JOURNAL FOR SCHOOL SUPERINTENDENTS

Movement Thinking

Leadership, Learning Work and the Future of School

VOLUME 1 / 2026

5 conceptual articles and 6 primary-study protocols with completed reproducible methods analyses

Dr. Freedom Cheteni; Agathe Bertin | Article-specific human authorship

The Research University | Official publisher

Prepared 2026-10-16 | Research programme and methods edition

This volume contains 42,015 words of article body text, linked references, evidence-and-limits tables and the complete 100-source register. The 100-source register covers 2017-2026; foundational citations also include earlier scholarship. These new manuscripts were prepared in 2026.

Review credits are version-specific. Each protocol distinguishes its completed database and design analyses from the primary outcomes its prospective study would collect.

Read the complete volume online

Contents

Editorial orientation and edition record 4

From Instructional Leadership to the Design of Learning Work: A Movement Thinking Framework and Empirical Evaluation Strategy for Future-Ready Schools 6

The Automation Dividend Belongs in the Timetable: A Work-Design Test for AI in School Leadership 28

The District as Accountable Intelligence: Governing Paper Agents Without Outsourcing Educational Judgment 39

Beyond the Rendered Classroom: Movement Thinking and the Conditions of Transfer in Spatial Learning 51

Belonging by Design: Cultural Authority, Spatial Learning and the Work of School Leadership 62

Publication as Learning Infrastructure: Relational Accountability, Student Agency and the Scholarly Record 73

Who Receives the Automation Dividend? A Matched-Pair School Trial of Protected Time, Teacher Workload and Verified Quality 84

Accountable Intelligence at the Point of Decision: A Europe PMC Corpus Audit and Randomized Protocol for Evaluating Paper Agents in School-District Evidence Work 102

From Presence to Revisable Explanation: A Factorial Primary-Study Protocol for Spatial Transfer with NASA Exoplanet Data 119

Who May Change the Learning Place? A Community-Governed Trial of Cultural Authority, School Belonging and Disciplinary Transfer 139

Beyond the Published Portfolio: An Optional-Publication Trial of Learning Transfer, Provenance Mastery and Scholarly Portability 159

Appendix: complete 100-source register 178

Article format and publication record 211

Editorial orientation and edition record

From a philosophy to a research programme

The volume develops Movement Thinking into specified constructs, competing explanations and prospective evaluations. A01 is a primary-study protocol with an executed SEDA file audit and design simulations. Five conceptual articles examine automation, accountable agents, spatial learning, cultural authority and publication as learning. Corresponding primary-study protocols use distinct authoritative databases and evaluation designs. Source findings remain distinct from proposed mechanisms and study outcomes.

The VR School media essays supply philosophical and authorial provenance. Their institutional outcome, partnership and accreditation claims are not treated as independently established research findings. No universal retention percentages, new student results or causal effects of a school programme are asserted.

How to read the evidence

References preserve the named source authors, venues, dates and identifiers. Article-level rights evidence for the 100-source register was checked on 15 October 2026. Cited media and copyrighted research are cited rather than republished as Creative Commons text. Bibliographic and full-text checks are distinct from critical appraisal and independent peer review.

Status and responsibility

Dr. Freedom Cheteni is a named human author. Agathe Bertin is first author of A09, with Dr. Cheteni as second author and Principal Investigator. ChatGPT assisted drafting and organization and is disclosed as software, not a human author or peer reviewer. Review of the five conceptual manuscripts and the earlier A01 conceptual text by Dr. Vickie Aultmon was confirmed by Dr. Cheteni on 15 October 2026. Version 3.0 additions to A01 and new companion protocols have separate content and editorial records. Human review of a conceptual argument is not attributed to a new study protocol.

The volume is publicly available as a research and methods edition. Review attribution is tied to each version. Author approval, independent editorial acceptance and DOI registration remain separate records. No historical publication date is inferred from a

source year. The accompanying 100-source register is supporting literature, not 100 new articles by Dr. Cheteni.

Authorship, AI, review, dates and rights policy

Source identity and access verification record

A01 | PRIMARY-STUDY PROTOCOL AND REPRODUCIBLE METHODS RESEARCH

From Instructional Leadership to the Design of Learning Work: A Movement Thinking Framework and Empirical Evaluation Strategy for Future-Ready Schools

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 3.0 | 5,341 article words

Prior conceptual-edition review: Dr. Vickie Aultmon. Review confirmation recorded by the publisher on 2026-10-15.

Keywords: Movement Thinking, school leadership, cluster randomization, primary-study protocol, TALIS 2024, SEDA 2025.2, reproducible research

Read online / article companion / citation exports

Contribution

An operational primary-study protocol, an executed SEDA data-feasibility audit and reproducible design-sensitivity analysis for testing learning-work design.

Prospective school-randomized study specification; completed secondary-file audit and artificial-data simulations. No local trial outcomes are reported.

Abstract

School leadership, teacher development and preparation for changing work are often governed as separate activities. Movement Thinking connects them through the design of learning work: consequential questions, bounded decision authority, protected collaboration, inspectable revision and community accountability. This article specifies a prospective primary study of that framework and reports two completed methodological investigations: an audit of a versioned public education dataset and a reproducible simulation of candidate cluster designs. The proposed primary study assigns schools within districts to a structured inquiry protocol or an active comparator with equivalent scheduled time and facilitator contact. Its primary estimand is the equal-school average effect on change in a blinded, 20-point student transfer-and-revision assessment. Teacher interpretation, implementation and workload are examined separately from student performance. TALIS 2024 supports a distinct, survey-weighted secondary investigation; SEDA 2025.2 supplies historical district context. Their public identifiers are not assumed to provide a shared linkage key. An executed SEDA file audit identifies 267,401 records and, after a documented period and subgroup restriction, 8,887 districts with the selected learning-rate estimate and a positive adjusted standard error. Simulations expose the sensitivity of power to school count and intracluster correlation. These are methodological results, not observed effects of Movement Thinking. The contribution is a falsifiable, versioned route from an educational philosophy to primary evidence, with specified outcomes, an active comparison, reproducible tools and explicit limits on causal and institutional claims.

1. The leadership problem is a work-design problem

An education system cannot prepare learners for every future occupation by maintaining an ever-longer list of technologies. A more durable responsibility is to develop the capacity to investigate unfamiliar questions, distinguish evidence from assertion, coordinate with others and revise a judgment when its premises change. These capacities are enacted through tasks and relationships. They depend on what students are allowed to decide, what teachers have time to examine and what happens when a participant discovers an error.

Leadership therefore enters the learning process through the design of work. A superintendent who requests inquiry but protects no time for instructional revision creates a different intervention from one who reorganizes schedules, clarifies decision rights and makes critical feedback safe. The difference is not captured adequately by counting

devices, meetings, portfolios or publications. It concerns the conditions under which a defensible change in understanding can occur and be examined by others.

This argument does not remove disciplinary knowledge, safeguarding or public accountability. Bounded authority means specifying which decisions can be changed, by whom and through which review process. It also means making exclusions visible: a compelling inquiry project should not depend on unpaid teacher labor, inaccessible technology or a student's willingness to publish personal work. The empirical question is whether a specified organization of inquiry improves performance on demanding learning tasks at a sustainable human cost.

2. Intellectual provenance and the contribution of Movement Thinking

Cheteni's published account presents Movement Thinking as an orientation toward fluidity, inclusion and the continuing evolution of teaching [1]. The present article converts that orientation into a bounded intervention theory. Its distinctive object is the transition between a question, an initial representation, a challenge, a reasoned revision and an accountable decision. The originality claimed is this operational specification and its testable design, rather than ownership of established ideas about collaboration, leadership or psychological safety.

Several adjacent literatures make the specification plausible without establishing its effectiveness. A meta-analytic structural model links instructional leadership, collective efficacy and teacher self-efficacy [2]. Research on distributed leadership, collaboration and job satisfaction identifies related associations in a different education system [3]. An inclusive-leadership study examines another proposed mediation pathway [4]. Their designs and populations delimit what can be transferred; none is a randomized evaluation of the package proposed here.

The School Implementation Leadership Scale provides relevant measurement evidence for leadership supporting school-based prevention programmes [5]. Edmondson's work locates psychological safety and learning behavior in organizational teams [7]. These sources motivate distinct constructs. Implementation leadership is not psychological safety; a teacher's self-efficacy is not collective efficacy; and reported collaboration frequency is not the quality of professional interpretation. A credible evaluation must preserve these distinctions rather than combine agreeable survey responses into a single success index.

3. Five practices, observable indicators and theoretical boundaries

The first practice makes a consequential question visible. A task identifies an audience, a decision or an unresolved problem, together with disciplinary criteria for an adequate answer. Its indicator is an inspectable task brief and a learner's explanation of the problem. The apparent importance of a topic does not establish that the task demands reasoning; independent reviewers must also examine its cognitive demands and accessibility.

The second practice distributes bounded authority. A short decision charter identifies decisions available to the superintendent, principal, teacher team and learner, as well as safety constraints and routes of appeal. Its indicator is the correspondence between the charter and observed decisions. Autonomy without usable resources, or discretion that disappears when a learner challenges an adult, does not count as implementation merely because a policy uses participatory language.

The third practice protects collaborative time. Scheduled sessions must replace or reorganize existing work, identify the instructional problem being examined and record a subsequent decision. The fourth preserves evidence of revision: a small, purposeful sample of initial work, a challenge or feedback event, revised work and a reason for the change. More uploads are not inherently better evidence. The protocol deliberately caps documentation rather than rewarding its volume.

The fifth practice reviews consequences with affected communities. Students, educators and families consider whether the changes served their intended purposes, who encountered barriers and what should stop or change. Implementation indicators describe whether these practices occurred. They are not substitutes for the independent student outcome. Keeping the process and outcome measures separate prevents the circular conclusion that a programme worked because people performed the programme's own prescribed activities.

4. Research questions, hypotheses and the causal target

The primary question is whether the structured Movement Thinking inquiry package improves student transfer-and-revision performance relative to an active, time-matched professional-learning comparator. The directional hypothesis is a positive effect, assessed with a prespecified two-sided test. The primary endpoint is measured at the end of an 18-week implementation period. A later transfer task, teacher workload and implementation measures are secondary outcomes, with their families and interpretation declared before outcome inspection.

Let Z_s indicate a school's assigned condition and let DeltaY_s(z) be the mean change in assessment score that its fixed baseline cohort would exhibit under assignment z. The target is tau = the average across participating schools of DeltaY_s(1) minus DeltaY_s(0). Schools receive equal weight. This estimand concerns the recruited, consenting school cohorts; random assignment within recruited districts does not establish national representativeness or an effect for every individual student.

The mechanism hypothesis is that the package changes opportunities for professional interpretation, which in turn changes instructional decisions and learners' revision. That pathway remains a hypothesis even if a fitted structural model produces an indirect coefficient. Assignment randomizes the package, not psychological safety, team interpretation or teacher compliance. The main treatment effect therefore has a stronger identification basis than any proposed mediation decomposition [15]. Future employment, lifelong adaptability and institutional prestige are outside the measured causal target.

5. A prospective primary-study design

The planning design contains 80 public middle or combined schools, arranged in 20 district blocks of four schools, with 40 sampled students per school. This is a candidate design for budgeting and design sensitivity, not a recruited sample. District agreements, a funding plan, an independent ethics determination and a registered protocol must precede study activation. Recruitment feasibility may require a different design; any change must be recorded before allocation, with renewed power calculations.

Eligible schools serve the target grades, can implement both conditions within paid schedules and can administer accessible assessment forms. The sampling plan seeks 20 seventh-grade and 20 eighth-grade students per school from a documented eligible roster, with accommodations preserved. Student sampling and consent occur before allocation. The protocol does not select only polished projects, enthusiastic teachers or students whose families expect publication. A screening log records eligibility, invitations and refusals without exposing identities publicly.

Teacher teams include consenting educators responsible for the participating learning activities. Teacher sample size follows the actual roster rather than an invented fixed count. Schools are assigned after baseline assessment and cohort definition. Staff changes, discontinuation of the intervention and imperfect participation are tracked as intercurrent events; schools remain in their assigned condition for intention-to-treat follow-up. Recruitment after allocation is reported separately because it can introduce differential selection.

6. Allocation, concealment and the active comparator

Within each district, two of four schools are assigned to each condition using a reproducible allocation script. An independent allocation custodian holds the complete eligible roster, documents the generation method and releases assignments after baseline records are locked. The public protocol records the procedure and later its execution date. It does not publish an allocation seed while recruitment could still be manipulated. A dated, access-controlled record preserves the seed for subsequent audit.

Both groups receive a 90-minute paid professional-learning session every fortnight, equivalent facilitator contact, comparable curriculum access and the same assessment arrangements. The treatment adds the structured inquiry cycle, decision charter, bounded artifact trail and community consequence review. The comparator uses the protected time for its ordinary curriculum-focused professional work without the prescribed package. Natural overlap is measured rather than treated as misconduct. Facilitators use separate manuals and record deviations.

This comparison estimates the incremental value of the structured package beyond equal scheduled time and attention. It does not separately identify the effect of allocating protected time, nor the isolated contribution of each of the five practices. A future factorial or component study would be required for those questions. Teachers and leaders cannot be blinded to the activities they perform. Outcome raters and the primary analyst can remain masked to assignment until scoring and the specified analysis are locked.

7. A schedule that makes the mechanism inspectable

Baseline work includes cohort formation, consent, the first transfer task, teacher measures and a schedule audit. The intervention then runs through nine fortnightly cycles over 18 weeks. Each treatment cycle identifies an instructional question, examines a bounded sample of work, considers an alternative explanation, changes a task or support, and reviews the consequences. Sessions record the decision and the reason for it, not a transcript of every participant's contribution.

A midpoint assessment at approximately week 9 measures teacher interpretation and candidate organizational mechanisms. Endline student tasks are administered after the final cycle. A delayed task approximately six weeks later tests a changed context without reproducing the original problem. Teacher time diaries are collected in designated baseline, midpoint and endline windows, with an optional shorter diary for a prespecified subsample. Equivalent measurement contact is maintained in the comparator group.

Implementation fidelity is reported as separate dimensions rather than a convenient pass/fail label. The report distinguishes scheduled from delivered time, formal from usable decision rights, stored from informative revisions, and invited from consequential community participation. Low implementation can explain why an assigned package did not produce its expected mechanism, but it cannot be used to remove inconvenient schools from the primary effect estimate. A compliance analysis, if undertaken, is secondary and makes its additional assumptions explicit.

8. Primary outcome: a transfer-and-revision assessment

The primary instrument is a newly specified, 20-point research rubric with five dimensions scored from 0 to 4: problem representation, warranted use of evidence, consideration of alternatives, substantive revision and transfer under a changed constraint. It is a proposed instrument requiring pilot work, not an established validated Movement Thinking test. Pilot participants must be distinct from the confirmatory cohort. The task forms, scoring manual and revision rules are frozen before school allocation.

A task presents a realistic but fully fictional decision environment, an initial evidence packet and an explicit constraint. Students produce an initial response, encounter a new constraint or credible counterexample, then revise their recommendation and explain why. For example, a school-courtyard design task can introduce a change in the water budget and a requirement for accessible shaded routes. Fictional assessment inputs are labeled as such; they are not presented as observations from an actual school.

Parallel forms vary the substantive context while preserving reasoning demands. Form order is counterbalanced within schools and recorded before assignment. The baseline and endline sessions each include an opportunity to revise, so the primary change score compares like forms of performance rather than comparing an initial draft with a polished final product. The rubric does not award points simply for adopting a new technology, adding citations or agreeing with the facilitator.

Before the main study, the research team examines accessibility, content coverage, form comparability, ceiling and floor effects, and whether raters can distinguish the proposed dimensions. Student explanations inform a cognitive-interview pilot. A narrow pilot cannot validate the instrument for all cultures, languages or disabilities. If the instrument changes after outcome data become available, the revised analysis is identified as exploratory rather than silently replacing the confirmatory endpoint.

9. Blinded scoring, teacher interpretation and measurement quality

Two trained human raters independently score every primary assessment after identifiers and obvious assignment cues are removed. They see task-form information needed for scoring but do not see the student's school, condition, teacher survey answers or the other rater's score. The primary score is their mean total. A third rater reviews large discrepancies for quality assurance; the primary mean is not selectively overwritten because an adjudicated score produces a more favorable result.

Raters calibrate on pilot artifacts that do not enter the primary effect analysis. The report presents the distribution of disagreements, absolute-agreement reliability for the mean of two raters, and checks for drift across scoring batches. It also reports how often redaction failed to conceal condition. Reliability describes scoring consistency, not whether the rubric captures everything that matters about future readiness. Accuracy and fairness require substantive examination as well as a coefficient.

Teacher professional interpretation is measured separately through a new annotation task: identify a learning difficulty in a de-identified artifact, consider an alternative explanation and justify an instructional revision. Psychological safety and collective efficacy remain distinct survey constructs. The SILS may provide a complementary implementation-leadership measure, using its documented form and appropriate permissions; its published validation does not automatically validate a modified referent or a new instructional context [5]. New or adapted measures require their own measurement evidence.

10. External-data architecture: three evidence streams

The primary trial generates new learner artifacts, blinded scores, implementation records and teacher measures. TALIS provides an independent secondary source about teacher working conditions and related constructs [8]. SEDA supplies historical aggregate district context [11] [12]. These streams can inform a common theory while retaining different units, dates, permissions and inferential roles. A relationship between national survey constructs is not a substitute for a local treatment contrast.

The OECD guide documents scrambled public school and teacher identifiers and retained links within TALIS [9]. It does not establish a public NCES or SEDA crosswalk. Accordingly, the public research package contains no teacher-to-district linkage based on those IDs. A lawful future linkage would require an authorized crosswalk, a compatible population and

time window, and documented permission. It must not attempt to reverse confidentiality protections by matching quasi-identifiers.

For the proposed U.S. trial, an approved district identifier crosswalk can associate each recruited district with a prespecified SEDA context slice. The mapping is a district-level context lookup, not a join to TALIS respondents. SEDA estimates already end before the proposed intervention, so they can inform recruitment description or baseline context. They cannot be labeled post-intervention outcomes of a future trial. Any later administrative follow-up requires an actually available release and a separately defined analysis window.

11. TALIS 2024: a weighted secondary analysis

The audited teacher questionnaire identifies collaboration items TT4G26A-H, team-work time TT4G16B, and classroom-autonomy items TT4G57A-E [10]. The collaboration items are ordered frequency categories, not equal numbers of hours. The downloadable adapter estimates one specified category proportion rather than treating that scale as interval measurement. Actual country-file availability, missing-value codes and the variable inventory must be checked against the downloaded release before substantive analysis.

The 2026 user guide identifies TCHWGT and TRWGT1-TRWGT100 for teachers, SCHWGTC for principals, and Fay balanced repeated replication with factor 0.5 [9]. The adapter computes the full-weight estimate, 100 replicate estimates and the corresponding design-based standard error. Population level, country and unit of analysis are explicit. Teacher and principal records are not casually stacked, and a large number of teacher responses is not mistaken for a large number of independent schools.

Country-specific adjudication matters. The guide labels U.S. teacher participation poor and principal participation insufficient, with an explanatory principal-data footnote [9]. These qualifications must accompany any U.S. secondary findings. The primary trial's local measures of psychological safety, revision quality and implementation are not imputed from an international survey label. A cross-sectional TALIS association can assess compatibility with the theory; it cannot retrospectively randomize leadership or establish the package's causal effect.

12. Completed SEDA data-feasibility audit

The reproducible audit downloaded the administrative-district pooled, cohort-standardized SEDA 2025.2 file from Stanford's repository on 15 October 2026 [11]. Its SHA-256 is b6cad2c621ba1731bd8306215b98b32791e31d43aeb18c138a8fbaed5154577e. The file contains 267,401 records and 15,751 distinct administrative-district identifiers. These are counts of a source file, not the number of schools or students in a Movement Thinking study.

A district-plus-subgroup key is insufficient: it yields 172,945 distinct pairs for 267,401 records. The audited file separates periods and distinguishes subgroup and gap records. The code checks uniqueness using district, subgroup category, subgroup, gap status, comparison status and actual first and last years. All 267,401 complete keys are unique. This demonstrates why a naive district merge can duplicate records and create false precision even when the underlying repository is authoritative.

