaces/MANUAL/pages/355599424/

JBI. (2025). Section 3.3.11: Assessing methodological quality. *JBI Manual for Evidence Synthesis* (2 March version). https://jbi-global.atlassian.net/wiki/spaces/MANUAL/pages/863928570/

Jett, N. (2018). *Radically early college entrants on radically early college entrance: A heuristic exploration* [Doctoral dissertation, University of North Texas]. UNT Digital Library. https://digital.library.unt.edu/ark:/67531/metadc1404578/

Jett, N., & Rinn, A. N. (2020). Student experiences and outcomes of early college: A systematic review. *Roeper Review*, *42*(2), 80–94. https://doi.org/10.1080/02783193.2020.1728801

Justman, J. (1953). Personal and social adjustment of intellectually gifted accelerants and non-accelerants in junior high schools. *School Review*, *61*(8), 468–478. https://doi.org/10.1086/442050

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*(17), 2693–2710. https://doi.org/10.1002/sim.1482

Kopelman, C. I. (2016). *Summum bonum: Investigating global life satisfaction in a sample of profoundly gifted students* [Doctoral dissertation, University of Iowa]. Iowa Research Online. https://doi.org/10.17077/etd.mxq3z72m

Kretschmann, J., Vock, M., Lüdtke, O., & Gronostaj, A. (2016). Skipping to the bigger pond: Examining gender differences in students' psychosocial development after early acceleration. *Contemporary Educational Psychology*, *46*, 195–207. https://doi.org/10.1016/j.cedpsych.2016.06.001

Kronborg, L. (2018). Gifted education in Australia and New Zealand. In S. I. Pfeiffer (Ed.), *APA handbook of giftedness and talent* (pp. 85–96). American Psychological Association.

Kulik, J. A. (2004). Meta-analytic studies of acceleration. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived: How schools hold back America's brightest students* (Vol. 2, pp. 13–22). University of Iowa.

LeBeau, B. C., Assouline, S. G., Foley-Nicpon, M., Lupkowski-Shoplik, A., & Schabilion, K. (2025). Likelihood of whole-grade or subject acceleration for twice-exceptional students. *Gifted Child Quarterly* [Advance online publication; Iowa Research Online repository preprint]. https://doi.org/10.1177/00169862241302813

Long, L. H. (1974). *Does migration interfere with children's progress in school?* (Revised conference report, ERIC Document Reproduction Service No. ED091475). https://files.eric.ed.gov/fulltext/ED091475.pdf

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later [with Online Supplemental Material and 2015 Corrigendum, *26*(4), 552–553]. *Psychological Science*, *25*(12), 2217–2232. https://doi.org/10.1177/0956797614551371

Mammadov, S., Hertzog, N. B., & Mun, R. U. (2018). An examination of self-determination within alumni of an early college entrance program. *Journal for the Education of the Gifted*, *41*(3), 273–295. https://doi.org/10.1177/0162353218781745

Mokher, C. G., Park-Gaghan, T. J., & Hu, S. (2021). Does the method of acceleration matter? Exploring the likelihood of college coursetaking success across four developmental education instructional strategies. *Teachers College Record*, *123*(9), 3–27. https://doi.org/10.1177/01614681211052000

Motamedi, Z., Afrooz, G., & Arjmandnia, A. (2022). Investigating and comparing the behavioral and emotional problems of students using grade skipping with normal students. *Monthly Journal of Psychological Science*, *20*(107), 2045–2054.

Noble, K. D., Vaughan, R. C., Chan, C., Childers, S., Chow, B., Federow, A., & Hughes, S. (2007). Love and work: Perceived impact of early university entrance on adult life. *Gifted Child Quarterly*, *51*(2), 152–166. https://doi.org/10.1177/0016986207299472

Noreen, G. D., Lubinski, D., & Benbow, C. P. (2025). In their own voice: Educational perspectives from intellectually precocious youth as adults. *Gifted Child Quarterly*, *69*(4), 307–332. https://doi.org/10.1177/00169862251339670

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. *BMJ, 372, Article n71.* https://doi.org/10.1136/bmj.n71

Park, G., Lubinski, D., & Benbow, C. P. (2013). When less is more: Effects of grade skipping on adult STEM productivity among mathematically precocious adolescents. *Journal of Educational Psychology*, *105*(1), 176–198. https://doi.org/10.1037/a0029481

Pevec, A. E. (1965). *An analysis of academic achievement and personality adjustment of accelerated high school students* [Unpublished master's thesis, as cited in Kulik, 2004].

Pollins, L. D. (1983). The effects of acceleration on the social and emotional development of gifted students. In C. P. Benbow & J. C. Stanley (Eds.), *Academic precocity: Aspects of its development* (pp. 160–178). Johns Hopkins University Press.