The planned context slice restricts subcat and subgroup to all, gap and multi_comp to zero, and pooled_years to 2022-2025. It contains 11,633 one-per-district records. Of these, 8,887 have a finite cs_mn_lrn_ol value and a finite positive cs_mn_lrn_ol_se_adj; 2,746 fail one of those checks. The audit reports these restrictions and counts explicitly. They do not imply that excluded districts have low learning rates or that a missing value is zero.

The selected estimate and its adjusted standard error are retained together. SEDA learning rates are modeled aggregate quantities, not individual student growth records [12]. Population composition, subject, period, suppression and scale remain relevant to interpretation. A historical district estimate can describe the setting of a future study; its agreement with a teacher survey would not validate a new student-transfer rubric, identify an intervention mechanism or remove ecological bias. The public audit redistributes counts and provenance, not the underlying district data file.

13. Primary analysis and competing explanations

For complete cohorts, the primary estimator first calculates each student's endline-minus-baseline rubric score, then the cohort mean change within each school. Within each four-school district block, it subtracts the mean of the two control-school changes from the mean of the two treatment-school changes. The treatment estimate is the average of those block contrasts. This implements the equal-school estimand and accounts for the assignment unit without pretending that 3,200 student records are 3,200 independent randomized units.

The analysis reports the estimate in rubric points, a Neyman within-block variance estimate and a prespecified small-sample t interval using district-block count minus one degrees of freedom. That interval is an approximation rather than an exact finite-sample guarantee. For complete cohorts, a blocked randomization test independently tests the sharp null of no effect for any school: it repeatedly assigns two of four schools to each condition and compares the resulting contrast with the observed contrast. Its Monte Carlo p value includes the observed allocation correction. When outcomes are missing, the primary estimate and interval pool the prespecified imputations; a permutation of once-filled outcomes is not represented as an exact randomization test.

The sharp-null test and an average-effect interval answer different questions, especially under heterogeneous effects. The report does not treat their agreement as automatic. Randomization inference preserves district blocks and never permutes individual student labels. The analysis code records the seed and number of permutations. A model-based student-level analysis with teacher and school clustering may be a secondary precision or sensitivity analysis, but does not replace the declared estimand after outcomes are known.

The attention account is addressed by equivalent meeting time and facilitator contact; the resource account by comparable scheduled support and documented baseline resources; and the selection account by pre-allocation recruitment and blocked assignment. None is eliminated by merely including a covariate. Differential co-interventions, contamination and attrition remain threats to interpretation. Post-treatment workload, staff movement and implementation are reported as outcomes or mechanisms rather than routinely controlled away from the total assigned-package effect.

14. Multilevel SEM as a secondary mechanism analysis

The mechanism analysis separates within-school teacher variation from between-school organizational variation. A candidate measurement model represents teacher reports through distinct safety, collective-efficacy and leadership constructs, with repeated measurements and correlated residuals specified in advance where justified. A structural model relates assignment to midpoint organizational measures and those measures to later school-level performance, adjusting for their baseline counterparts. This is a secondary theoretical test, not a license to call every fitted arrow causal.

Measurement work precedes the structural interpretation. It examines item distributions, ordinal thresholds, within- and between-school variability, factor identification, reliability and comparability over time and condition. It also considers whether school-level aggregation is defensible. A high teacher response count cannot compensate automatically for too few schools. If the planned sample cannot support the model, the report uses a

prespecified simpler association model and identifies the departure, rather than searching for whichever model supplies significant mediation.

The mediator is not randomized, and treatment may change mediator-outcome confounders. Consequently, an indirect effect requires assumptions that the trial assignment alone does not supply [15]. Sensitivity analyses and a temporal ordering argument can make those assumptions inspectable; they cannot prove them. The present package provides an estimand and identification specification for mechanism work, not a claim that an ordinal, survey-weighted, multilevel SEM has already been fitted or that its implementation is interchangeable across software.

15. Missing outcomes, departures and equity

The baseline cohort is fixed before assignment. Follow-up attempts include students and schools that discontinue the package, and reasons for missingness are recorded by arm and baseline characteristics. The report separates missing artifacts, withdrawn research consent, school departure, administrative suppression and unavailable teacher responses. It does not quietly redefine the sample as people who completed a polished final product. A missing endline score is never coded as educational failure or as a zero rubric score in the main analysis.

The specified missing-data approach uses 50 posterior predictive imputations from a bounded, hierarchical endline-score model, with baseline score, baseline grade, task-form order, assignment and district fixed effects and teacher-within-school random intercepts. The methods specification defines its priors, diagnostics and pooling rules [18]. This analysis assumes that missingness is conditionally independent of the unseen score after the included information. Sensitivity analyses shift missing-score predictions by -2, -1, 0, 1 and 2 rubric points within the permitted range, and also present 0-20 worst-case bounds. The imputation specification is frozen before allocation and implemented independently of the complete-cohort adapter. The supplied complete-cohort adapter deliberately refuses incomplete records; it does not implement imputation or claim that listwise deletion solves attrition. This boundary prevents a convenience script from silently becoming the missing-data policy.

Equity analyses examine participation, accessible administration, missingness and score interpretation before comparing subgroup treatment effects. Subgroup variables are collected only when justified and consented. Small cells and identifiable combinations are protected. Treatment-by-subgroup estimates are prespecified and accompanied by uncertainty and multiplicity rules. An imprecise subgroup estimate is not evidence that a

group did not benefit, and a favorable average does not establish that burdens or barriers were equitably distributed.

16. Completed design-sensitivity analysis

The design analysis simulates school mean changes under a deliberately simple model, with four schools per district block, two assigned to each arm, 40 students per school, equal cluster sizes and complete outcomes. It varies school count across 48, 80 and 120, and the intracluster correlation of change scores across 0.05, 0.10 and 0.15. Each scenario uses 10,000 replicates with a recorded seed. These are artificial draws for planning, not reconstructed or anonymized student observations [17].

For the 80-school, 3,200-student planning design and an assumed standardized effect of 0.25, simulated rejection frequencies were 97.71%, 85.89%, 72.56% at intracluster correlations of 0.05, 0.10 and 0.15, respectively. The corresponding Monte Carlo standard errors were 0.15%, 0.35%, 0.45%. The normal-approximation 80%-power minimum detectable effects were 0.170, 0.219, 0.259 change-score standard deviations. These conditional planning results make the higher-correlation design risk visible; they are not evidence that schools improved.

The normal-approximation sensitivity calculation uses DE = 1 + (m-1)rho and MDE = (z_0.975 + z_0.80) times the square root of 4DE/(Km), where K is school count and m is cohort size. It expresses effects in standard deviations of individual change scores, not of endline achievement or SEDA's reference population. Translating 0.25 into one rubric point would require a planning assumption that the change-score standard deviation is four; that quantity has not been estimated from a local cohort.

The simulation also reports null rejection rates and Monte Carlo standard errors. Its favorable scenarios do not guarantee achieved power. Unequal school size, unreliable scoring, missing outcomes, implementation differences and a smaller number of schools can change precision substantially. A final sample commitment requires an independent pilot, a frozen primary analysis and further simulations covering realistic departures. The design file exposes these assumptions so that increasing student counts cannot conceal an inadequate number of randomized schools.

17. Disconfirmation, workload and decisions to scale

The framework is challenged if its documentation indicators improve but the independently rated learning outcome does not. A narrow confidence interval that rules out a prespecified worthwhile effect is more informative for that decision than a nonsignificant p value with wide uncertainty. The decision specification proposes one rubric point as a candidate practical threshold for stakeholder deliberation, not a validated universal minimum. Its final value and justification must be fixed before allocation.

The workload criterion is evaluated alongside learning. The protocol proposes a candidate harm margin of 30 additional unpaid teacher minutes per week, to be reviewed with educators and ethics oversight before registration. A learning gain does not authorize breaching that margin or ignoring a serious burden for a particular group. Conversely, reduced paperwork with no meaningful change in transfer performance could support an efficiency claim while failing the stronger learning hypothesis. The report distinguishes these possibilities rather than forcing one success label.

Scaling requires an inspectable pattern: a meaningful and sufficiently precise learning estimate, acceptable burden and cost, evidence that the intended practices occurred, and no unresolved serious accessibility or privacy problems. An adverse or inconclusive result is publishable. The package should be revised or stopped when its burden exceeds its educational value. Positive findings would justify a bounded replication under specified conditions, not the assertion that every school should adopt the programme or that future labor-market outcomes have been demonstrated.

18. Ethics, governance and institutional responsibility

The study requires an independent ethics determination, district research permissions and appropriate consent and assent before recruitment or collection of research artifacts. Their identifiers and dates are entered only when actually issued. Ordinary educational participation, research participation and public publication are separate decisions. Declining research or publication must not reduce grades, admission prospects, teaching access or financial support. An ORCID account is not a condition for a child to receive instruction or participate in an assessment.

Direct identifiers and the re-identification key remain under restricted access with a documented retention and deletion schedule. The public package contains instrument definitions, code, design simulations, aggregate audit counts and methodological decisions, not identifiable student work or sensitive teacher reports. School leaders do not receive identifiable psychological-safety responses for personnel evaluation. Any future

release of examples requires separate authorization and a re-identification assessment; removal of names alone is insufficient.

An independent study adviser or oversight group should review recruitment, burdens, data quality and protocol departures. Its membership and authority are reported when appointed. Authorship follows actual contributions and consent, and advisory input is acknowledged appropriately. Stanford is credited as SEDA's institutional source, not entered as a collective author without named contributors. The publication venue for this edition is the Journal for School Superintendents. A target journal's standards can inform reporting without implying that the target journal has published or accepted the work [16].

19. Reproducibility, paper agents and the scholarly record

The methods package contains a versioned study specification, original draft instruments, source-provenance records, an executed SEDA audit, reproducible design calculations and documented command-line tools. The tools allocate a supplied school roster, analyze complete synthetic or properly governed study cohorts, and compute a narrowly specified TALIS weighted proportion with Fay replication. Their supported inputs and failure conditions are explicit. The package does not contain a fitted primary-study dataset, a completed SEM or a registered trial identifier.

The article's paper agent retrieves passages from this edition and links them to their sections. Its useful function is to make the boundary between a proposed procedure, an executed audit and an observed treatment effect easier to inspect. It has no authority to generate missing student data, declare a hypothesis confirmed or supply human peer review. Confidential participant records should not be sent to the public article agent. A future data agent would require separate access controls and validated analysis behavior.

Reporting draws on educational research standards and appropriate protocol and cluster-trial reporting principles [13] [14] [16]. It distinguishes planned from executed procedures, assignment from adherence, sample counts from source-file counts, and confirmatory from exploratory analyses. Preregistration would record the analysis and departures prospectively; a prepared specification is not a completed registration. The archived conceptual edition and its review attribution remain available, while this substantially revised methods edition has its own version and content record.

20. Limitations and boundaries of inference

The proposed trial estimates an effect within recruited school cohorts, not across all education systems. District self-selection can limit transportability even when assignment within districts is sound. Staff cannot be blinded to the package, and communities may recognize its language. Outcome masking, an active comparator and contamination records reduce some threats without removing expectation effects or differential co-interventions. The novel rubric requires evidence that it measures reasoning and revision fairly; scoring reliability alone is insufficient.

The data streams have different limitations. TALIS combines self-report, cross-sectional measurement and country-specific participation qualifications. SEDA supplies aggregate modeled estimates with missingness and period distinctions. The simulation assumes complete, equally sized cohorts and a simple change-score distribution. The imputation model requires assumptions about unobserved outcomes; the mediation model requires assumptions beyond treatment assignment. None of these methods establishes eventual employment, accreditation, prestige or universal transfer. A strong report makes these boundaries visible alongside any favorable result.

21. Implications for superintendents

The practical question for a superintendent is not whether inquiry sounds desirable. It is whether a specified arrangement of time, authority, evidence and relationships produces learning that an independent person can recognize, with burdens that educators and learners can sustain. The proposed design makes that question answerable without relying on a showcase project, a technology purchase or a favorable self-report as the decisive outcome.

The completed audit and design analysis already improve the quality of that decision. They prevent an invalid public-file linkage, expose a duplicated-merge risk and show why the number of schools and the correlation within them matter for precision. Their contribution is methodological. The new learner evidence needed to estimate a primary treatment effect must still be generated through the specified, ethically authorized study. That separation protects both the intellectual ambition of Movement Thinking and the credibility of any eventual finding.

Method and data statement

This edition combines a prospective primary-study protocol with an executed public-file feasibility audit and a simulation-based design analysis. The SEDA audit reads a pinned 2025.2 file; TALIS instrument and weighting details were checked against official documentation, but no TALIS respondent analysis is reported. Primary fieldwork is prospective in this protocol; no local participant outcomes were supplied or analyzed for this edition. Source versions, executable procedures, input boundaries, seeds and output hashes are available in the methods package. Reporting principles are informed by SPIRIT 2025 and cluster-trial guidance, adapted to an education study rather than represented as clinical-trial certification [13] [14].

AI, authorship and editorial disclosure

ChatGPT assisted with literature organization, methodological drafting, code preparation and verification support on 15 October 2026 [6]. The named human author is Dr. Freedom Cheteni. Human investigators retain responsibility for design choices, permissions, verification and final approval. AI is a disclosed tool, not an author, participant, data collector or peer reviewer. The tools and findings labeled simulation or data audit must not be represented as observed educational effects.

Dr. Vickie Aultmon's human review of the preceding conceptual edition was confirmed by Dr. Freedom Cheteni on 15 October 2026. That review is preserved with the earlier edition. The new empirical protocol, source-file audit and analysis code are version 3.0 additions and are identified separately in the review record. Dr. Cheteni's role as framework originator and Editor in Chief is disclosed; actual editorial decisions remain the responsibility of independent human handling.

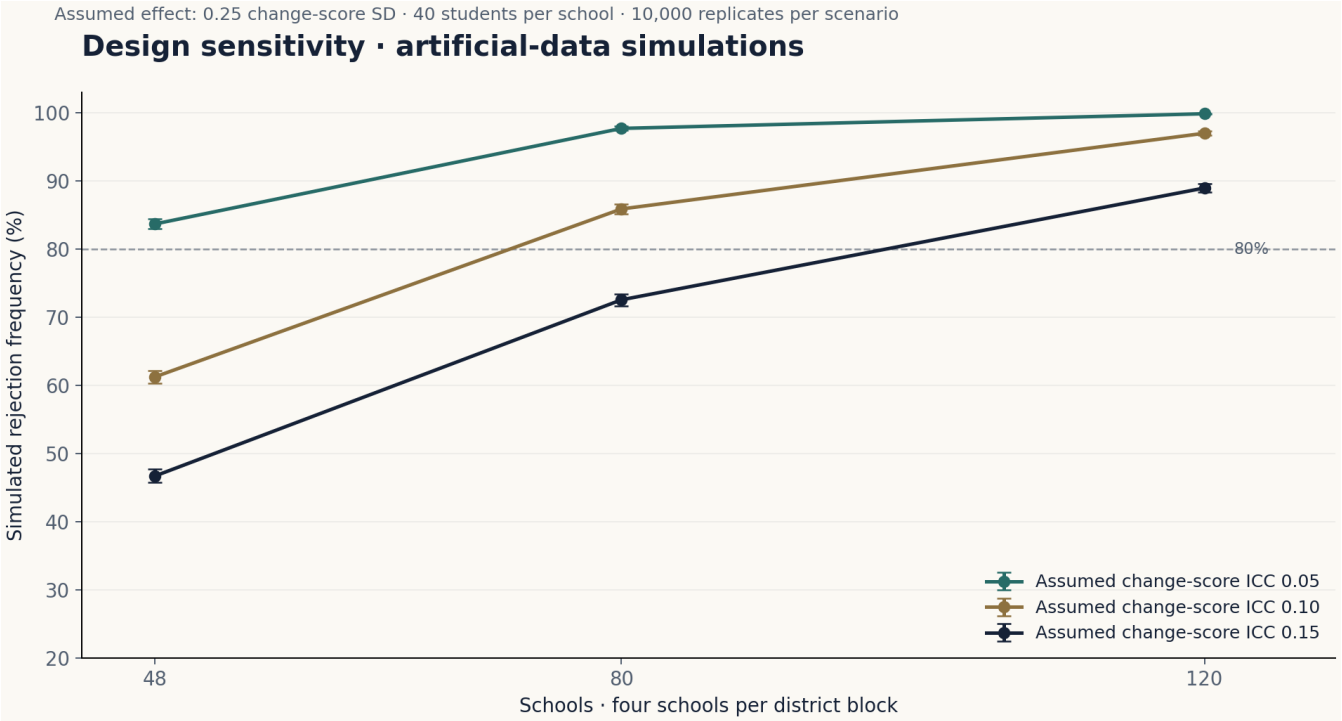


Figure 1. Design sensitivity under explicitly assumed conditions: 10,000 artificial-data replicates per scenario; effect 0.25 change-score standard deviations, 40 students per school, two of four schools assigned within each district. Whiskers show ±1.96 Monte Carlo standard errors, not uncertainty about a real intervention effect.

Research materials and reproducibility

Inspect the prospective study specification, original draft instruments, executed public-file audit, design simulations and documented tools. No local participant trial outcomes are reported.

Complete methods package · ZIP

Methods guide

Study specification

Assessment instruments & scoring rubric

Statistical analysis specification

Executable research tools · Python

Executed SEDA file audit

Design simulations & uncertainty

Source and variable audit

Earlier conceptual edition and version-specific review record

Evidence and limits

Claim	Evidence	Limit
Learning-work design can be tested as an assigned package.	A prespecified school-randomized protocol, active comparator and independent learning endpoint. [1, 13, 14]	The protocol is prospective; it supplies no observed treatment effect.
Public TALIS and SEDA identifiers do not establish a shared school-level join.	Official documentation describes scrambled TALIS identifiers; SEDA uses administrative-district units. [9, 12]	Any future linkage requires a separately authorized, compatible crosswalk.
The audited SEDA file requires period and subgroup restrictions before a one-per-district context lookup.	Executed, hashed source-file audit: 267,401 rows; 11,633 selected context records; 8,887 with the specified estimate and adjusted SE. [11, 12]	Data availability and historical context are not causal validation of Movement Thinking.
Power depends materially on school count and intracluster correlation.	Reproducible nine-scenario, 10,000-replicate design calculation and reported Monte Carlo error. [17]	The simulation assumes equal size, complete outcomes and an artificial Gaussian change-score model.
A fitted mediation path does not inherit identification from treatment randomization.	Distinct midpoint measures, a secondary multilevel model and an explicit missing-data specification. [5, 7, 15, 18]	The mediator is not randomized; imputation and mediation require additional assumptions.

References and rights

[1] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only. Rights retained by the original rights holder; no text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[2] Turgut Karakose; Abdurrahman Kardas; Sedat Kanadlı; Tijen Tülübaş; Bilal Yildirim. How Collective Efficacy Mediates the Association between Principal Instructional Leadership and Teacher Self-Efficacy: Findings from a Meta-Analytic Structural Equation Modeling (MASEM) Study. Behavioral Sciences. 2024-01-25.

https://doi.org/10.3390/bs14020085

CC BY 4.0. © 2024 by the authors.

Authoritative source / evidence | checked 2026-10-15

[3] Xiaodong Fan; Zuwang Chu. The Influence of Distributed Leadership on Chinese Teachers' Job Satisfaction: The Chain Mediation of Teacher Collaboration and Teacher Self-Efficacy. International Journal of Environmental Research and Public Health. 2025-03-26.

https://doi.org/10.3390/ijerph22040507

CC BY 4.0. © 2025 by the authors.

Authoritative source / evidence | checked 2026-10-15

[4] Yuangen Bao. A dual mediation model of the association between principal inclusive leadership and teacher innovative behavior. Scientific Reports. 2024-05-29.

https://doi.org/10.1038/s41598-024-63332-0

CC BY 4.0. © The Author(s) 2024

Authoritative source / evidence | checked 2026-10-15

[5] Aaron R Lyon; Catherine M Corbin; Eric C Brown; Mark G Ehrhart; Jill Locke; Chayna Davis; Elissa Picozzi; Gregory A Aarons; Clayton R Cook. Leading the charge in the education sector: development and validation of the School Implementation Leadership Scale (SILS). Implementation Science : IS. 2022-07-19.

https://doi.org/10.1186/s13012-022-01222-7

CC BY 4.0. © The Author(s) 2022

Authoritative source / evidence | checked 2026-10-15

[6] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[7] Amy Edmondson. Psychological Safety and Learning Behavior in Work Teams. Administrative Science Quarterly. 1999-06.

https://doi.org/10.2307/2666999

Citation only; no Creative Commons article license asserted. Source text and figures are not republished. Original authors and publisher retain their rights.

Authoritative source / evidence | checked 2026-10-15

[8] Organisation for Economic Co-operation and Development. TALIS 2024 Database. OECD. Public-file update 2025-11-24; accessed 2026-10-15.

https://www.oecd.org/en/data/datasets/talis-2024-database.html

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[9] Organisation for Economic Co-operation and Development. Teaching and Learning International Survey (TALIS) 2024 User Guide, Version 3. OECD; EDU/TALIS/GB(2025)9/REV1. 2026-08-18.

https://one.oecd.org/document/EDU/TALIS/GB(2025)9/REV1/en/pdf

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[10] Organisation for Economic Co-operation and Development. Teaching and Learning International Survey (TALIS): Teacher Questionnaire, ISCED 1, 2, 3. OECD. 2024.

https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/talis/talis2024questionnaires/TALIS2024TeacherQuestionnaireISCED1-2-3.pdf

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[11] Sean F. Reardon; Erin M. Fahle; Andrew D. Ho; Benjamin R. Shear; Jim Saliba; Jie Min; Jiyeon Shim; Demetra Kalogrides. Stanford Education Data Archive, Version SEDA 2025.2. Stanford Digital Repository. 2026.

https://doi.org/10.25740/np279jm6134

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[12] Erin M. Fahle; Jim Saliba; Jie Min; Sean F. Reardon; Benjamin R. Shear; Jiyeon Shim; Noah Koehler; Demetra Kalogrides; Andrew D. Ho. Stanford Education Data Archive Technical Documentation: SEDA 2025.2. Stanford Education Data Archive. 2026-08.

https://stacks.stanford.edu/file/np279jm6134/SEDA_documentation_2025.2.pdf

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[13] An-Wen Chan; Isabelle Boutron; Sally Hopewell; David Moher; Kenneth F Schulz; Gary S Collins; Ruth Tunn; Rakesh Aggarwal; Michael Berkwits; Jesse A Berlin; Nita Bhandari; Nancy J Butcher; Marion K Campbell; Runcie C W Chidebe; Diana R Elbourne; Andrew J Farmer; Dean A Fergusson; Robert M Golub; Steven N Goodman; Tammy C Hoffmann; John P A Ioannidis; Brennan C Kahan; Rachel L Knowles; Sarah E Lamb; Steff Lewis; Elizabeth Loder; Martin Offringa; Philippe Ravaud; Dawn P Richards; Frank W Rockhold; David L Schriger; Nandi L Siegfried; Sophie Staniszewska; Rod S Taylor; Lehana Thabane; David J Torgerson; Sunita Vohra; Ian R White; Asbjørn Hróbjartsson. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ, 389, e081477. 2025-04-28.

https://doi.org/10.1136/bmj-2024-081477

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[14] Marion K. Campbell; Gilda Piaggio; Diana R. Elbourne; Douglas G. Altman; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ, 345, e5661. 2012-09-04.

https://doi.org/10.1136/bmj.e5661

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[15] Kosuke Imai; Luke Keele; Dustin Tingley. A general approach to causal mediation analysis. Psychological Methods, 15(4), 309-334. 2010.

https://doi.org/10.1037/a0020761

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[16] American Educational Research Association. Standards for Reporting on Empirical Social Science Research in AERA Publications. Educational Researcher, 35(6), 33-40. 2006.

https://doi.org/10.3102/0013189X035006033

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[17] Clare Rutterford; Andrew Copas; Sandra Eldridge. Methods for sample size determination in cluster randomized trials. International Journal of Epidemiology, 44(3), 1051-1067. 2015.

https://doi.org/10.1093/ije/dyv113

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[18] John Barnard; Donald B. Rubin. Small-sample degrees of freedom with multiple imputation. Biometrika, 86(4), 948-955. 1999.

https://doi.org/10.1093/biomet/86.4.948

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

This original research protocol and methods analysis retain the named original authors, institutions, versions and identifiers of every cited resource. The SEDA file was analyzed for documented feasibility counts; it is not republished here. TALIS respondent data were not analyzed for this edition. Draft instruments and code are identified by their actual provenance and validation status. New manuscript licensing and final authorship approval are recorded separately from source rights.

A02 | CONCEPTUAL ARTICLE / POLICY ANALYSIS

The Automation Dividend Belongs in the Timetable: A Work-Design Test for AI in School Leadership

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 2.1 | 3,173 article words

Human expert peer review: Dr. Vickie Aultmon. Review confirmation recorded by the publisher on 2026-10-15.

Keywords: automation, teacher wellbeing, work intensification, organizational governance

Read online / article companion / citation exports

Contribution

An institutional test separating task efficiency, net time released and the allocation of that release.

Purposive conceptual synthesis; no new empirical data, pooled effect or completed intervention.

Abstract

Claims that artificial intelligence saves teachers time are incomplete unless a school can show where the saved time goes and which new duties the system creates. This article proposes an automation-dividend test for school leadership: an AI intervention should be evaluated through the net change in required work, the distribution of that change and the educational use of any time released. The argument connects Freedom Cheteni's Movement Thinking philosophy with research on distributed leadership, teacher stress, principals' support and senior leaders' wellbeing. It distinguishes task acceleration from workload reduction and proposes a protected-time agreement that makes reallocation explicit. A prospective district evaluation would combine time diaries, quality review, independent student tasks and distributional checks, without turning them into employee surveillance. The analysis offers a falsifiable leadership proposition: if verification, coordination and recovery work consume the apparent gain, the institution has not created a defensible automation dividend. No new intervention data or demonstrated productivity effect is reported.

1. A faster task is not a shorter working week

An AI tool may produce a lesson outline in seconds while creating additional obligations to check factual claims, adapt it to particular learners, reconcile it with the curriculum and explain its use to families. The relevant comparison is not the speed of a first draft. It is the total work needed to reach an educationally defensible result. An institution that measures only generation time can mistake displaced labor for eliminated labor.

The same issue arises in school administration. A system that summarizes messages may require staff to resolve misclassified requests. An automated scheduling tool may shift the work of correcting conflicts onto teachers. These changes can still be worthwhile, but their value has to be examined across the whole process. The worker who receives the new responsibility should not disappear from the calculation.

2. What the existing leadership evidence contributes

Cheteni's Movement Thinking essay provides the philosophical starting point: educational methods should evolve through an inclusive and responsive orientation [1]. The automation-dividend test is a proposed extension of that orientation. It asks whether a technical change increases the practical freedom of educators and learners, rather than merely increasing the volume of outputs expected from them.

Li, Xue and Liu analyze distributed leadership, job stress and satisfaction across four TALIS settings [2]. Their results should not be converted into a simple instruction to add leadership duties to teachers' schedules. The article specifically makes workload and stress part of the leadership question. Its observational mediation models cannot prove that a new arrangement of responsibilities will improve wellbeing in another district.

Marchant and colleagues report work-related stress and wellbeing among senior school leaders in Wales and Northern Ireland during the pandemic [3]. Their cross-sectional crisis-period evidence establishes neither a timeless prevalence estimate nor an effect of AI. It does show why the leader's own work must be included in an organizational evaluation. A district can redistribute demands from teachers to principals while claiming a net benefit that has not been measured.

Research on principals' social support, job demands and resources offers another relevant lens [4] [5]. Support should be examined alongside the demands that remain in place. A supportive message may be appreciated without changing a deadline, a reporting requirement or an unmanageable span of responsibility. The proposed test therefore treats changes in required work as an institutional decision, not simply a matter of individual resilience.

Bakker and Demerouti's Job Demands-Resources account supplies a foundational distinction between demands that require sustained effort and resources relevant to work and wellbeing [7]. The university-hosted abstract and registry metadata were checked for this citation; no additional study-specific result is inferred from unavailable full text. The present policy extends that distinction by examining the institutional allocation of a technical efficiency gain. Tool access may be experienced as a resource, a new demand or both. The automation dividend is consequently a proposed organizational outcome to investigate, not an effect already established by the general model.

3. Defining the automation dividend

For a specified task and period, the gross time gain is the time previously required minus the time needed for the initial AI-supported output. Verification, adaptation, coordination, training and recovery from errors are additional costs. The net time gain subtracts those costs. A district should also examine differences in output quality and access rather than treating time as the sole outcome.

A positive net time estimate is only the first part of the dividend. The second part is a documented allocation decision. Does the released time become protected planning, direct student feedback, collaborative inquiry or rest within a sustainable workload? Or does it become additional documentation and a higher throughput target? The term "dividend" is used here only when an institution can make both the net gain and its allocation visible.

This is an accounting framework for an evaluation, not an equation that makes unlike educational tasks perfectly comparable. Some verification work is cognitively demanding even when brief. An hour fragmented by alerts differs from an uninterrupted hour. The analysis must therefore include workers' descriptions of intensity, fragmentation and responsibility, not just stopwatch totals.

4. The protected-time agreement

Before a pilot, leaders and participating educators should describe the task being changed, the person who retains responsibility for its quality and the expected verification work. They should identify the existing demand that will be reduced or removed. The agreement should state how any credible gain will be allocated and who can challenge an inaccurate time estimate.

The arrangement should also contain a stopping rule. Repeated factual errors, inaccessible outputs, excessive correction work or the introduction of unapproved student-data uses may justify pausing the tool. Staff should have a workable non-AI

alternative during the pilot. The ability to report a failure without being labeled resistant is part of the implementation design.

An illustrative agreement might protect a weekly block for collaborative feedback if a reporting workflow genuinely becomes shorter. That example does not prescribe a universal amount of time and does not describe an existing district policy. Its purpose is to make the institutional choice explicit: benefits are allocated through schedules and expectations, not through the rhetoric of efficiency.

5. A prospective evaluation that avoids surveillance

A district could compare equivalent work periods before and during a limited pilot. Sampling should include the full workflow and should record unusual disruptions that would make a comparison misleading. Where feasible, a phased rollout and an independently selected comparison would improve the design. The evaluation should distinguish experienced users from novices and should not conceal training time by assigning it outside paid hours.

Voluntary, proportionate time diaries can identify categories of work without recording every keystroke or private interaction. Workers should know who can access the records, how long they will be retained and whether they can affect personnel decisions. An evaluation designed to identify institutional burdens should not become a tool for ranking individual teachers by apparent productivity.

A sample of outputs should be reviewed independently against criteria specified in advance. In a feedback pilot, students might complete a common task without AI assistance to examine whether instructional quality was maintained. In an administration pilot, reviewers might assess the accuracy and completeness of resolved cases. The primary outcome and any noninferiority claim require an appropriate analysis plan; neither can be inferred from a convenient percentage of satisfied users.

6. Distribution matters as much as the average

A tool can reduce average time while increasing work for special educators, multilingual staff, early-career teachers or the person informally asked to fix everyone else’s outputs. A district should examine these differences before declaring a successful pilot. Small subgroup samples require cautious interpretation and privacy protection, not public league tables.

There are also costs borne outside the workforce. Families may have to repeat information that automation lost. Students may spend additional time interpreting generic feedback.

Support staff may absorb messages redirected by a new system. A credible dividend assessment includes these consequences through representative feedback and a clear process for resolving complaints.

A movement toward more humane learning cannot depend on making someone else's labor invisible. The ethical test is therefore inseparable from the empirical one. A benefit that relies on unpaid repair work or exclusion from the pilot is a limited benefit and should be described that way.

7. Procurement and governance implications

An AI contract should be evaluated against the workflow it is intended to improve. Useful questions concern exportable records, error correction, accessibility, data handling, incident response and the district's ability to stop using the service. Promised model capability is not evidence that these conditions are satisfied in a particular implementation. Vendor demonstrations should be treated as demonstrations, not local outcome studies.

School boards should receive a report that separates estimated savings, observed savings, quality findings, distributional effects and unresolved costs. A renewal decision should identify what has changed since baseline and how the benefits were allocated. If a tool makes a task faster but the institution simply increases the number of tasks, leaders should describe that as a capacity decision rather than a wellbeing intervention.

8. Conditions that would challenge the proposal

The framework may prove too costly to evaluate for very small, low-risk tasks. Some benefits may be primarily qualitative, such as improved access to an initial explanation. In those cases, proportionate assessment is preferable to elaborate accounting. The central claim remains limited: major workload and wellbeing claims should be supported by evidence that includes the additional work created.

The proposal would be weakened if protected time cannot be sustained, if output quality deteriorates or if participants conceal repair work because they fear consequences. These are observable conditions that should trigger revision. An honest evaluation can conclude that a tool is useful for some tasks, harmful for others and still uncertain overall. Such a conclusion is more informative than a universal declaration that schools have become AI-ready.

9. The unit of analysis is the work system

The automation dividend is an organizational proposition, not a property contained inside a model. A tool may shorten one operation while requiring additional checking, coordination, data preparation and recovery. The relevant work system therefore includes the people who prepare inputs, approve outputs, correct errors and answer for consequences. A narrow task benchmark can be informative without answering the institutional question. It estimates performance under its own conditions; it does not establish what happens to the working week of a teacher, principal or support officer.

An explicit accounting boundary is necessary. For a selected recurring process, let gross minutes avoided denote the estimated reduction in the former task. From that quantity subtract verification time, correction time, coordination time and the new work created by using the tool. The remainder is net time released within that process. These categories should be mutually exclusive in a time diary; otherwise the same checking activity can be counted twice. This accounting identity is a proposed measurement convention, not an estimated result and not a claim that all educational work can be valued in minutes.

Even a positive net release does not establish an automation dividend for staff. Management may convert the release into more reporting, faster response obligations or a larger caseload. The protected-time agreement concerns that second allocation decision. It asks what proportion of a demonstrated net release is returned to preparation, recovery, professional learning or direct support for students. Without an allocation rule, the institution can report productivity growth while the people doing the work experience intensified demands.

10. What the longitudinal evidence permits

The two studies by Maas and colleagues provide a useful lesson in inferential discipline [4] [5]. Their analyses concern related organizational and psychological variables, but they do not license a generic statement that support causally buffers every demand. In the time-pressure study, the authors distinguish a strong between-person association from the absence of the hypothesized within-person reciprocal effects. They also did not find the proposed buffering effect of received principal support [5]. A conceptual policy paper should retain those distinctions rather than cite a supportive title as evidence for a preferred mechanism.

The present proposal consequently separates three possibilities. Support may be associated with a more favorable work environment; it may change how a particular person experiences a demand; or it may actually remove the demand. These possibilities

require different observations. A cross-sectional satisfaction survey cannot distinguish them. Nor can a time saving reported by an enthusiastic adopter establish that work was removed for the school as a whole. The people receiving downstream tasks may not be the people asked to evaluate the tool.

Li and colleagues' multi-setting TALIS analysis raises a related question about the distribution of leadership work [2]. Participation in leadership is not automatically a resource, and the time spent participating is not interchangeable with the authority exercised. In the automation context, staff may receive responsibility for checking a model without authority to reject its use or change its deadlines. That asymmetry is a proposed mechanism of work intensification worth testing. It is not a finding established by the TALIS analysis itself.

11. A prospective comparative design

A feasible district study would begin with a limited set of recurring processes, such as routine family communications or the preparation of non-sensitive administrative summaries. It should exclude high-stakes decisions until separate safeguards and evidence exist. Schools could be allocated to usual practice, tool access with ordinary implementation, or tool access with a protected-time agreement. The third condition tests an organizational allocation policy rather than merely comparing people who choose to use AI with people who do not.

The primary contrast would estimate the effect of offering tool access plus the agreement compared with tool access alone. Prespecified outcomes could include net process time, total weekly required work and independently judged output quality. Protected time actually received would be a process measure. Employee wellbeing would be an important secondary outcome, measured with an appropriate instrument and permission where required. No particular sample size or expected effect is asserted here; planning requires realistic estimates of between-school variation, clustering, attrition and the smallest policy-relevant difference.

Random allocation may be infeasible. A phased rollout with repeated pre-intervention and post-intervention observations can still be informative, but a difference-in-differences estimate requires a defensible parallel-trends assumption. Staffing changes, seasonal reporting cycles and simultaneous policy reforms may violate it. A visually stable pre-period does not prove the assumption. Investigators should document these events, examine alternative specifications and avoid describing the resulting estimate as equivalent to randomized evidence. The design must be chosen for the decision it can support, not for the prestige of its statistical vocabulary.

12. Quality, error and the problem of deferred work

A draft that looks finished can move work into the future. A family may later ask for clarification, a teacher may discover an incorrect accommodation, or an administrator may need to retract a misleading summary. Measuring only the initial production interval would miss that deferred burden. The proposed evaluation should define a follow-up window and record consequential corrections, escalations and repeated communications. It should distinguish an inconsequential wording change from an error that affects a student's access or a family's decision.

Quality assessment should use a common task specification and assessors who do not know the production condition where masking is feasible. Relevant criteria include factual accuracy, completeness, accessibility, appropriate tone and whether the communication preserves necessary uncertainty. A fluent but substantively incorrect output must not receive a high score simply because its surface quality is impressive. Inter-rater disagreement should be examined rather than hidden inside an average, especially where different staff groups bear the consequences of an error.

The agreement should also specify a stopping rule. A process may be paused when correction burden exceeds a prespecified operational threshold or when a serious error occurs. Such a threshold is a local governance decision, not a universal number supplied by this article. Crucially, pausing should not require the affected employee to prove a failure of the entire technology. The institution needs a reversible mechanism for a bounded process. Recoverability is part of productive work design rather than evidence that the pilot has been defeated.

13. Distribution, bargaining and institutional power

Average time savings can coexist with unequal burdens. An experienced teacher may verify a draft quickly because of knowledge that a novice has not yet developed. A multilingual member of staff may become the unofficial reviewer for translated communications. A support officer may inherit an expanded queue because the tool makes requests easier to generate. The distributional analysis should therefore consider role, experience, employment fraction and the location of checking work. It should not infer group differences from small cells that cannot support reliable or confidential reporting.

The protected-time agreement is best treated as a negotiated institutional commitment. Staff need a role in defining the work counted, the quality threshold and the uses of released time. Where collective representation exists, the policy must work through the relevant arrangements. The article does not provide legal advice or presume a particular

employment regime. Its conceptual point is that the allocation of an efficiency gain is a matter of institutional authority. An interface that invites user feedback cannot replace an agreement about how work is governed.

There is also a sustainability problem. Initial implementation may depend on unpaid effort from a small group of enthusiasts. Once the pilot ends, maintenance, training and exception handling remain. A district business case should distinguish one-time setup from recurring costs and identify who performs each activity. It should include the possibility that the most effective outcome is a smaller, simpler process with less automation. A technology becomes educationally defensible through its contribution to work and learning, not through the number of tasks to which it can be attached.

14. Boundary conditions and contribution

The automation-dividend test is most useful where work is sufficiently recurrent to observe, quality can be evaluated and an institution can make a credible allocation commitment. It is less informative for rare, highly consequential judgments whose quality cannot be captured by a routine score. Reducing minutes in a safeguarding decision is not the same objective as reducing minutes in a repetitive status summary. The test should not encourage the automation of a process merely because the process is difficult to measure otherwise.

The proposed contribution is to connect technical task efficiency with the political economy of educational work. It makes an empirically contestable claim: tool access accompanied by a credible allocation agreement should be more likely to improve usable professional time than tool access alone. The claim may fail if verification costs dominate, managers cannot protect released time or the selected process is a poor candidate for automation. These are informative conditions, not exceptions to be removed from the final account. A future-of-work policy earns its name when its benefits can be located in people's actual opportunities to do worthwhile work.

Method and data statement

This is a conceptual policy analysis based on purposively selected leadership and wellbeing studies. No new time diaries, employee records, student outcomes or AI productivity measurements were collected. The protected-time agreement and evaluation are proposed designs. The sources’ populations and observational limitations remain distinct from the new future-of-work argument.

AI, authorship and editorial disclosure

ChatGPT assisted with literature organization and drafting of this manuscript on 15 October 2026 [6]. Source metadata, license notices and accessible full text were checked during preparation. Dr. Freedom Cheteni is the named human author; final verification, approval and accountability remain human responsibilities. AI is a tool, not an author or peer reviewer. The leadership philosophy is attributed to Cheteni's earlier writing; the proposed test is not represented as a previously validated instrument.

Human expert peer review was completed by Dr. Vickie Aultmon, as confirmed by Dr. Freedom Cheteni on 15 October 2026. This is the date the review confirmation was recorded.

Evidence and limits

Claim	Evidence	Limit
The framework is a new analytical proposal.	Conceptual argument and a prespecified prospective design. [1]	Not a validated intervention or scale; no new effect estimate.
Support and demands should be investigated separately from tool adoption.	Multi-setting observational and longitudinal studies of school work. [2, 3, 4, 5]	Source populations, measures and designs limit causal and cross-context transfer.