Richardson, T. M., & Benbow, C. P. (1990). Long-term effects of acceleration on the social-emotional adjustment of mathematically precocious youths. *Journal of Educational Psychology*, *82*(3), 464–470. https://doi.org/10.1037/0022-0663.82.3.464

Robinson, N. M., & Janos, P. M. (1986). Psychological adjustment in a college-level program of marked academic acceleration. *Journal of Youth and Adolescence*, *15*(1), 51–60. https://doi.org/10.1007/BF02140783

Rogers, K. B. (2004). The academic effects of acceleration. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived: How schools hold back America's brightest students* (Vol. 2, pp. 47–57). University of Iowa.

Rogers, K. B. (2007). Lessons learned about educating the gifted and talented: A synthesis of the research on educational practice. *Gifted Child Quarterly*, *51*(4), 382–396. https://doi.org/10.1177/0016986207306324

Rogers, K. B. (2015). The academic, socialization, and psychological effects of acceleration: Research synthesis. In S. G. Assouline, N. Colangelo, J. VanTassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered: Evidence trumps the excuses holding back America's brightest students* (Vol. 2, pp. 19–30). Belin-Blank Center, University of Iowa.

Sayler, M. F., Boazman, J., Natesan, P., & Periathiruvadi, S. (2015). Subjective well-being of gifted American college students: Initial psychometric validation of the Personal Wellbeing Index–Adult. *Gifted Child Quarterly*, *59*(4), 236–248. https://doi.org/10.1177/0016986215597748

Sayler, M. F., & Brookshire, W. K. (1993). Social, emotional, and behavioral adjustment of accelerated students, students in gifted classes, and regular students in eighth grade. *Gifted Child Quarterly*, *37*(4), 150–154. https://doi.org/10.1177/001698629303700403

Schuur, J., van Weerdenburg, M., Hoogeveen, L., & Kroesbergen, E. H. (2021). Social–emotional characteristics and adjustment of gifted accelerated university students: A systematic review. *Gifted Child Quarterly*, *65*(1), 29–51. https://doi.org/10.1177/0016986220969392

Shepard, S. J., Foley Nicpon, M., & Doobay, A. F. (2009). Early entrance to college and self-concept: Comparisons across the first semester of enrollment. *Journal of Advanced Academics*, *21*(1), 40–57. https://doi.org/10.1177/1932202X0902100103

Singh, S. P., Menon, S., Singh, S., Nadeau, A. J., & Jiao, J. (2021). Perceived depression, anxiety, and stress outcomes among students of early entrance college programs. *Cureus, 13(12), Article e20360.* https://doi.org/10.7759/cureus.20360

Smedsrud, J. (2018). Mathematically gifted accelerated students participating in an out-of-school mathematics mentor program: A Norwegian perspective. *Frontiers in Psychology, 9, Article 1359.* https://doi.org/10.3389/fpsyg.2018.01359

Stanley, J. C. (1973). *Accelerating the educational progress of intellectually gifted youths* (Revised presidential address, ERIC Document Reproduction Service No. ED087816). https://files.eric.ed.gov/fulltext/ED087816.pdf

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: Findings of two second-order meta-analyses. *Review of Educational Research*, *86*(4), 849–899. https://doi.org/10.3102/0034654316675417

Steenbergen-Hu, S., & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. *Gifted Child Quarterly*, *55*(1), 39–53. https://doi.org/10.1177/0016986210383155

Stride, C. F. (1996). *The historical development of the Texas Academy of Mathematics and Science 1987–1992* [Doctoral dissertation, University of North Texas]. UNT Digital Library. https://digital.library.unt.edu/ark:/67531/metadc279092/

Swiatek, M. A., & Benbow, C. P. (1991). Ten-year longitudinal follow-up of ability-matched accelerated and unaccelerated gifted students. *Journal of Educational Psychology*, *83*(4), 528–538. https://doi.org/10.1037/0022-0663.83.4.528

Thomas, T. A. (1987). *CSUS Academic Talent Search follow-up report: After the first four years* (ERIC Document Reproduction Service No. ED287253). California State University, Sacramento.

Thomas, T. A. (1993). *The achievement and social adjustment of accelerated students: The impact of Academic Talent Search after seven years* (ERIC Document Reproduction Service No. ED368146). California State University, Sacramento.

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, *102*(4), 860–871. https://doi.org/10.1037/a0019454

Wells, R., Lohman, D. F., & Marron, M. (2009). What factors are associated with grade acceleration? An analysis and comparison of two U.S. databases. *Journal of Advanced Academics*, *20*(2), 248–273. https://doi.org/10.1177/1932202X0902000203

Wu, J. (2018). *An investigation of early college entrants' social development* [Doctoral dissertation, University of Iowa]. Iowa Research Online. https://doi.org/10.17077/etd.ozihjd9e

Wu, J., Assouline, S. G., McClurg, V. M., & McCallum, R. S. (2022). An investigation of an early college entrance program's ability to impact intellectual and social development. *Roeper Review*, *44*(2), 95–109. https://doi.org/10.1080/02783193.2022.2043503

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CleanednewT05FinalS1SupplementQ1.pdf