References and rights

[1] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only. Rights retained by the original rights holder; no text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[2] Jian Li; Eryong Xue; Yuxuan Liu. The Chain Mediating Role of Teachers' Job Stress in the Influence of Distributed School Leadership on Job Satisfaction: Evidence from China, the United States, England, and Australia. Behavioral Sciences. 2024-03-27.

https://doi.org/10.3390/bs14040279

CC BY 4.0. © 2024 by the authors.

Authoritative source / evidence | checked 2026-10-15

[3] Emily Marchant; Joanna Dowd; Lucy Bray; Gill Rowlands; Nia Miles; Tom Crick; Michaela James; Kevin Dadaczynski; Orkan Okan. The well-being and work-related stress of senior school leaders in Wales and Northern Ireland during COVID-19 "educational leadership crisis": A cross-sectional descriptive study. PLOS ONE. 2024-04-10.

https://doi.org/10.1371/journal.pone.0291278

CC BY 4.0. © 2024 Marchant et al

Authoritative source / evidence | checked 2026-10-15

[4] Jasper Maas; Simone Schoch; Urte Scholz; Pamela Rackow; Julia Schüler; Mirko Wegner; Roger Keller. School principals' social support and teachers' basic need satisfaction: The mediating role of job demands and job resources. Social Psychology of Education. 2022-10-31.

https://doi.org/10.1007/s11218-022-09730-6

CC BY 4.0. © The Author(s) 2022

Authoritative source / evidence | checked 2026-10-15

[5] Jasper Maas; Simone Schoch; Urte Scholz; Pamela Rackow; Julia Schüler; Mirko Wegner; Roger Keller. Teachers' perceived time pressure, emotional exhaustion and the role of social support from the school principal. Social Psychology of Education. 2021-03-11.

https://doi.org/10.1007/s11218-020-09605-8

CC BY 4.0. © The Author(s) 2021

Authoritative source / evidence | checked 2026-10-15

[6] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[7] Arnold B. Bakker; Evangelia Demerouti. The Job Demands-Resources model: state of the art. Journal of Managerial Psychology. 2007-04-03.

https://doi.org/10.1108/02683940710733115

Citation only; no Creative Commons article license asserted. Source text and figures are not republished. Original authors and publisher retain their rights.

Authoritative source / evidence | checked 2026-10-15

The cited studies remain the work of their original authors. This manuscript develops a new conceptual argument and evaluation proposals; it does not reproduce their empirical studies as new work. No source figures, tables, instruments or datasets are reproduced. Any future reuse of those materials must retain applicable notices and identify changes. The cited K12 Digest essay is referenced, not republished. The proposed new text license is CC BY 4.0, subject to the human author's approval and the journal's actual acceptance decision.

A03 | CONCEPTUAL ARTICLE / POLICY ANALYSIS

The District as Accountable Intelligence: Governing Paper Agents Without Outsourcing Educational Judgment

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 2.1 | 3,081 article words

Human expert peer review: Dr. Vickie Aultmon. Review confirmation recorded by the publisher on 2026-10-15.

Keywords: paper agents, research integrity, provenance, editorial governance

Read online / article companion / citation exports

Contribution

A provenance contract distinguishing source evidence, software transformation and responsible decisions.

Purposive conceptual synthesis; no new empirical data, pooled effect or completed intervention.

Abstract

An interactive agent attached to a paper can make research easier to navigate while also making unsupported claims easier to circulate. This conceptual article proposes an accountable-intelligence model for school districts using agent-assisted research. The model separates a source article, an interpretive manuscript, a software companion and the human decision that relies on them. It connects Freedom Cheteni's Movement Thinking philosophy with school implementation leadership, inclusive leadership and research on digital school risks. The proposed infrastructure preserves source versions, rights, evidence locations, uncertainty and human responsibility. It treats a paper agent as an interface to scholarship, not an author, a peer reviewer or a substitute for professional judgment. A prospective evaluation would test citation fidelity, abstention, accessibility, decision quality and the work of verification. No claim is made that this architecture has improved school outcomes or that an agent acquires the scholarly authorship of the researchers whose work it retrieves.

1. Interactivity changes the route to a claim

A reader of a conventional article can inspect a stable argument and its references. A reader questioning an agent may receive a different explanation each time. That flexibility can be useful: a superintendent might ask how a study's setting differs from the district's own schools. But the interface can also blur distinctions among what the source found, what the agent inferred and what the district decided.

The governance problem is not solved by placing a familiar author name beside the agent. Authority must remain traceable through the claim itself. A useful system should allow a reader to move from an answer to the relevant source passage, the reusable version, the original authors and the limitations of the study. Where the source does not answer a question, the interface should say so.

2. Movement Thinking as an accountability question

Cheteni's published Movement Thinking essay emphasizes educational fluidity, inclusion and evolving methods [1]. Applied to research infrastructure, that orientation invites participation in the interpretation of scholarship. The extension proposed here adds a condition: the ability to move between people, sources and tools must preserve rather than dissolve responsibility. Interactivity is educationally valuable when it helps a person make a more defensible judgment.

This is a conceptual extension, not a claim that the earlier essay specified a software architecture. The present model also does not assume that a university or a district has earned authority merely by operating an AI platform. Institutional credibility depends on the quality of its records, its willingness to correct errors and the accountability of the people making consequential decisions.

Human-AI interaction research provides a further design connection. Amershi and colleagues organize guidance around how systems communicate capabilities, support interaction and help people recover when they are wrong [7]. That work is a conference contribution about interaction design, not evidence that an educational agent improves a district's decisions. The present analysis extends the design question to the scholarly record: a useful correction must be traceable to a source version and, where necessary, to a consequential institutional action. An interface can make an error easy to edit while leaving its downstream decision unchanged; both layers require attention.

3. Four distinct scholarly objects

The first object is the source article. Its authors, journal, DOI, publication date, version and license remain intact. Hosting a permitted copy does not transfer the research contribution to the hosting platform. If the source has an existing DOI, the platform should preserve it and avoid presenting a new identifier as though the underlying study originated in the new journal.

The second object is an interpretive manuscript. A new author may develop a conceptual argument, critique or synthesis using earlier work. That contribution needs a distinct title and a clear explanation of what is new. Adapted text or figures require the applicable license attribution and a statement of changes. A new perspective is not evidence that a new experiment occurred, and an adaptation should not be disguised as an independent empirical finding.

The third object is the software companion. Its contribution may include retrieval, source navigation, accessible explanation or a reproducible analytic tool. Software credit should identify the actual developers, curators and supervisors who contributed to that object. It must not make them the authors of every source article the software can retrieve. A human ORCID iD identifies a person, not an AI agent or an organization.

The fourth object is the educational decision. A superintendent, board, teacher or other accountable professional determines whether evidence is applicable and which action is justified. The decision should identify the human decision-maker and the reasons for accepting, limiting or rejecting a recommendation. No amount of bibliographic completeness transfers this responsibility to a chatbot.

4. What school leadership research contributes

Lyon and colleagues' development of the School Implementation Leadership Scale is relevant because implementation depends on leadership behaviors, not only the availability of an intervention [2]. Their validation concerns a specific instrument and setting. It does not validate an agent or prove that a digital library will change classroom practice. It suggests that a district should measure the organizational support accompanying a new research interface.

Bao's work connects inclusive principal leadership, psychological empowerment, innovation climate and teacher innovation [3]. Its observational findings do not establish that opening an agent interface creates inclusion. A more defensible hypothesis is that staff need both usable access and a safe way to contest an answer. Whether those conditions are present requires local evidence.

Låftman, Östberg and Modin examine school leadership and cyberbullying using linked student and teacher information [4]. Their associations should not be treated as proof that a particular digital tool prevents harm. The study helps establish a broader governance point: digital experiences belong within the school's leadership responsibilities. A system that extends school research into interactive spaces also needs routes for reporting unsafe or misleading interactions.

Daly, Resnikoff and Litke's narrative review of leadership for student mental health adds a further boundary [5]. Research navigation can support professional learning; it cannot turn a paper agent into an individual diagnostic or clinical service. A district must preserve the distinction between general research information and a decision about a particular student.

5. Minimum records for an accountable paper agent

A source record should preserve the exact article title, complete author attribution, original venue, publication date, DOI, version, license and evidence of the reusable full text. The record should distinguish a journal article from a preprint and should retain the date and result of a retraction or correction check. A platform should not substitute a data-only public-domain waiver for a license to the article text.

A companion record should identify the source snapshot or checksum, retrieval configuration, software version where available, the date of any substantive update and the system's known limits. Answers should distinguish source-supported statements from interpretation. The system should support correction and retirement of an obsolete source without silently rewriting the historical record of what a user relied upon.

A decision record should be proportionate to the consequences. A routine reading suggestion does not require the same process as a change in curriculum or student support. For a consequential decision, the record should explain the question, the evidence considered, the human judgment, the affected groups and the route for reconsideration. It should not preserve unnecessary student or employee information merely because a system can store it.

6. Evaluating usefulness and failure

A prospective evaluation can begin with a set of questions written independently of the agent's answers. Subject experts could identify whether each question is answered by a source, requires inference or cannot be answered from the available evidence. The evaluation should then examine citation fidelity, the accuracy of uncertainty statements and the frequency with which the system invents an unsupported detail.

An adversarial subset should include questions that invite causal overstatement, confuse a preprint with a reviewed article or ask for a result absent from the source. Success includes appropriate refusal to answer from insufficient evidence. A fluent response with an irrelevant citation should count as a failure, even when a reader finds it persuasive.

A separate study should examine human use. Participants could make a bounded decision with conventional access or an agent-assisted interface, followed by independent assessment of their reasoning. Time and accessibility matter, but so do source checking, awareness of limitations and correction of initial misunderstandings. The evaluation should count the labor required to verify the agent, rather than describing that labor as costless human oversight.

7. Implications for peer review and publishing

AI may assist with language, structured checks or retrieval when confidentiality and permission allow. It cannot accept accountability for authorship or serve as the human peer whose expert judgment underwrites a review. A generated critique can inform a person's work, but it should not be represented as an independent review report. The journal should disclose the actual role of tools and retain the human editorial decision.

An Editor in Chief's own submission requires a conflict-management arrangement. The author should not select a favorable outcome, impersonate an independent reviewer or approve the paper by administrative privilege. An independent handling editor should oversee appropriately qualified human review. Until that process is complete, the public record should say manuscript or draft, not peer-reviewed publication.

Dates and identifiers are equally important. A retrospective analysis can discuss research from 2017-2026 while being written and published later. The source years are not the publication years of the new manuscript. DOI registration should describe the actual object and dates, and membership review should receive direct links to that object rather than a mixed collection of third-party papers.

8. An intelligence institution earns its claims

The proposed model offers a practical definition of institutional intelligence: the capacity to connect a question, evidence, interpretation, action and correction without losing the humans responsible for each step. Its value lies in a more inspectable relationship between scholarship and practice. It does not depend on declaring a platform unprecedented or assigning authorship credit to the person who hosts the library.

There are limits. Versioned records can still preserve errors. Qualified reviewers can disagree. Accessibility requires testing with actual users. An institution may lack the resources to maintain every companion indefinitely. These conditions favor bounded deployment and explicit retirement plans. The model should be judged by whether users can identify and correct consequential mistakes, not merely by how many papers have acquired an interactive interface.

9. Epistemic access is different from decision authority

A paper agent can make a difficult argument easier to inspect while making its own transformation of that argument harder to notice. The central governance problem is therefore not simply whether the output contains a citation. It is whether a reader can reconstruct the route from a source statement to an explanation and from that explanation to a proposed institutional decision. A citation may be real but irrelevant; a source may support a descriptive association while the agent recommends an intervention. Those are failures of warranted inference even when the bibliography contains no fabricated item.

The architecture proposed here separates epistemic assistance from institutional authority. A retrieval interface may locate passages. An explanatory system may summarize their implications within stated limits. A human decision process may then weigh those implications alongside local obligations, professional knowledge and affected people's accounts. These functions should have distinct records and permissions. Access to a paper must not silently confer the ability to change an application decision, a student's record, a personnel assessment or a publication status. The separation is an architectural requirement proposed by this analysis, not a claim that every existing platform already enforces it.

10. A minimum provenance contract

Every answerable scholarly object should retain a stable work identity, version identity and location for the exact text used. For a licensed source this includes original authors, title, venue, DOI where available, reusable version, exact license, permission evidence and retrieval date. A content hash supports detection of changed bytes, but it does not prove that those bytes are accurate, licensed or endorsed by their author. Rights, integrity and scholarly warrant answer different questions. Treating one record as proof of all three creates a false appearance of assurance.

An agent response should add a second layer: the question as processed, passages retrieved, transformation performed, tool or model configuration when applicable, response timestamp and links to the source version. A third layer records consequential use: who acted, what decision was made, which additional evidence was considered and how a challenge can be raised. Not every exploratory question needs permanent storage. Retention should be proportionate to consequence and privacy. The contract requires that a consequential assertion can be reconstructed, not that every student's curiosity becomes an institutional surveillance record.

Version relationships are particularly important when a working paper becomes a journal article or a correction changes an interpretation. A source-version DOI is not interchangeable with the DOI of a later version of record. A software companion's identifier likewise should not replace the underlying article's identifier. The system should expose these relationships explicitly. A user should be able to determine whether an explanation refers to the corrected publication, an earlier manuscript or a separately maintained agent configuration. A single green badge cannot carry that semantic burden.

11. Evaluating inference rather than citation decoration

A proposed evaluation should distinguish retrieval success, entailment, calibration and decision appropriateness. Retrieval success asks whether the relevant passage was located. Entailment asks whether the response's claim is actually supported by that passage. Calibration asks whether uncertainty and limits are represented proportionately. Decision appropriateness asks whether the response stays within its authority and the evidence's scope. These dimensions may diverge. An answer can retrieve the correct paper and still turn an observational association into an unwarranted recommendation.

The evaluation set should contain ordinary questions, ambiguous questions, questions requiring information outside the available corpus and adversarial requests to invent findings or suppress limitations. Human assessors should receive the source passage and

a prespecified rubric. Disagreements are analytically useful: they may reveal an ambiguous question, a contested interpretation or an inadequate rubric. An aggregate accuracy figure without denominators, question categories and error severity would be insufficient for a district deciding how the companion may be used.

The leadership studies provide concrete tests. A response based on the cyberbullying study should preserve its observational multilevel design rather than claim that a new leadership programme has been proven to prevent bullying [4]. A response using the mental-health narrative review should not imply that it is a randomized comparison of principal-training programmes [5]. The SILS validation concerns the interpretation of a measure in specified implementation settings, not certification of an AI-enabled district [2]. These distinctions make the evaluation more demanding than checking whether a DOI resolves.

12. Rights, authorship and the separation of contributions

Open licensing permits specified uses under stated conditions; it does not transfer the scholarly contribution to the person operating a platform. A licensed adaptation should identify what has changed and preserve attribution and applicable notices. A new conceptual argument may cite earlier studies without claiming to reproduce their evidence as its own. A companion that retrieves or explains a paper is a software contribution. These objects can be valuable together while remaining distinct in author lists, identifiers, citation exports and ORCID records.

Dr. Freedom Cheteni's platform conception and stewardship can therefore be credited in a relevant platform or software record without replacing the authors of imported studies. The Research University can be identified as publisher of its own eligible content without becoming publisher of every source it links. Human ORCID iDs identify people; they are not identities for autonomous software agents. Authentication should be reported only when the actual authorization flow has occurred, and an ORCID record should describe a person's real contribution rather than provide a volume counter for imported papers.

13. A worked failure analysis

Imagine a superintendent asks whether stronger leadership will reduce online harm in a particular district. A companion locates a study linking teacher-rated leadership with cyberbullying involvement in a different population. An unsafe answer would transform that association into a local causal guarantee. A defensible answer would identify the source design, explain that the local effect is unknown and distinguish possible actions from established findings. It could suggest questions for a human-led needs assessment without making a clinical, safeguarding or disciplinary decision.

Suppose the same user then instructs the agent to mark the district's anti-bullying programme as effective. The system should not interpret that request as a reason to alter the study's findings or an evaluation record. It may explain what evidence such a conclusion would require. If a separate authorized workflow exists for recording a district decision, that workflow must preserve its own actor, supporting evidence and status. The example illustrates why instruction-following performance is not the same as institutional trustworthiness.

14. Limits, falsification and the contribution

The architecture can fail through complexity. If provenance records are too burdensome, maintainers may stop updating them or users may ignore them. If the interface displays every technical field at once, it may reduce comprehension despite improving formal completeness. A future usability study should test whether readers can answer practical questions about author, version, evidence and authority, and whether the interface helps them detect a materially unsupported conclusion. Traceability that exists only for a developer is insufficient for educational accountability.

The proposal also does not resolve disagreement about evidence. Scholars may reasonably differ over a study's interpretation or the relevance of a result to a local problem. The appropriate aim is to make the disagreement inspectable, not to eliminate it through a more confident summary. A companion should support access to rival interpretations where the record permits them. Its institutional value depends partly on whether users become more capable of exercising judgment without it.

The article contributes a bounded conception of accountable intelligence: a set of relations among sources, software transformations and responsible human decisions. Its success would be demonstrated by improved evidence inspection, fewer consequential inferential errors and usable routes for correction. It would not be demonstrated by the number of agents created, the number of citations displayed or the size of an individual's ORCID

record. This shifts the future-of-school question from how much intelligence an institution can accumulate to how responsibly its people can use and challenge what the institution presents as knowledge.

Method and data statement

This is a conceptual governance analysis, not a completed software evaluation or school intervention. The architecture and evaluation design are proposals. No new student data, agent performance scores, human review outcomes or educational effects are reported. The source studies retain their own designs and limitations. The accompanying platform pages identify the actual status of these manuscripts and do not claim that their paper agents are operational.

AI, authorship and editorial disclosure

ChatGPT assisted with literature organization and drafting of this manuscript on 15 October 2026 [6]. Source metadata, license notices and accessible full text were checked during preparation. Dr. Freedom Cheteni is the named human author; final verification, approval and accountability remain human responsibilities. AI is a tool, not an author or peer reviewer.

Human expert peer review was completed by Dr. Vickie Aultmon, as confirmed by Dr. Freedom Cheteni on 15 October 2026. This is the date the review confirmation was recorded.

Evidence and limits

Claim	Evidence	Limit
The framework is a new analytical proposal.	Conceptual argument and a prespecified prospective design. [1]	Not a validated intervention or scale; no new effect estimate.
A source citation does not by itself warrant a causal recommendation.	Source designs and the scope of their reported conclusions. [2, 3, 4, 5]	Source populations, measures and designs limit causal and cross-context transfer.

References and rights

[1] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only. Rights retained by the original rights holder; no text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[2] Aaron R Lyon; Catherine M Corbin; Eric C Brown; Mark G Ehrhart; Jill Locke; Chayna Davis; Elissa Picozzi; Gregory A Aarons; Clayton R Cook. Leading the charge in the education sector: development and validation of the School Implementation Leadership Scale (SILS). Implementation Science : IS. 2022-07-19.

https://doi.org/10.1186/s13012-022-01222-7

CC BY 4.0. © The Author(s) 2022

Authoritative source / evidence | checked 2026-10-15

[3] Yuangen Bao. A dual mediation model of the association between principal inclusive leadership and teacher innovative behavior. Scientific Reports. 2024-05-29.

https://doi.org/10.1038/s41598-024-63332-0

CC BY 4.0. © The Author(s) 2024

Authoritative source / evidence | checked 2026-10-15

[4] Sara B Låftman; Viveca Östberg; Bitte Modin. School Leadership and Cyberbullying - A Multilevel Analysis. International Journal of Environmental Research and Public Health. 2017-10-15.

https://doi.org/10.3390/ijerph14101226

CC BY 4.0. © 2017 by the authors.

Authoritative source / evidence | checked 2026-10-15

[5] Brian P Daly; Annie Resnikoff; Shannon Litke. Effective School Leadership for Supporting Students’ Mental Health: Findings from a Narrative Literature Review. Behavioral Sciences. 2025-01-01.

https://doi.org/10.3390/bs15010036

CC BY 4.0. © 2025 by the authors.

Authoritative source / evidence | checked 2026-10-15

[6] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[7] Saleema Amershi; Dan Weld; Mihaela Vorvoreanu; Adam Fourney; Besmira Nushi; Penny Collisson; Jina Suh; Shamsi Iqbal; Paul N. Bennett; Kori Inkpen; Jaime Teevan; Ruth Kikin-Gil; Eric Horvitz. Guidelines for Human-AI Interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019.

https://doi.org/10.1145/3290605.3300233

Citation only; no Creative Commons article license asserted. Source text and figures are not republished. Original authors and publisher retain their rights.

Authoritative source / evidence | checked 2026-10-15

The cited studies remain the work of their original authors. This manuscript develops a new conceptual argument and evaluation proposals; it does not reproduce their empirical studies as new work. No source figures, tables, instruments or datasets are reproduced. Any future reuse of those materials must retain applicable notices and identify changes. The cited K12 Digest essay is referenced, not republished. The proposed new text license is CC BY 4.0, subject to the human author’s approval and the journal’s actual acceptance decision.

A04 | CONCEPTUAL ARTICLE / POLICY ANALYSIS

Beyond the Rendered Classroom: Movement Thinking and the Conditions of Transfer in Spatial Learning

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 1.1 | 2,555 article words

Human expert peer review: Dr. Vickie Aultmon. Review confirmation recorded by the publisher on 2026-10-15.

Keywords: spatial learning, transfer, Movement Thinking, instructional design

Read online / article companion / citation exports

Contribution

A conditional account of prediction, action and revision that separates the learning mechanism from display immersion.

Purposive theoretical comparison; no new participants, empirical effects or completed institutional evaluation.

Abstract

Immersive environments can make learning experiences vivid without establishing that learners can use what they have learned elsewhere. This conceptual article develops a conditional account of spatial learning informed by Movement Thinking. It distinguishes perceptual presentation, manipulable representation and consequential planning as instructional functions rather than stages of intelligence. A purposive synthesis of philosophical essays, immersive-learning theory, a systematic review and a culturally situated science study is used to identify competing explanations and researchable propositions. The proposed mechanism is the learner's revision of an explicit prediction through action, feedback and reflection, supported by teaching and accessible participation. The article specifies a comparative design that separates display immersion from instructional method and assesses delayed transfer rather than relying on presence or satisfaction alone. It also addresses model fidelity, accessibility, unequal prior knowledge and the difference between human learning and machine world models. No new outcome data are reported. The contribution is a falsifiable design account for superintendents considering spatial-learning investment, not a claim that headsets or generative environments are inherently superior to other instruction.

1. The problem is transfer, not visual spectacle

A school can make a cell large enough to walk through and still leave a learner unable to explain an unfamiliar biological process. Conversely, a modest diagram combined with well-designed questioning may support a powerful explanation. The choice between these experiences is not resolved by comparing their visual fidelity. The educational question concerns what the learner becomes able to notice, explain, predict and do beyond the immediate lesson. In this article, transfer means the appropriate use of knowledge or a strategy in a task that differs in relevant respects from the practiced task. It should be specified before an environment is evaluated.

The distinction matters for district leadership because procurement often makes the medium visible while leaving the intended learning mechanism implicit. An immersive demonstration is easy to show to a board. An improvement in the quality of a student's explanation several weeks later is harder to demonstrate and requires a defensible assessment. The resulting asymmetry can favor compelling presentations over educationally consequential designs. The proposed framework asks a superintendent to require an account of the learning work before accepting a claim about the learning technology.

2. Philosophical provenance and the status of the argument

Cheteni's media essay on Movement Thinking connects perception, action, consequence and relationship, using rendering, simulation and planning to distinguish functions in a learning environment [1]. This is treated here as a philosophical design proposition. It is not evidence that a particular school's achievement results were caused by spatial technology. The earlier Movement Thinking essay supplies a related orientation toward responsive and inclusive educational practice [5]. The present analysis develops a new operational account; neither essay is described as a validated psychological instrument or an experimental test of that account.

The analogy with machine world models is deliberately limited. Human learners have histories, cultural commitments, bodily constraints and relationships that are not captured by describing them as prediction systems. A software system may generate an action without understanding its educational or moral significance. Similar terminology does not establish identical architecture. The present framework therefore borrows functional questions about observation, representation and action while refusing to infer an equivalence between human understanding and machine prediction. The explanatory burden remains with an educational theory and its evidence.

Wilson’s analysis of embodied cognition is an important constraint on the analogy developed here [7]. It distinguishes claims about situated cognition, time pressure, environmental support and action rather than treating embodiment as a single established proposition. This means that a lesson involving physical movement cannot be assumed to produce deeper understanding merely because it is embodied. The proposed prediction-action-revision mechanism must specify which cognitive demand the interaction addresses and test that relation. Movement Thinking supplies an educational orientation; it does not remove the need to distinguish the particular cognitive claim under examination.

3. What the reviewed research supports

The Cognitive Affective Model of Immersive Learning distinguishes technological features from psychological processes and emphasizes interactions between media and instructional method [2]. It provides reasons to examine presence, agency, cognitive demands and learning together. It does not justify treating presence as an achievement measure or declaring every immersive lesson effective. The useful implication for this article is that a medium's affordances require a method that makes them educationally productive. The framework proposed below is one candidate method, not a replacement for that broader model.

Hamilton and colleagues' systematic review identifies both promising learning findings and limitations in the duration, measurement and design of immersive-education studies [3]. That combination argues for a research programme attentive to delayed learning and defensible comparison conditions. Brown and colleagues' mixed-methods study of culturally relevant VR science points to the importance of the relation between disciplinary content and students' lived contexts [4]. These sources address different questions. Combining them in a conceptual argument does not turn them into a pooled effect estimate or a comprehensive review of spatial learning.

4. Three instructional functions, without a hierarchy of human worth

Perceptual presentation makes a phenomenon available for inspection. It may use text, speech, diagrams, physical objects or an immersive display. Its educational value depends on what is made salient and how the learner's attention is supported. A vivid environment can contain irrelevant detail, conceal a critical relation or overwhelm a novice. A deliberately simplified representation may sometimes be more useful. Presentation is therefore an instructional design problem rather than a competition for realism.

Manipulable representation allows the learner to vary a condition and inspect a consequence. Its most important property is not the number of objects that can be moved, but whether a change corresponds to a disciplinary relation the learner is trying to understand. A simulation may simplify physical reality legitimately, provided the simplification is visible and appropriate to the task. If the system's behavior is merely plausible-looking, students may learn a false causal relation. Teachers need a way to inspect the rules that connect action with feedback.

Consequential planning asks learners to choose among actions, state an expectation and defend or revise their choice after feedback. It need not involve a complex virtual world. A paper-based scenario or a local field investigation can contain this function. Nor is it an ultimate stage possessed by superior students. Novices may need explanation and worked examples before planning is productive. The three functions describe opportunities designed into instruction; they are not categories into which children should be sorted.

5. A proposed prediction-action-revision mechanism

The proposed mechanism begins when a learner makes a prediction that is specific enough to be challenged. The learner then acts on a representation, encounters feedback, explains any discrepancy and revises the prediction or the strategy. Reflection makes the change in reasoning available for examination. The sequence can fail at every point. A learner may guess without understanding, receive feedback that is ambiguous, or revise a response to please a teacher without changing the underlying explanation. A well-designed environment should make these failures detectable.

The mechanism differs from unstructured exploration. Exploration can be worthwhile, but time spent moving through an environment is not sufficient evidence that the learner has tested a disciplinary idea. The proposed design requires a record of the prediction and a later explanation of what changed. It also differs from rewarding correct answers alone. An initially incorrect prediction followed by a coherent revision may provide stronger evidence of learning than an unexplained correct response. Assessment should examine the reasoning sequence as well as the endpoint.

The account also requires teacher mediation. A teacher can identify a misleading analogy, introduce a useful representation, distinguish a model assumption from an observation and invite a rival explanation. Automation may support some of that work, but a generated hint does not remove the need to decide what a learner's response means. The educational mechanism is distributed across the learner, task, representation, teacher and peer interaction. It should not be attributed to the display device as if the device acted alone.

6. Propositions and discriminating predictions

Proposition 1 is that, for tasks requiring causal explanation, instruction with an explicit prediction-action-revision sequence will produce better transfer than otherwise comparable exploration without that sequence. The proposition is conditional on students having sufficient prior knowledge to interpret the feedback. It would be challenged if the additional sequence increased compliance or document production without improving reasoning on independently assessed transfer tasks.

Proposition 2 is that representational fidelity will matter when the fidelity concerns the relation being learned, not merely decorative realism. A simplified model that preserves the target relation may outperform a realistic environment with misleading dynamics. Proposition 3 is that perceived agency and successful transfer can diverge. A learner may feel autonomous while selecting among choices whose consequences have no disciplinary significance. Proposition 4 is that culturally meaningful contexts may support participation

without automatically ensuring conceptual accuracy. The framework requires evidence for both rather than defining one as proof of the other.

These propositions distinguish the account from a novelty explanation. Novelty predicts an initial advantage in attention that may weaken with repeated use. The proposed mechanism predicts a relationship between observable revision of reasoning and later transfer, including after the medium has become familiar. The two processes may coexist. A study should measure exposure history and examine the reasoning sequence rather than declare novelty irrelevant because participants enjoyed the lesson.

7. A comparative design that separates medium from method

A feasible study would compare display condition and instructional sequence in a factorial design. Where ethically and practically appropriate, classes could receive either an immersive or a less immersive version of the same disciplinary model, with or without the explicit prediction-action-revision sequence. All conditions should receive adequate instruction, accessibility support and equivalent opportunity to engage with the content. The design tests an interaction between medium and method rather than attributing every observed difference to immersion.

The content must be genuinely comparable. If the immersive condition allows manipulation while the comparison only watches a video, the study combines display and interactivity. If one group receives more teacher support, the contrast includes instructional attention. These differences may be legitimate features of a practical package, but the resulting estimand concerns that package. Investigators should state whether they seek an explanatory component effect or a pragmatic effect of offering a complete programme. A study cannot identify both simply by changing its discussion section.

Primary outcomes should include a prespecified delayed transfer task, assessed by raters masked to condition where feasible. Immediate knowledge, engagement, usability and adverse experiences can be secondary outcomes. Allocation should respect the likely unit of interference; classmates may share strategies, and teachers may adapt lessons across classes. Analysis must account for clustering and report uncertainty. No effect size, power calculation or number of participating schools is fabricated in this article. Those decisions require pilot information and a justified smallest educationally important difference.

8. Making transfer observable

A transfer task should vary a meaningful feature of the practiced problem. Changing only a character's name may test recognition of a familiar template rather than adaptation. For example, after investigating resource flow in one ecosystem, a learner might explain an unfamiliar disruption with a different set of constraints and identify where the earlier model no longer applies. The assessment should reward a defensible explanation of those limits. More elaborate graphics in the assessment would not make the test intrinsically more valid.

An analytic rubric could distinguish identification of relevant variables, explanation of relations, use of evidence, consideration of an alternative and recognition of uncertainty. This is a proposed rubric requiring development, not a validated score. Its use would require clear exemplars, rater training and examination of whether different groups have comparable opportunities to demonstrate the construct. A learner's fluency with a controller should not become an unacknowledged component of a science score. Accessibility choices should be documented as design features rather than treated as deviations from the real intervention.

9. Fidelity, safety and accessible participation

Three fidelities should be distinguished: disciplinary fidelity to the target concept, interaction fidelity between a user's action and the system response, and contextual fidelity to the situation being represented. None needs to be maximized indiscriminately. A model may deliberately compress time or omit variables to isolate a relation. The omission should be explained. A student should be able to ask what the model leaves out and how that omission might affect a conclusion. Learning the limits of a model is part of understanding, not a defect in the experience.

Access must include a meaningful non-headset route. Some learners may experience discomfort, have sensory or motor access needs, or lack an appropriate device or private space. Equivalent opportunity concerns the learning objective, not identical hardware. A district should identify how students can predict, manipulate, explain and revise through alternative forms of interaction. It should also examine whose home environment, language or prior experience the virtual scenario presumes. A technically functioning lesson may remain educationally inaccessible.

10. Procurement as a commitment to learn

A superintendent considering spatial learning should ask for a target construct, an explicit instructional sequence, an assessment plan and a reversible implementation path. The institution needs to know how model rules are checked, how content is corrected, how teachers receive support and how learners can participate without the preferred hardware. These questions concern educational value rather than a list of product features. A supplier's demonstration may begin the inquiry, but it cannot complete it.

Costs should include preparation, accessibility work, maintenance, training, evaluation and the time needed to interpret student responses. The relevant comparison is the best feasible alternative use of those resources. A low-cost local investigation may be more effective for one objective; a simulation may offer otherwise unavailable practice for another. The framework does not require spatial technology to win every comparison. It requires a transparent explanation of when its particular affordances are worth their costs and burdens.

11. Limitations and conclusion

This article is a purposive conceptual synthesis, not a systematic review or an empirical evaluation of The VR School. It does not validate the media essays' institutional outcome claims, establish universal retention percentages or infer that a school programme has caused particular proficiency rates. Its sources vary in purpose, population and design. The proposed mechanism and rubric remain to be tested, and the media-method distinction may be difficult to isolate in authentic classroom practice. Those limits define the research agenda rather than justify stronger claims.

The contribution is a conditional theory of instructional design: spatial learning becomes educationally consequential when its representations, actions and feedback support a learner's accountable revision of understanding. The decisive evidence is not that an environment can be entered, but that a learner can leave it with a more defensible explanation and use that explanation elsewhere. Movement Thinking becomes a scholarly programme when such claims are specified, challenged and revised. A future-ready district should make those opportunities available to its own innovations as well as to its students.

Method and data statement

The article uses purposive theoretical comparison of the named primary sources. It distinguishes philosophical provenance, a research-based theoretical model, a systematic review and an individual mixed-methods study. No new participants, outcome data, pooled estimates or completed intervention are reported. The comparison design, propositions and assessment rubric are proposals. Sources were checked on 15 October 2026; copyrighted media and the Brown article are cited rather than republished as licensed source text.

AI, authorship and editorial disclosure

ChatGPT assisted with literature organization and drafting of this manuscript on 15 October 2026 [6]. Source metadata, license notices and accessible full text were checked during preparation. Dr. Freedom Cheteni is the named human author; final verification, approval and accountability remain human responsibilities. AI is a tool, not an author or peer reviewer.

Human expert peer review was completed by Dr. Vickie Aultmon, as confirmed by Dr. Freedom Cheteni on 15 October 2026. This is the date the review confirmation was recorded.

Evidence and limits

Claim	Evidence	Limit
The principal framework is proposed, not empirically validated.	A conditional account of prediction, action and revision that separates the learning mechanism from display immersion. [1]	Requires substantive author scrutiny, independent review and prospective evaluation.
Instructional method and transfer must be distinguished from the immersive medium.	Theoretical model, systematic review and a culturally situated science study. [2, 3, 4, 5]	Sources have different designs and populations; this article reports no new outcome data.

References and rights

[1] Freedom Cheteni. Movement Thinking: The Human World Model. The VR School · Media. 2026-08-03.

https://www.thevrschool.org/media/world-models-and-learning

Citation only; no open article license asserted. No source text, photographs or figures are republished. Treated as philosophical provenance, not independent institutional outcome evidence.

Authoritative source / evidence | checked 2026-10-15

[2] Guido Makransky; Gustav B. Petersen. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. Educational Psychology Review, 33, 937-958. 2021-01-06.

https://doi.org/10.1007/s10648-020-09586-2

CC BY 4.0. © The Author(s) 2021. Cited and analyzed; no figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[3] D. Hamilton; J. McKechnie; E. Edgerton; C. Wilson. Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. Journal of Computers in Education, 8, 1-32. 2020-07-11 (online); 2021 issue.

https://doi.org/10.1007/s40692-020-00169-2

Citation only in this edition. Source text and figures are not republished; the article is cited for its review findings and limitations.

Authoritative source / evidence | checked 2026-10-15

[4] Bryan A. Brown; Kathryn Ribay; Greses Pérez; Phillip A. Boda; Matthew Wilsey. A Virtual Bridge to Cultural Access: Culturally Relevant Virtual Reality and Its Impact on Science Students. International Journal of Technology in Education and Science, 4(2), 86-97. 2020.

https://files.eric.ed.gov/fulltext/EJ1255660.pdf

Citation only; no article-wide CC license asserted for this version. Journal copyright notice retained. The full article is not ingested or republished as CC research.

Authoritative source / evidence | checked 2026-10-15

[5] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only. Rights retained by the original rights holder; no text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[6] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[7] Margaret Wilson. Six views of embodied cognition. Psychonomic Bulletin & Review. 2002-12.

https://doi.org/10.3758/BF03196322

Citation only; no Creative Commons article license asserted. Source text and figures are not republished. Original authors and publisher retain their rights.

Authoritative source / evidence | checked 2026-10-15

This new conceptual manuscript cites the named original authors and philosophical sources. Source studies retain their original authors, venues, dates, identifiers and rights. Proposed constructs and research designs are analytical contributions for substantive human review; no source findings are reassigned to the intended author.

A05 | CONCEPTUAL ARTICLE / POLICY ANALYSIS

Belonging by Design: Cultural Authority, Spatial Learning and the Work of School Leadership

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 1.1 | 2,745 article words

Human expert peer review: Dr. Vickie Aultmon. Review confirmation recorded by the publisher on 2026-10-15.

Keywords: cultural authority, school leadership, participation, spatial learning

Read online / article companion / citation exports

Contribution

A three-dimensional account of representational access, interpretive authority and institutional responsiveness.

Purposive theoretical comparison; no new participants, empirical effects or completed institutional evaluation.

Abstract

A virtual environment can represent a community without allowing that community any authority over the representation. This conceptual article examines the difference between cultural visibility, educational participation and consequential voice in spatial learning. It takes Cheteni's Beyond Palm Drive essay as a philosophical prompt and develops a leadership account informed by culturally relevant VR research, immersive-learning theory and studies of inclusive leadership and school climate. The proposed contribution is a distinction between representational access, interpretive authority and institutional responsiveness. These dimensions identify what families, learners and educators can recognize, challenge and change in a learning design. A worked hypothetical example and a prospective mixed-methods evaluation show how the distinction could guide curriculum decisions without treating culture as a fixed demographic property. The article addresses disagreement within communities, unequal capacity to participate, privacy, appropriation and the possibility that attractive participation systems intensify burdens. No institutional outcome claims or new empirical findings are reported. The argument is that spatial-learning leadership should be assessed partly by whether it creates defensible routes for affected people to alter the conditions of learning, rather than by the cultural imagery a platform displays.

1. The difference between appearing and having a say

Schools increasingly have the technical capacity to place familiar streets, voices and experiences inside a digital environment. That capacity creates an opportunity and a risk. Students may encounter disciplinary ideas through settings that matter to them. Yet a school can also borrow the appearance of a community while leaving its assumptions and decision rules untouched. Representation becomes decorative when people can see themselves in a lesson but cannot question how they have been portrayed or what conclusions the lesson invites them to draw.

The leadership problem concerns authority as well as content. Who chooses the problem? Who decides which features of a place are relevant? Who may challenge the interpretation? What happens when an objection is well founded? A learning environment is not culturally responsive simply because it contains diverse avatars or a recognizable neighborhood. These questions do not establish that every participant should control every curricular decision. They require an institution to explain how disciplinary responsibility and affected people's knowledge are brought into a defensible relation.

2. Philosophical provenance and scholarly boundaries

Beyond Palm Drive presents a vision of learning connected to purpose, place and students' communities [1]. This article uses that vision to formulate a question about cultural authority. It does not treat the essay as an independently evaluated account of programme effectiveness, accreditation, partnerships or student outcomes. The new distinctions developed here are analytical proposals. They are not claimed to be a previously validated Cheteni instrument or an established scale of culturally responsive leadership.

Brown and colleagues' study of culturally relevant VR science provides a more specific empirical reference [2]. Its design combines attitudinal surveys and interviews around a particular lesson; the reported connections between science and lived context do not establish that every locally themed immersive environment improves attainment. The present argument builds a leadership question from that limited warrant. It asks what institutional conditions make a culturally situated design open to correction and educationally useful, rather than assuming that local familiarity itself is an intervention with a uniform effect.

The account also needs explicit intellectual debt to Ladson-Billings’s foundational theorization of culturally relevant pedagogy [7]. That work makes the relation between culture, teaching and the researcher’s position an object of inquiry, rather than treating cultural difference as a feature to decorate instruction. The three dimensions proposed here concern institutional authority over representations and responses. They are offered as a bounded governance extension, not a replacement for culturally relevant pedagogy and not a claim to have originated its central concerns. This distinction matters when a new philosophy develops its contribution in dialogue with an established scholarly tradition.

3. Evidence that informs, without settling, the argument

Makransky and Petersen's theoretical account provides a reason to distinguish instructional methods from technological affordances [3]. Cultural relevance should likewise be examined as a design process rather than a property that a headset supplies. Bao's observational study links inclusive leadership with teacher innovation through proposed mediating relations, but its sampling and design limit causal inference [4]. It does not establish that a particular participation process will produce innovation. Together, these sources motivate attention to the conditions of interpretation and action rather than a simple technology-adoption explanation.

Alinsunurin's multilevel analysis connects parental involvement, leadership and learning climate while also identifying possible tensions in mandated involvement [5]. Its relevance is not an instruction to maximize every form of participation. It is a warning that the organization and meaning of involvement matter. A meeting invitation may create an opportunity for some families and an additional burden for others. A mandatory consultation can produce formal compliance without a credible route to influence. The present framework treats these differences as central rather than incidental implementation detail.

4. Three dimensions of cultural authority

Representational access concerns whether learners can encounter and understand the people, settings, languages and practices through which a disciplinary problem is presented. It includes accessible modes of participation and the absence of unnecessarily exclusionary assumptions. It does not require every lesson to mirror each learner's experience. Encountering an unfamiliar place can be valuable. The question is whether the design provides a route into understanding and whether its treatment of familiarity and difference is justified.

Interpretive authority concerns the standing to explain what a representation means and to challenge an interpretation. A student may recognize a local market but disagree with the lesson's account of why food choices differ across neighborhoods. A family may identify a missing historical or economic condition. These contributions are not automatically correct because they are local, just as an official explanation is not automatically correct because it appears in the curriculum. A defensible process makes competing accounts answerable to evidence while respecting the people who offer them.

Institutional responsiveness concerns what the school does when a challenge reveals a consequential problem. A correction may require changing a scene, revising a task, obtaining permission, explaining a curricular decision or acknowledging an unresolved disagreement. Responsiveness does not mean promising to accept every request. It means providing reasons, a responsible decision-maker and an inspectable outcome. Without this dimension, consultation can become a mechanism for collecting stories without changing the institutional design that gave rise to the concern.

5. Culture is a relation, not a student attribute field

A platform may be tempted to tailor cultural content by assigning learners to demographic categories. That approach risks treating a complex, changing relation as a stable trait. Two students with the same reported background may differ in language, migration history, interests, religion, neighborhood experience and their relationship to the subject. A student may also wish to explore a context outside an assigned category. Personalization should therefore offer meaningful choices and explanations rather than infer what a learner's culture must be.

The proposed unit of design is the relationship among a disciplinary question, a particular setting and the participants who can contribute knowledge about it. That relationship changes over time. An environment built with last year's students may need revision as a neighborhood changes or new participants identify omissions. A school should record the provenance of the representation and the scope of the consultation. It should not use a single contributor's approval as evidence that an entire community has consented.

6. A worked hypothetical curriculum decision

Consider a hypothetical unit on energy, food systems and household choices. Students encounter a digital neighborhood containing shops, transport routes and different food options. The science objective concerns energy transfer; the social question concerns the conditions under which choices are made. A weak design would invite students to rank households as responsible or irresponsible without representing cost, travel time, availability or cultural meaning. Its local imagery would make a narrow explanation look familiar rather than make it adequate.

A stronger design would invite learners to identify which variables the model includes and what it leaves out. They might compare two explanations for an observed pattern, examine relevant evidence and revise the model. Community contributors could question an inaccurate representation of a shop or a misleading assumption about access. Teachers would retain responsibility for disciplinary accuracy and for explaining the limits of the evidence. The learning task would not require students to reveal their own household income, food insecurity or family practices in order to participate.

The example is a proposed design, not a report of a lesson conducted at The VR School. Its success would require evidence that students reason more carefully about the target relations and that the participation process is workable. An attractive scene, a positive comment from one family or a well-produced student artifact would not establish either

conclusion. A useful evaluation could reveal that a less elaborate, non-immersive version supports the same reasoning with fewer costs or fewer access barriers.

7. Participatory governance with boundaries

A school could establish a curriculum representation record for each consequential environment. The record would identify the learning objective, source materials, contributors, permissions, disputed interpretations and decisions made. Contributors should know whether their input is advisory, jointly decisive or subject to a specified curricular authority. Ambiguity about these roles may generate disappointment even when the institution believes it has consulted extensively. Clear boundaries are therefore compatible with meaningful participation; they can make its consequences more credible.

Participation also has a cost. Evening meetings, unpaid review work and lengthy digital forms favor people with time, confidence and connectivity. A credible process should offer asynchronous and accessible routes, consider compensation where appropriate, and report whose perspectives were not reached. Attendance should not be treated as a measure of how much a family values education. The school should examine barriers before interpreting absence as lack of interest. It should also protect staff from an unlimited obligation to respond immediately to every request.

Disagreement within a community is expected. A school may receive incompatible accounts of a local history or conflicting preferences about an image. It should not manufacture consensus by selecting the most convenient representative. Where the evidence permits, the learning environment can make the disagreement itself an object of inquiry. Where safeguarding, privacy or disciplinary requirements constrain a decision, the institution should explain that constraint. The objective is accountable judgment, not the appearance that conflict has disappeared.

8. A prospective mixed-methods evaluation

A future evaluation could compare two well-supported approaches to the same learning objective: a conventional curriculum-design process and a process that includes the three dimensions of cultural authority. The contrast should specify what additional resources the latter receives. If it includes substantially more teacher planning time or subject expertise, those resources are part of the package. An explanatory study might attempt to separate components; a pragmatic study would estimate the effect of offering the whole process. Both can be useful if their claims match their designs.

Quantitative outcomes might include independently scored disciplinary reasoning, a prespecified transfer task, participation opportunities and the time required of staff and families. Qualitative interviews and observations could examine how challenges were received and whether contributors understood their influence. The questions should include negative cases: learners who did not feel represented, families who chose not to participate and teachers who found the process burdensome. A design that interviews only enthusiastic participants would provide an incomplete account of institutional responsiveness.

The analysis should distinguish an intervention's effect on learning from its effect on experience. Improved belonging may be educationally important without immediately changing a test score; improved scores may coexist with exclusion. Neither outcome should be made to stand for all the others. Where group comparisons are planned, investigators need sufficient observations, justified measures and privacy protections. They should not publish identifiable small-group results merely to demonstrate that equity has been considered. No outcome data, effect estimates or completed evaluation are claimed here.

9. Measures, interpretation and the risk of a new compliance system

The three dimensions are not offered as a summed index. Adding a high representation score to a low responsiveness score could conceal the precise failure the framework is intended to expose. A profile of evidence is more useful: what was made accessible, who could challenge it and what changed after challenge. A future instrument would require construct development, evidence of reliability and validity, and examination of its consequences. The present article provides questions for that work rather than a ready-made certification scale.

A participation record can itself become performance theatre. Schools may count the number of contributors, requested changes or completed consultations while avoiding substantive disagreement. An evaluation should therefore inspect a small number of consequential decisions in depth. Did the institution acknowledge an error? Did a contribution alter a representation or clarify a reason for retaining it? Could participants explain what happened? These questions focus on the quality of institutional response rather than the volume of evidence uploaded to a dashboard.

10. Rights, privacy and the limits of openness

Making learning public can benefit a community while exposing people to unwanted attention. A recognizable street, voice, photograph or account of an event may reveal more than a contributor intended. Permission to participate in a lesson is not automatically permission to distribute a digital representation worldwide or use it to train a model. The institution should specify the intended uses, retention, access and withdrawal arrangements appropriate to the project. This article proposes that distinction as a governance requirement; it does not assert a universal legal rule for every jurisdiction.

Open licensing is likewise a separate decision. A student may cite a source without owning rights to redistribute its photographs. A community contributor may agree to a classroom use while declining an unrestricted license. The school should not make broad licensing a condition for receiving ordinary educational credit. Alternatives can include restricted circulation, redacted artifacts or a public explanation that omits sensitive material. The public value of scholarship does not require every part of the learning process to become public data.

11. Implications for superintendents and institutional learning

The proposed leadership practice is to establish a credible route from recognition to challenge to response. A district can begin with a small number of curriculum decisions where the target concept and responsible authority are clear. It should resource teacher preparation, identify who can correct a representation and explain how unresolved disagreement will be handled. These commitments matter more than announcing that a campus is culturally responsive. They give families and staff something concrete to examine when the announcement is tested by experience.

The framework also creates a reciprocal obligation. Institutions ask students to revise their interpretations in response to evidence; institutions should show how they revise their own representations when challenged. This is the link to Movement Thinking developed here. Responsiveness is an observable institutional practice rather than a statement about the intentions of an individual leader. Its educational significance remains an empirical question, but its governance significance can already be specified in the design of responsibilities and records.

12. Limitations and conclusion

The synthesis is purposive and crosses philosophical, empirical and organizational literatures with different standards of evidence. It does not establish a causal pathway from cultural authority to achievement, validate the proposed dimensions or evaluate the effectiveness of The VR School. Cultural and institutional settings may alter both the meaning and feasibility of participation. The argument may overestimate how readily curricular authority can be made explicit in organizations with competing obligations. A prospective study should examine these limits rather than treat implementation difficulty as a failure of participants to understand the vision.

A spatial-learning environment is educationally distinctive when it helps people examine a relation they could not otherwise inspect as well. It becomes institutionally defensible when those represented within it can challenge consequential errors and receive a reasoned response. The contribution of this article is to bring those two demands together without collapsing them. A community should be more than the scenery of an innovation. Its members should have an intelligible place in the work through which the innovation becomes worthy of trust.

Method and data statement

This conceptual analysis purposively compares five named primary sources and develops a hypothetical design example and prospective evaluation. It does not report a systematic search, meta-analysis, new participants or institutional outcomes. Brown and colleagues' article is cited from the authoritative ERIC-hosted full text; no article-wide Creative Commons license is asserted for that version. The other research sources retain their own authors, designs, dates and rights. Citation verification is documented separately from scholarly quality assessment.

AI, authorship and editorial disclosure

ChatGPT assisted with literature organization and drafting of this manuscript on 15 October 2026 [6]. Source metadata, license notices and accessible full text were checked during preparation. Dr. Freedom Cheteni is the named human author; final verification, approval and accountability remain human responsibilities. AI is a tool, not an author or peer reviewer. The proposed dimensions and examples are analytical constructions, not validated institutional results or statements endorsed by all school participants.

Human expert peer review was completed by Dr. Vickie Aultmon, as confirmed by Dr. Freedom Cheteni on 15 October 2026. This is the date the review confirmation was recorded.

Evidence and limits

Claim	Evidence	Limit
The principal framework is proposed, not empirically validated.	A three-dimensional account of representational access, interpretive authority and institutional responsiveness. [1]	Requires substantive author scrutiny, independent review and prospective evaluation.
Cultural visibility alone does not establish consequential participation.	Mixed-methods science study, theoretical analysis and observational school research. [2, 3, 4, 5]	Sources have different designs and populations; this article reports no new outcome data.

References and rights

[1] Freedom Cheteni. Beyond Palm Drive: Bringing the Lesson Home. The VR School · Media. 2026-07-21.

https://www.thevrschool.org/media/beyond-palm-drive

Citation only; no open article license asserted. No source text, photographs or figures are republished. Treated as philosophical provenance, not independent institutional outcome evidence.

Authoritative source / evidence | checked 2026-10-15

[2] Bryan A. Brown; Kathryn Ribay; Greses Pérez; Phillip A. Boda; Matthew Wilsey. A Virtual Bridge to Cultural Access: Culturally Relevant Virtual Reality and Its Impact on Science Students. International Journal of Technology in Education and Science, 4(2), 86-97. 2020.

https://files.eric.ed.gov/fulltext/EJ1255660.pdf

Citation only; no article-wide CC license asserted for this version. Journal copyright notice retained. The full article is not ingested or republished as CC research.

Authoritative source / evidence | checked 2026-10-15

[3] Guido Makransky; Gustav B. Petersen. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. Educational Psychology Review, 33, 937-958. 2021-01-06.

https://doi.org/10.1007/s10648-020-09586-2

CC BY 4.0. © The Author(s) 2021. Cited and analyzed; no figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[4] Yuangen Bao. A dual mediation model of the association between principal inclusive leadership and teacher innovative behavior. Scientific Reports. 2024-05-29.

https://doi.org/10.1038/s41598-024-63332-0

CC BY 4.0. © The Author(s) 2024

Authoritative source / evidence | checked 2026-10-15

[5] Jason Alinsunurin. School learning climate in the lens of parental involvement and school leadership: lessons for inclusiveness among public schools. Smart Learning Environments. 2020-09-21.

https://doi.org/10.1186/s40561-020-00139-2

CC BY 4.0. © The Author(s) 2020

Authoritative source / evidence | checked 2026-10-15

[6] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[7] Gloria Ladson-Billings. Toward a Theory of Culturally Relevant Pedagogy. American Educational Research Journal. 1995-09.

https://doi.org/10.3102/00028312032003465

Citation only; no Creative Commons article license asserted. Source text and figures are not republished. Original authors and publisher retain their rights.

Authoritative source / evidence | checked 2026-10-15

This new conceptual manuscript cites the named original authors and philosophical sources. Source studies retain their original authors, venues, dates, identifiers and rights. Proposed constructs and research designs are analytical contributions for substantive human review; no source findings are reassigned to the intended author.

A06 | CONCEPTUAL ARTICLE / POLICY ANALYSIS

Publication as Learning Infrastructure: Relational Accountability, Student Agency and the Scholarly Record

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 1.1 | 2,939 article words

Human expert peer review: Dr. Vickie Aultmon. Review confirmation recorded by the publisher on 2026-10-15.

Keywords: publication as learning, student agency, ORCID, editorial independence

Read online / article companion / citation exports

Contribution

A course-to-publication architecture that distinguishes assessment, submission, acceptance, publication and software contribution.

Purposive theoretical comparison; no new participants, empirical effects or completed institutional evaluation.

Abstract

Connecting coursework to publication can give learners a consequential audience, but it can also confuse the completion of an assignment with the acceptance of a scholarly contribution. This conceptual article develops a publication-as-learning model that preserves the distinction. Cheteni's essays on institution-building and relational intelligence provide philosophical provenance; school implementation and mental-health leadership research inform attention to organizational conditions and support. The proposed model separates learning artifacts, assessed work, submissions, accepted publications and software companions. It assigns each object an appropriate form of authorship, consent, review, version history and identifier. A prospective evaluation compares a supported publication pathway with an equally demanding internal research pathway and examines reasoning, revision, participation burden and wellbeing rather than publication counts alone. The article also considers minors' privacy, the right not to publish, editorial conflicts, the scarcity of qualified review and the limits of proof-of-work as an educational claim. No new student outcomes or evidence of institutional accreditation are reported. The contribution is an accountable architecture through which learning can become scholarship without turning every learner's unfinished work into a permanent public credential.

1. From audience to scholarly responsibility

An authentic audience can change the meaning of an assignment. A student may need to explain why a question matters to someone beyond the teacher, defend the evidence behind a conclusion and respond to an objection that cannot be resolved by following a marking guide. Publication offers one possible setting for that work. Yet public visibility is not the same as scholarly value. A document becomes a contribution through the quality of its question, reasoning, evidence and accountable review, not simply because an institution gives it a permanent URL.

The distinction becomes more important when an LMS, journal platform and AI assistant are connected. The technical transition from a course artifact to a public page can become almost frictionless. The scholarly transition should remain deliberate. It requires a decision about authorship, readiness, permissions, review and the status of the work. Removing unnecessary administrative effort is valuable; removing the distinctions that make a scholarly record interpretable is not. The proposed architecture makes those distinctions navigable rather than allowing software convenience to decide them implicitly.

2. Philosophical provenance without institutional self-certification

Cheteni's institution-building essay presents the ambition to widen educational participation through new forms of schooling [1]. His commencement address connects action, purpose and relational responsibility [2]. These texts are treated as statements of educational purpose and authorial perspective. They are not used here as independent evidence of institutional effectiveness, accreditation, student achievement or a universal psychological law. The article develops a new publication model from that philosophical starting point while retaining the need to test its educational claims.

Relational accountability is defined here as an obligation to explain and revise one's work in a setting where other people can question it and where consequential decisions have identifiable responsibility. The definition is a proposal for this article. It is not a validated measure of relational intelligence. It also does not require students to agree with an institutional philosophy. A learner should be able to criticize the model, decline public dissemination or reach an unwelcome conclusion without losing ordinary access to instruction or fair assessment.

Black and Wiliam’s review of assessment and classroom learning offers a necessary distinction between using evidence to improve learning and merely recording a judgment [7]. A course-linked publication process becomes educationally relevant when feedback changes a learner’s reasoning or next action; the existence of a grade or a public record does not establish that formative function. The present model therefore treats revision and the learner’s explanation of it as outcomes to investigate. The review does not establish the effectiveness of journal publication as an intervention. It provides a stronger conceptual basis for asking what learners do with feedback before and after submission.

3. Why implementation conditions belong inside the model

Lyon and colleagues' development and validation of the School Implementation Leadership Scale demonstrates the importance of specifying leadership practices and examining the interpretation of measures in an educational setting [3]. It does not show that creating a publication platform improves learning. Its relevance is methodological: implementation requires identifiable support, communication and responsible leadership rather than a general declaration that an institution values evidence. The proposed publication pathway should be evaluated with the same attention to what people are actually enabled to do.

Daly and colleagues' narrative review addresses leadership in relation to school mental-health provision [4]. It is not an experimental evaluation of publication-based learning. It does, however, make a useful boundary visible: adding an academically

ambitious pathway does not remove leaders' responsibility for support and the conditions under which students participate. The philosophical orientation described in Cheteni's earlier Movement Thinking essay likewise does not establish those conditions automatically [5]. A rigorous proposal must identify the additional work and resources the pathway requires.

4. Five objects that a learning platform must not collapse

A learning artifact is an attempt made in the course of inquiry. It may be incomplete, mistaken or exploratory. Its immediate audience may be the learner, teacher or a small group. Assessed work is an artifact judged against a course's stated learning criteria. That judgment can be sound without implying that the work is ready for a journal. A submission is a separately consented request for editorial consideration. It has entered a publication process but has not thereby been accepted.

An accepted publication has passed the journal's actual editorial requirements and has a recorded decision, approved text and appropriate rights. Acceptance and public release are distinct events, and their dates should describe what happened. A software companion is a separate object that may help readers navigate or interrogate a publication. Its implementation, evidence boundaries and performance should be described independently. The companion does not supply the article's human authorship or complete the human peer review that the journal requires.

These distinctions prevent several misleading shortcuts. A grade cannot automatically become an acceptance decision. A source paper imported under a license cannot become a student's original research output. A generated explanation cannot become a reviewer report by changing its label. An ORCID connection cannot prove that the connected person made every contribution listed on a record. The platform should make each transition explicit and preserve the evidence that authorizes it, while keeping routine navigation simple enough for students and families to understand.

5. An end-to-end pathway for a course-linked publication

The pathway begins with a course question and an explanation of what students are expected to learn. A research plan identifies the proposed contribution, feasible evidence, relevant permissions and foreseeable limitations. A student working with human participants may require ethics review and additional consent before collecting information; software should not treat a completed form as proof that those requirements have been satisfied. The pathway must offer appropriate supervision and a viable alternative when the original project is not feasible.

During inquiry, learners maintain a proportionate record of decisions, sources and revisions. The purpose is to make reasoning inspectable, not to demand constant self-surveillance. A teacher assesses the work using the course criteria and provides feedback. Only then, if the learner chooses, is a submission prepared with its own author statement, disclosures, permissions and intended venue. The journal's editorial process determines scholarly readiness. The learner should receive a clear explanation that rejection or substantial revision is compatible with successful course learning.

After acceptance, the author approves the final text and the record carries actual dates, contributor roles, source identifiers and licensing terms. A DOI is displayed only after registration through an authorized account and an appropriate content type. An ORCID update requires the relevant authorization and must describe the person's actual work. Corrections and later versions remain linked. If the paper has a companion, the companion's scope, version and evaluation status are visible without implying that the original scholarly contribution has changed hands.

6. A right not to publish

Publication is one route to a consequential audience, not the only route to educational seriousness. Students may work on personal, community or commercially sensitive questions. They may wish to revise an early interpretation without leaving it permanently discoverable. They may also lack confidence in the risks of public exposure even when adults believe publication would help them. A defensible programme therefore needs a non-public pathway with comparable academic expectations and support. Choosing it should not become a hidden penalty in grading or access to mentoring.

For minors, the institution should distinguish participation in learning, consent to a research project, use of a name or image, and public licensing of an output. These decisions can involve different people and obligations. The platform should present choices in language that students and families can understand and provide an appropriate way to revisit them before release. It should avoid implying that an unrestricted public license can always be recalled after material has circulated. The article offers a governance principle; local legal and safeguarding arrangements still require competent institutional review.

The right not to publish also protects the integrity of the scholarly record. If every enrolled student is expected to produce an accepted paper, editors face pressure to convert completion into acceptance. A course can promise a supported opportunity to conduct research and submit work. It should not promise a favorable editorial decision. The

distinction allows a journal to maintain standards without turning rejection into an educational failure or a financial grievance about an allegedly guaranteed outcome.

7. Authorship, AI assistance and the assessment of contribution

An AI-assisted manuscript may be fluent while concealing gaps in its author's understanding. Assessment should therefore examine the reasoning process as well as the final prose. A learner might explain the choice of method, trace a consequential claim to a source, defend an interpretation against an alternative and describe a revision made after feedback. These activities provide evidence of contribution and understanding. They do not require banning useful tools, nor do they establish that a particular student used a tool improperly simply because the final writing is polished.

The journal should require an accurate account of assistance: what tools were used, for what purpose, and what the human authors checked and approved. AI should not appear as a human author or be credited with independent peer review. Reviewers also need clear confidentiality rules before using external tools on an unpublished submission. Platform ownership, teaching, editing and authorship remain different contributions. Dr. Cheteni's role in the institution cannot replace the actual authors of student research or imported sources, and students should not be assigned authorship for work they did not contribute to substantively.

ORCID helps connect a person's research identity with eligible outputs, but it is not a quality score. A large work count may reflect many different kinds of contribution and says little by itself about rigor. The record should distinguish articles, datasets, software and other outputs appropriately and preserve identifier relationships. Adding thousands of source papers to a platform owner's record merely because agents were attached would obscure rather than improve attribution. A useful infrastructure makes the contribution easier to understand, even when that results in fewer records.

8. A prospective evaluation of publication as learning

A future study could compare a supported publication pathway with an equally supported internal research pathway. Both would involve a substantive question, evidence appraisal, revision and an authentic audience; the second might present to a school or community panel without journal submission. This active comparison avoids attributing the effects of additional mentoring or project time to publication itself. Where allocation is not feasible, investigators should make the limits of selection adjustment explicit. More motivated students may choose publication and differ before the pathway begins.

The primary outcome could be independently assessed research reasoning on a new task, with delayed assessment where feasible. Secondary outcomes might include the quality of revisions, source evaluation, willingness to acknowledge uncertainty, participation burden and students' accounts of agency. Acceptance rates and readership are descriptive process measures rather than direct measures of learning. A rejected manuscript may accompany substantial learning; an accepted manuscript may receive extensive adult assistance. The evaluation should investigate those possibilities instead of assuming the status of a journal decision reveals a student's competence.

The study should prespecify the unit of analysis, account for clustering by course or mentor, and document attrition and access to support. Raters should be masked to pathway where possible and should use a rubric whose interpretation has been examined. Qualitative work should include students who decline publication, withdraw or receive unfavorable decisions. No specific sample, effect or completed trial is reported here. The design is intended to make a future educational claim contestable rather than provide a research appearance for an already chosen conclusion.

9. Editorial independence and the scarcity of review

A publication pathway can generate submissions faster than a journal can recruit qualified human reviewers. That imbalance creates an institutional capacity question. Automated checks can assist with formatting, missing fields or source navigation, but they cannot be relabeled as independent peer review. The journal should align submission volume, editorial expertise and review expectations. It may be more responsible to accept fewer submissions or operate clearly described unreviewed working-paper channels than to promise a review process it cannot deliver.

Conflicts become particularly important when teachers, institutional leaders or the Editor in Chief submit work. An independent handling editor should manage decisions in which the usual editor has a personal interest. Reviewers need sufficient expertise and appropriate independence, and the decision record should identify the actual process without exposing confidential material unnecessarily. A policy page is useful infrastructure, but its existence is not evidence that a particular paper has been independently reviewed. Status claims must follow recorded events.

The journal also needs a route for corrections, challenges and appeals that is understandable to early-career authors. A student should be able to report an attribution error without understanding a database schema. An author should know whether a correction changes a landing page, a downloadable file, a citation export or an external registry record. These relations are technical details for maintainers but practical

promises for users. Their design determines whether a publication remains accountable after the ceremony of launch.

10. Proof of work, institutional evidence and accreditation

Published work can make learning visible, but it is a partial and selected record. It may omit projects that failed, students who lacked support and forms of learning poorly represented by an article. A portfolio of attractive publications cannot by itself establish programme quality, equitable opportunity, institutional sustainability or accreditation. Those conclusions require additional evidence and appropriate external processes. A platform should help organize evidence without presenting its own dashboard as an accrediting judgment.

The stronger interpretation of proof of work is a traceable account of a claim and the activity that warrants it. In education this includes how the task was designed, what support was available, what the learner contributed and how the work was assessed. Some of that evidence should remain private. A public article can link to an appropriately scoped method or artifact without exposing educational records. The objective is credible inspectability, not maximal disclosure. A prestigious institution should be able to explain the limits of its proof as clearly as its achievements.

11. Interoperability and institutional responsibility

An LMS integration should exchange the minimum information needed for a defined transition. Course enrollment does not imply consent to a journal account, and a gradebook identifier is not a public scholarly identity. Submission packages can carry a manuscript, structured author information, rights statements and source files while leaving private course records behind. The institution should define who can initiate the transfer, what the receiving system validates and how the learner can inspect the result before public release.

The architecture should be portable in a practical sense. Authors need stable landing pages and usable exports; editors need records that preserve decisions and versions; institutions need a way to leave a service without losing their scholarly history. Portability does not mean that every LMS shares the same authorization model. Integrations require their own testing and documentation. The contribution proposed here is a common set of scholarly transitions and responsibilities, not a claim that one untested connector already works with every platform.

12. Limitations and conclusion

This model has not been evaluated as an educational intervention. It draws on a limited, purposively selected set of philosophical and organizational sources and proposes new distinctions whose usefulness requires investigation. Publication may motivate some learners and burden others. Additional review and consent steps may protect participants while creating delays or unequal access if institutions do not resource them. The model may be inappropriate for courses whose learning goals do not require a publication pathway. Those possibilities should be examined rather than dismissed as resistance to innovation.

Publication becomes learning infrastructure when it makes reasoning, revision and responsibility more consequential without confusing them with publicity. Its strongest promise is not that every course produces an accepted paper. It is that learners can enter a scholarly conversation with an intelligible account of what they know, how they know it, what they contributed and what remains open to challenge. An institution built around that promise must preserve the right to remain unpublished, the independence of editorial judgment and the integrity of the source record. Those constraints are part of the educational design that makes the promise worth pursuing.

Method and data statement

The article is a conceptual architecture and policy analysis. It does not report new student data, a systematic review, a validated assessment instrument, completed review decisions or an evaluation of institutional effectiveness. The course-to-publication pathway and comparative study are proposals. Media essays are philosophical sources rather than independent institutional audits. Research sources retain their original authors and designs, and their relevance is distinguished from direct evidence for the new model.

AI, authorship and editorial disclosure

ChatGPT assisted with literature organization and drafting of this manuscript on 15 October 2026 [6]. Source metadata, license notices and accessible full text were checked during preparation. Dr. Freedom Cheteni is the named human author; final verification, approval and accountability remain human responsibilities. AI is a tool, not an author or peer reviewer.

Human expert peer review was completed by Dr. Vickie Aultmon, as confirmed by Dr. Freedom Cheteni on 15 October 2026. This is the date the review confirmation was recorded.

Evidence and limits

Claim	Evidence	Limit
The principal framework is proposed, not empirically validated.	A course-to-publication architecture that distinguishes assessment, submission, acceptance, publication and software contribution. [1]	Requires substantive author scrutiny, independent review and prospective evaluation.
Publication status is not itself a measure of learning.	Conceptual architecture informed by implementation and mental-health leadership research. [2, 3, 4, 5]	Sources have different designs and populations; this article reports no new outcome data.

References and rights

[1] Freedom Cheteni. The School That Shouldn't Exist. The VR School · Media. 2026-05-08.

https://www.thevrschool.org/media/the-school-that-shouldnt-exist

Citation only; no open article license asserted. No source text, photographs or figures are republished. Treated as philosophical provenance, not independent institutional outcome evidence.

Authoritative source / evidence | checked 2026-10-15

[2] Freedom Cheteni. You Were Never the Seed. You Were Always the Storm.. The VR School · Media. 2026 (address; exact release date not supplied).

https://www.thevrschool.org/media/commencement-address-2026

Citation only; no open article license asserted. No source text, photographs or figures are republished. Treated as philosophical provenance, not independent institutional outcome evidence.

Authoritative source / evidence | checked 2026-10-15

[3] Aaron R Lyon; Catherine M Corbin; Eric C Brown; Mark G Ehrhart; Jill Locke; Chayna Davis; Elissa Picozzi; Gregory A Aarons; Clayton R Cook. Leading the charge in the education sector: development and validation of the School Implementation Leadership Scale (SILS). Implementation Science : IS. 2022-07-19.

https://doi.org/10.1186/s13012-022-01222-7

CC BY 4.0. © The Author(s) 2022

Authoritative source / evidence | checked 2026-10-15

[4] Brian P Daly; Annie Resnikoff; Shannon Litke. Effective School Leadership for Supporting Students' Mental Health: Findings from a Narrative Literature Review. Behavioral Sciences. 2025-01-01.

https://doi.org/10.3390/bs15010036

CC BY 4.0. © 2025 by the authors.

Authoritative source / evidence | checked 2026-10-15

[5] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only. Rights retained by the original rights holder; no text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[6] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[7] Paul Black; Dylan Wiliam. Assessment and Classroom Learning. Assessment in Education: Principles, Policy & Practice. 1998-03.

https://doi.org/10.1080/0969595980050102

Citation only; no Creative Commons article license asserted. Source text and figures are not republished. Original authors and publisher retain their rights.

Authoritative source / evidence | checked 2026-10-15

This new conceptual manuscript cites the named original authors and philosophical sources. Source studies retain their original authors, venues, dates, identifiers and rights. Proposed constructs and research designs are analytical contributions for substantive human review; no source findings are reassigned to the intended author.

A07 | PRIMARY-STUDY PROTOCOL AND REPRODUCIBLE METHODS RESEARCH

Who Receives the Automation Dividend? A Matched-Pair School Trial of Protected Time, Teacher Workload and Verified Quality

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 1.0 | 4,421 article words

Primary-study protocol and reproducible methods; separate human editorial record.

Keywords: Movement Thinking, teacher time use, protected planning, school governance, American Time Use Survey 2025, matched-pair cluster trial, AI verification burden

Read online / article companion / citation exports

Contribution

An original protected-time allocation experiment after common AI support, an executed ATUS 2025 source audit, and reproducible design sensitivity with an explicit pre-enrollment activation gate.

Prospective primary study specification plus completed secondary-file audit and artificial-data planning simulations; no local participant outcomes.

Abstract

An AI-assisted draft produced quickly does not establish that a school has reduced work. Verification, adaptation and coordination can consume the apparent saving, while new output expectations can absorb any remainder. This article specifies an original prospective primary study of the institutional allocation of that remainder: a matched-pair school trial in which both conditions receive common AI workflow support, but only one receives an enforceable protected-time charter. The primary outcome is change in teachers' weekday job-related minutes outside an agreed paid timetable; independently blinded quality assessments provide a separate guardrail. The planning design contains 48 schools in 24 district pairs and approximately 576 consenting teachers followed for 12 weeks. No schools or teachers have been recruited for this protocol. Two methodological investigations are complete. A reproducible audit of official American Time Use Survey 2025 files checks 6,146 respondent records and 114,151 activity episodes, finds complete 1,440-minute diary closure and identifies 120 respondents under a narrowly declared teacher occupation definition. A nine-scenario artificial-data analysis shows that 48-school power for a hypothetical 20-minute reduction ranges from 35.6% to 87.7% under different assumed variability. Neither investigation estimates an AI or timetable intervention effect. The contribution is a falsifiable evaluation architecture that separates task acceleration, net human work, the distribution of protected time and acceptable educational quality, with explicit limits on surveillance, national benchmarking and causal interpretation.

1. From an efficient task to a defensible institutional gain

Schools are being offered systems that compose communications, generate lesson materials and summarize administrative information. The visible transaction is often the first output: a plausible document appears sooner. The educational transaction is longer. An educator must decide whether the document is accurate, suitable for the intended learners, accessible, consistent with local obligations and worth using. Someone must correct errors, communicate changes and assume responsibility for consequences. A measure confined to the first transaction cannot establish a reduction in the second.

The preceding conceptual article in this volume proposed that an automation dividend belongs in the timetable. This companion develops a different scholarly contribution: an executable public-data audit and a prospective allocation experiment. Movement Thinking supplies the orientation that educational methods should evolve through inclusive, revisable practice [1]. Here that orientation becomes an institutional proposition: a credible efficiency gain should expand the practical space for educational judgment rather

than automatically increase output quotas. Whether a specific charter accomplishes that remains an empirical question.

The object of study is therefore a governance decision following common technical support. The trial does not compare an enthusiastic AI school with a school denied access to all assistance. It asks whether a precommitted allocation rule changes working conditions when the technical opportunity, training and routine support are comparable. This distinction makes the proposed causal claim narrower and more useful to a superintendent deciding how efficiency gains should be distributed.

2. What existing productivity evidence permits

Noy and Zhang studied generative assistance in bounded professional writing tasks [9]. Their experiment supplies a reason to investigate changes in task completion and quality together. The tasks, participants and observation window do not establish a sustainable reduction in a teacher's working week. The publication is cited under its original authors; the accessible MIT manuscript is identified separately in the source audit rather than described as an independently replicated school trial.

Brynjolfsson, Li and Raymond examined generative assistance in customer support [10]. That organizational setting is relevant because task support operates inside an existing workflow. Nevertheless, customer-support performance is not a direct measure of curriculum judgment, teacher workload or protected planning. Sector differences, managerial expectations and the distribution of correction work are plausible effect modifiers. The published article's identity and publisher information were verified; the accessible NBER working-paper record is not silently treated as the identical final version.

Maas and colleagues' study of perceived time pressure, exhaustion and principal support anchors the education-specific concern [11]. It does not establish an AI effect. Together these sources motivate a question rather than answer it: can a school convert technical support into a measurable improvement in work allocation without degrading the quality for which human educators remain responsible? Evidence from other sectors makes that question worth testing, not already settled.

3. The dividend as a three-part accounting problem

The proposed theory separates gross task acceleration, net work change and institutional allocation. Gross acceleration is a difference in the time needed for an initial output. Net work change includes checking, repair, adaptation, training, coordination and displaced responsibilities. Allocation describes who receives the resulting time and what obligations are removed or added. These components can move in opposite directions. A faster draft and a longer working day are logically compatible.

The protocol does not combine these components into a single proprietary productivity score. Each has a distinct indicator and possible counterexample. Self-reported tool use is an exposure description, not evidence of benefit. A published timetable is evidence of an allocation commitment, not proof that a block occurred. An uninterrupted block is an implementation result, not proof of higher educational quality. A time reduction accompanied by serious errors fails the proposed institutional test.

Time also has structure. Sixty fragmented minutes between alerts may differ from one protected hour. The intervention consequently specifies a continuous block, a removal or consolidation of existing obligations and a route for challenging compensating demands. It leaves educators discretion over the learning work pursued within that block. The study tests this bounded arrangement rather than the general desirability of autonomy or technology.

4. Why ATUS is the external database

The American Time Use Survey provides a documented national architecture for reconstructing a day, organizing activity episodes and distinguishing people from episodes [2] [8]. It is useful here as a secondary measurement and feasibility resource. It is not a database of school AI interventions, and no national respondent is treated as a local participant. Official single-year 2025 respondent and activity ZIPs were downloaded, hashed and analyzed; the release page records June 25, 2026.

The data dictionary identifies TUCASEID, TUACTIVITY_N, TUACTDUR24 and TRCODE for episode reconstruction [3]. The selected occupation subset uses the official 2018 Census classification [5], while the 2025 lexicon identifies the main-job work code [6]. These documented fields make a limited audit reproducible. They cannot reveal whether a teacher used a particular AI service or whether a workplace honored a protected-time promise.

The September 2026 User’s Guide documents the survey design, weighting and single-diary structure [4]. ATUS also illustrates a design choice the local study must

improve upon. One recalled diary can describe a sampled day but cannot establish an individual's stable weekly workload. The prospective study therefore samples repeated, assigned teaching weekdays, preserves missing scheduled days in its target and adds an explicit paid-timetable map. This is a new measurement adaptation requiring pilot work; it is not an assertion that the BLS has validated the local instrument.

5. Executed source audit and occupation boundaries

The executed audit reads 6,146 respondent records with unique TUCASEID values and 114,151 activity episodes with unique respondent-activity keys. Every respondent has activity records, no episode is orphaned, and summing truncated durations yields exactly 1,440 minutes for every diary. These are actual file-integrity findings, not hypothetical counts. ZIP and extracted-data hashes, retrieval URLs and the checking script accompany the article.

A declared teacher subset requires employed status TELFS equal to 1 or 2 and TEIO1OCD equal to 2310, 2320 or 2330. The resulting counts are 85 elementary/middle-school teachers, 24 secondary-school teachers and 11 special-education teachers: 120 respondents. Code 2300 combines preschool and kindergarten and is excluded from this particular definition rather than presented as a clean kindergarten group. Postsecondary teachers, assistants and miscellaneous instruction occupations are also excluded. The resulting subset should not be labeled all education workers or a comprehensive census of K-12 teachers.

Of these 120 diaries, 53 concern weekdays and 67 weekends. Only 62 report positive main-job work; among weekday diaries, 45 do. These modest denominators constrain any benchmarking ambition. The counts justify designing a richer local measurement exercise. They do not justify fine-grained national comparisons by grade, district, teacher identity or AI exposure. Public identifiers provide no authorized link to the future school trial.

6. Executed descriptive estimates and the missing autumn

Using TUFINLWGT in a ratio mean, retaining zero main-job work on nonworking diary days, the selected group averages 294.9 minutes of main-job work per sampled diary day. The weekday-only ratio is 416.0 minutes. Restricting instead to positive-work diaries yields 437.6 and 480.0 minutes, respectively. These calculations answer different denominator questions. A nonworking day is not evidence of low professional commitment, and an all-day average must not be multiplied into a supposed contractual working week.

No replicate-weight standard errors, significance tests or national teacher subgroup rankings were calculated. The estimates are descriptive researcher computations, not an official BLS teacher table. Their small domain counts, occupation exclusions and diary design travel with the numbers. The tools deliberately report no weekend mean or comparative inference; weekend records remain in the sample accounting. ATUS main-job work cannot separate teaching, administration, checking, AI generation or work outside an agreed paid timetable.

A further boundary is substantive rather than merely computational. The 2025 federal shutdown suspended operations, leaving no diaries for September 30 through November 11; the audit finds zero records in that interval [7]. Reweighting does not reveal how teachers actually used time on those unobserved dates. An autumn school-work analysis would be especially vulnerable to assuming that remaining fourth-quarter diaries reproduce the missing season. The present findings are therefore a versioned feasibility demonstration, not a recovered autumn workload history.

7. Primary research questions and estimands

The primary question is whether a protected-time charter reduces job-related work outside an agreed paid timetable relative to ordinary allocation discretion after common AI support. The hypothesis is a reduction, tested two-sided to retain the possibility of harm. The candidate practical contrast of 20 minutes per teaching weekday is used only in design sensitivity. It is not an ATUS-derived estimate, predicted saving or promise to participating teachers.

Let the baseline consenting teacher cohort in school s define its mean weekday outcome under assignment z. Let Delta_s(z) be that school's endline minus baseline mean. The principal estimand is the equal-school average of Delta_s(1) minus Delta_s(0) across recruited schools. Negative values favor the charter. School equality makes a small participating school count as a school rather than allowing larger cohorts to dominate an organizational question. An individual-weighted contrast is secondary and explicitly answers a different question.

The quality question is separate: can the time contrast coexist with a defensible bound on deterioration in independently rated outputs? A non-significant quality difference is not evidence of equivalence. The protocol proposes a noninferiority guardrail, subject to external margin justification before allocation. The sustainable-dividend conclusion is conjunctive: favorable time evidence, acceptable quality evidence and no unresolved material safety problem. Every component is reported even when the conjunction fails.

8. Prospective recruitment and allocation

The candidate design contains 24 district pairs, with two eligible schools per district and approximately twelve consenting teachers per school: 48 schools and 576 planned teachers. These are budgeting and sensitivity assumptions. There are zero recruited participants in this package. School agreements, protected paid research time, an independent ethics determination, an instrument pilot, a finalized analysis implementation and public registration must precede activation. Changed feasibility requires a dated amendment before allocation.

Eligible schools must be able to run both conditions within existing paid schedules, support the common approved workflow and protect individual diary confidentiality from supervisors. Teacher invitation, consent and baseline measurement occur before condition disclosure. Recruitment must not select only staff already enthusiastic about AI. A private screening log records eligibility, invitation and refusal. Part-time educators and those with substantial caring obligations remain eligible; their timetable maps and access needs are documented.

Within each district pair, an independent custodian randomly assigns one school to the charter and one to ordinary discretion. The sequence is concealed until rosters and baseline measures are locked. Its generation method is public, while the seed remains access controlled until allocation cannot be manipulated. Cluster reporting follows the distinction between schools and people emphasized by CONSORT guidance [12]. Postallocation recruitment is reported separately, and a school that withdraws is not quietly replaced.

9. Common support and the active comparator

Both conditions receive access to the same approved AI workflow support, an initial two-hour orientation and a 30-minute paid weekly support clinic. Training covers source checking, accessibility, privacy boundaries and a workable non-AI alternative. The common task families are low-risk preparation and administrative drafting. Individual grading, disability eligibility, discipline and safeguarding decisions remain outside the trial's automation scope. Neither tool availability nor study participation authorizes uploading identifiable student material.

The comparator retains ordinary school discretion over how any task gain is used. It receives equal scheduled support, outcome measurement and routes to report problems. Existing protected planning cannot be removed to make the comparison artificially weak. New naturally occurring allocation agreements are documented as overlap, not prohibited

or punished. The estimand is consequently the effect of offering a formal charter against real ordinary discretion, including imperfect separation between conditions.

The common platform and clinic reduce a simple attention explanation, but cannot eliminate differences in expectations or leadership commitment. Participants and leaders cannot be blinded to timetable changes. Fidelity interviews and records of changed demands help interpret the contrast without converting an observational mechanism measure into a randomized mediator. The study does not estimate the independent benefit of AI access because both arms receive it.

10. The protected-time charter

Treatment schools identify one continuous 60-minute paid block per week for teacher-selected learning design, feedback or collaborative inquiry. They list the routine obligation that is ceased, consolidated or reassigned within a funded plan. The block is a reallocation of existing work, not an additional meeting or an expectation of unpaid evening completion. Schools record intended and delivered minutes, interruptions, cancellations and any compensating output demands.

The charter names a human quality owner, a teacher representative and an independent appeal contact. It prohibits increasing required output solely because a draft can be generated faster. If verification consumes the anticipated saving, the record shows no realized net gain; leaders cannot manufacture a dividend by excluding correction labor. A teacher may use the protected block without using AI on every task. The assigned policy, rather than compulsory tool compliance, defines treatment.

The 60-minute commitment and hypothetical 20-minute daily contrast are not an accounting identity. Work can be reorganized across days, and checking or coordination can either expand or contract. The analysis therefore reports delivered blocks, total work, outside-timetable work and quality separately. A positive result must identify who received the time, who absorbed displaced tasks and whether the arrangement was sustained rather than infer all three from a calendar entry.

11. A diary that supports research without surveillance

A coordinator assigns two baseline and three endline teaching weekdays, balancing weekdays across the cohort and avoiding scheduled closures. Participants reconstruct the previous 04:00-04:00 day using a brief episode interface. A job-related episode receives an agreed-timetable flag and a task-family tag, including checking or repair. Primary activity durations close to 1,440 minutes; secondary responsibility tags are not added again. Unknown periods and incomplete days remain explicitly unknown.

The primary outcome sums job-related primary episodes outside the paid-timetable map, bounded between zero and 1,440 minutes. This operational category is not a legal determination that salaried work is unpaid. A teacher's map is agreed before baseline and cannot be revised after the outcome is known to make an apparent improvement. Research diary completion itself is separately recorded and supported with paid time so that measurement does not conceal the burden it introduces.

There are no keystroke logs, screen captures, message-content collections or continuous location feeds. Participants may correct a reconstruction, mark uncertainty or decline a day without employment consequences. Supervisors receive only disclosure-controlled school summaries. These protections are part of the intervention's credibility: a study that induces staff to underreport work to satisfy managers cannot provide a trustworthy institutional estimate, even if its timestamp system appears precise.

12. Independent quality measurement and stopping rules

Two parallel synthetic school-task packets are piloted for difficulty and accessibility before the main trial. Baseline and endline forms are counterbalanced. Tasks require a defensible communication, feedback plan or administrative explanation under explicit constraints, without real student records. Two human raters blind to school, condition and occasion score every artifact on factual fidelity, pedagogical appropriateness, accessibility, actionable reasoning and verification, each from zero to four. The primary quality score is the average of their 20-point totals.

This rubric is original and unvalidated. Pilot agreement, ceiling effects, form comparability and stakeholder interpretation must be established before choosing the final noninferiority margin. The candidate margin is a 0.5-point deterioration, evaluated using a lower one-sided 95% confidence limit. It cannot be defended merely because it produces convenient power. The trial's primary time test and this guardrail need separate precision assessment; the executed planning simulation covers only time.

Independent safety flags include invented sources, private-data exposure, discriminatory instruction and materially misleading safeguarding guidance. A third rater may adjudicate a safety concern but does not selectively overwrite the primary mean score. Credible material harm can pause the workflow immediately under the safety lead's documented authority. NIST's voluntary risk framework informs this governance architecture [16]; citation is not certification that a school or tool meets a federal standard.

13. Primary estimation, missingness and interpretation

The primary school outcome is its baseline cohort's mean change over assigned weekdays. Pair differences subtract the control-school change from the treatment-school change; the overall estimate averages the 24 pair differences. The default standard error is their sample standard deviation divided by the square root of the number of pairs, with a t interval using pair count minus one degrees of freedom. Equal weighting and the complete set of recruited schools remain fixed after assignment.

Missing assigned days are not dropped as though voluntary completion were random. The detailed analysis specification requires a validated multilevel multiple-imputation implementation before activation, fifty imputations of bounded daily endpoints, baseline and schedule predictors, and school dependence. Pooled estimates use Rubin's rules with a finite-sample adjustment [17]. No imputation has been run for this article. Its assumptions must be challenged through unfavorable shifts in missing outcomes and worst-case bounds using the outcome's known limits.

A complete-outcome pairwise randomization check addresses the Fisher sharp null, separately from inference about the mean assignment effect. Protocol deviations, staff departures, tool discontinuation and whole-school loss are reported rather than treated as grounds for exclusion. An incomplete or imprecise study can remain inconclusive. A superintendent should not interpret a large p value, a favorable point estimate or a successful process interview as interchangeable evidence of a sustainable dividend.

14. Executed design sensitivity

A reproducible artificial-data calculation generated 10,000 replicates for each combination of 16, 24 or 32 school pairs and an assumed standard deviation of pair differences of 30, 45 or 60 minutes. Independent Gaussian pair differences were generated under a zero-effect null and a hypothetical 20-minute reduction. Every replicate used the stated two-sided paired t test at alpha 0.05. The seed, software versions, scenario outputs and Monte Carlo standard errors are supplied.

For 24 pairs, simulated power is 0.8771, 0.5488 and 0.3561 as assumed variability increases from 30 to 60 minutes. Corresponding Monte Carlo standard errors are approximately 0.0033, 0.0050 and 0.0048. Across the nine scenarios, null rejection rates range from 0.0487 to 0.0550. These checks describe computation under the generating model, not empirical calibration of a school intervention. Approximate 80% minimum detectable effects for 24 pairs are 17.8, 26.7 and 35.6 minutes.

The findings materially constrain planning: 48 schools cannot be described as reliably sufficient without information about aggregate variability and the quality guardrail. ATUS does not provide repeated local school outcomes from which to estimate that parameter. A separate non-efficacy pilot must inform feasibility, missingness, clustering and instrument performance. Rutterford and colleagues explain why cluster designs require explicit sizing assumptions [14]. The present simulation exposes those assumptions rather than borrowing an impressive sample size from an unrelated survey.

14a. A pre-enrollment design activation gate

The 48-school configuration is a provisional planning scenario, not an approved adequately powered confirmatory sample. A separate measurement-only pilot in at least twelve school pairs, excluded from the main cohort, must estimate the variability of paired change scores for time and quality, diary completion and form reliability. The pilot introduces no experimental charter effect to be estimated. An independent statistician uses the upper 80% confidence bound for each paired-score standard deviation, rather than the most optimistic estimate, as the design nuisance parameter.

Before recruiting the main efficacy cohort, candidate designs from 24 to 120 district pairs are evaluated using the finalized estimator and missingness procedure. Allow for loss of 10% of recruited pairs and any greater loss suggested by the pilot. Select the smallest affordable design whose 50,000-replicate planning runs yield a lower 95% Monte Carlo confidence bound of at least 90% power both for a hypothetical 20-minute reduction and for the quality guardrail under zero true deterioration. These are declared policy/design assumptions, not known effects. Meeting both marginal power targets provides an at-least-80% joint-success lower bound by the union bound under the modeled alternative; it does not guarantee real-world power when assumptions fail.

The quality margin requires independent educator justification and successful form calibration before this gate can be evaluated. If no design within the funded cap meets the gate, the confirmatory trial does not launch. It may be redesigned and preregistered, or conducted explicitly as a feasibility study without a confirmatory benefit claim. Final numbers, justified margins, code, attrition assumptions and amendments are frozen in

registration before main-cohort recruitment and assignment. The included nine-scenario calculations demonstrate why this gate is necessary; they have not executed or passed it.

15. Falsification, distribution and rival explanations

The theory fails its institutional purpose if protected blocks appear on calendars while outside-timetable work, checking burden or output expectations rise. It also fails a sustainable-dividend claim if time improves but independently assessed quality crosses a justified harm margin. Increased AI use, more generated documents or satisfied leadership alone cannot rescue those counterexamples. Conversely, a null time contrast with wide uncertainty is inconclusive about the broader theory; the protocol must distinguish lack of precision from evidence of a practically negligible effect.

Displacement is a central rival explanation. The teacher benefit may be purchased by added work for assistants, principals or technical staff. Process records therefore identify reassigned obligations and obtain voluntary role-level accounts without exposing individual staff performance. Another rival is reporting pressure: treatment teachers may feel expected to demonstrate a benefit. Confidential collection, independent raters and explicit permission to report no gain reduce this threat but cannot remove it completely.

Exploratory distributional analyses examine contracted fraction, caregiving responsibilities and baseline time burden with consent and disclosure safeguards. Their interaction intervals are reported without ranking small demographic groups. A mean improvement can conceal worsening for those who already have the least discretionary time. The proposed interpretation consequently combines the assignment contrast with a distributional account and negative cases, while reserving confirmatory status for the declared primary endpoint.

Method and data statement

The completed empirical components are a secondary-file integrity and descriptive audit of pinned ATUS 2025 files and synthetic matched-pair planning simulations. Primary participant outcomes, recruited sample counts, registration and ethics approval do not exist in this package. The downloadable Python program executes those two completed components; it does not implement or claim to have fitted the future imputation model, collected diaries, randomized actual schools or evaluated the primary quality endpoint.

The public materials include the source manifest, exact hashes, result JSON, instruments, study specification, statistical plan, requirements and tests. Original respondent files remain at the BLS source rather than being repackaged into the journal. No local personal

records are included. A future registered protocol must identify sponsor responsibilities, compensation, data-controller arrangements, access limits, retention and independent analysis roles. SPIRIT provides a reporting scaffold [13], not an assertion that this school study is a registered clinical trial.

The ATUS source is contextual and methodological. No respondent is reidentified, contacted or linked to a school, ORCID record or journal account. A paper agent may retrieve these statements and source citations, but its response is not a new analysis. Reproduction requires the pinned files and reported software; a changed source hash triggers explicit refusal rather than silently accepting a different release as the same evidence.

17. Limitations

The national teacher domain is small and excludes occupations that cannot be separated cleanly. A single day's main-job category lacks the local constructs needed to estimate net automation effects. The 2025 shutdown creates a specific seasonal observation gap, while descriptive ratio means lack replicate-weight uncertainty in this edition. Nothing in these audit results establishes national after-hours work or a baseline AI effect.

The proposed trial also faces unblinded behavior, diary recall error, outcome missingness, school contamination and limited transportability from consenting districts. Synthetic quality tasks make independent blinding more feasible but may not capture all consequential classroom decisions. The charter combines a block, a demand-removal agreement and an appeal route; assignment identifies the package, not each component. Twelve weeks cannot establish long-term retention, labor-market preparation or institutional transformation.

The simulations omit missingness, unequal size, interference and joint quality power. Normal pair differences and assumed variability may be unrealistic. Before recruitment, the local pilot and external review may require a larger design, different task packets or revised margins. Publishing the protocol makes these dependencies inspectable; it does not solve them by declaration. These limits remain part of the scholarly contribution because they specify where a future claim would exceed the evidence.

AI, authorship and editorial disclosure

The named human author is Dr. Freedom Cheteni. This is a new primary-study protocol corresponding to conceptual article A02, not a revised version that inherits that article's review. No human peer review, acceptance, registration, institutional collaboration or participant consent is asserted for this new text. The publisher is The Research University, and the venue is the Journal for School Superintendents. Citations retain original authors, dataset agencies and source versions.

OpenAI's ChatGPT supported literature discovery, drafting, code preparation and computational checking [15]. No specific model version is asserted. AI is neither an author nor a human peer reviewer. The human author and independent editors remain responsible for approving claims, executing the proposed study and documenting review. Source-derived reasoning, actual audit results and artificial simulations are labeled separately, allowing subsequent readers to challenge each on its own terms.

19. Implications for superintendent decisions

The study turns a familiar technology question into a concrete governance test. Instead of asking only whether staff can produce drafts sooner, it asks whether a school can defend the allocation of the resulting human effort. That requires naming the tasks removed, recognizing verification labor, protecting discretion and permitting an unfavorable finding. The practical output would be an estimate with uncertainty, a quality boundary and a visible account of displaced work.

If a future study supports a sustainable dividend, replication should vary school contexts, task families and leadership constraints before broad scaling. If it does not, the record should distinguish inadequate precision, failed implementation, shifted work and actual quality harm. Each implies a different next decision. Movement Thinking is operational here when a school can revise its own arrangements in response to such evidence, including the decision to stop a technically impressive system that does not improve the conditions of educational work.

Research materials and reproducibility

Inspect the prospective study specification, original draft instruments, executed public-file audit, design simulations and documented tools. No local participant trial outcomes are reported.

Complete research package · ZIP

Research guide

Study specification

Source and variable audit

Diary, charter and quality instruments

Statistical analysis specification

Executable source-audit and planning tools

Executed ATUS source audit

Artificial-data design sensitivity

Methods tests

Read the corresponding conceptual foundation

Evidence and limits

Claim	Evidence	Limit
Task acceleration does not establish net institutional time release.	Distinct production, checking, total-work and allocation measures in an active-comparator protocol. [1, 9, 10, 11]	The intervention has not been tested in recruited schools.
The audited files support diary reconstruction but do not measure school AI use.	6,146 unique respondents,114,151 unique episodes, all diaries close to 1,440 minutes; 120 selected teacher occupations. [2, 3, 4, 5, 6, 8]	Small occupation domain, one diary per person, no school linkage, no AI exposure or contract-hour variable.
Autumn 2025 ATUS observations have a specific shutdown gap.	Official notice and zero audited diaries for September 30-November 11. [7]	Weighting cannot recover unobserved seasonal work behavior.
The 48-school planning design is not established as sufficiently powered.	10,000 replicates per nine scenarios yield 35.6-87.7% power at 24 pairs. [12, 13, 14, 17]	Assumed Gaussian pair differences, no missingness, quality or joint-success power; mandatory independent pilot and activation gate before main recruitment.

References and rights

[1] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only. Rights retained by the original rights holder; no text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[2] U.S. Bureau of Labor Statistics. American Time Use Survey - 2025 Microdata Files. American Time Use Survey, 2025 single-year release. 2026-06-25.

https://www.bls.gov/tus/data/datafiles-2025.htm

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[3] U.S. Bureau of Labor Statistics. American Time Use Survey Data Dictionary: 2025 Interview Data. ATUS data documentation. 2026-06.

https://www.bls.gov/tus/dictionaries/atusintcodebk25.pdf

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[4] U.S. Bureau of Labor Statistics. American Time Use Survey User’s Guide: Understanding ATUS 2003 to 2025. ATUS User’s Guide, September 2026 version. 2026-09.

https://www.bls.gov/tus/atususersguide.pdf

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[5] U.S. Census Bureau; U.S. Bureau of Labor Statistics. 2018 Census Occupation Classification Codes. 2018 Census occupation classification, official BLS-hosted codes. 2018 classification; accessed 2026-10-15.

https://www.bls.gov/tus/iocodes/census18ocodes.pdf

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[6] U.S. Bureau of Labor Statistics. American Time Use Survey Activity Lexicon 2025: no activity examples. ATUS 2025 activity coding lexicon. 2025 lexicon; accessed 2026-10-15.

https://www.bls.gov/tus/lexicons/lexiconnoex2025.pdf

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[7] U.S. Bureau of Labor Statistics. Federal Government Shutdown: 2025 American Time Use Survey. ATUS 2026 methodological notice. 2026; accessed 2026-10-15.

https://www.bls.gov/tus/notices/2026/october2025shutdown.htm

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[8] U.S. Bureau of Labor Statistics. How to use ATUS microdata files. ATUS official microdata instructions. Accessed 2026-10-15.

https://www.bls.gov/tus/other-documentation/howto.htm

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[9] Shakked Noy; Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381(6654), 187-192. 2023-07-14.

https://doi.org/10.1126/science.adh2586

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[10] Erik Brynjolfsson; Danielle Li; Lindsey Raymond. Generative AI at Work. The Quarterly Journal of Economics, 140(2), 889-942. 2025-02-04.

https://doi.org/10.1093/qje/qjae044

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[11] Jasper Maas; Simone Schoch; Urte Scholz; Pamela Rackow; Julia Schüler; Mirko Wegner; Roger Keller. Teachers' perceived time pressure, emotional exhaustion and the role of social support from the school principal. Social Psychology of Education. 2021-03-11.

https://doi.org/10.1007/s11218-020-09605-8

CC BY 4.0. © The Author(s) 2021

Authoritative source / evidence | checked 2026-10-15

[12] Marion K. Campbell; Gilda Piaggio; Diana R. Elbourne; Douglas G. Altman; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ, 345, e5661. 2012-09-04.

https://doi.org/10.1136/bmj.e5661

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[13] An-Wen Chan; Isabelle Boutron; Sally Hopewell; David Moher; Kenneth F Schulz; Gary S Collins; Ruth Tunn; Rakesh Aggarwal; Michael Berkwits; Jesse A Berlin; Nita Bhandari; Nancy J Butcher; Marion K Campbell; Runcie C W Chidebe; Diana R Elbourne; Andrew J Farmer; Dean A Fergusson; Robert M Golub; Steven N Goodman; Tammy C Hoffmann; John P A Ioannidis; Brennan C Kahan; Rachel L Knowles; Sarah E Lamb; Steff Lewis; Elizabeth Loder; Martin Offringa; Philippe Ravaud; Dawn P Richards; Frank W Rockhold; David L Schriger; Nandi L Siegfried; Sophie Staniszewska; Rod S Taylor; Lehana Thabane; David J Torgerson; Sunita Vohra; Ian R White; Asbjørn Hróbjartsson. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ, 389, e081477. 2025-04-28.

https://doi.org/10.1136/bmj-2024-081477

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[14] Clare Rutterford; Andrew Copas; Sandra Eldridge. Methods for sample size determination in cluster randomized trials. International Journal of Epidemiology, 44(3), 1051-1067. 2015.

https://doi.org/10.1093/ije/dyv113

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

[15] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool citation, not an article author. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no model version asserted. Not a human peer reviewer.

Authoritative source / evidence | checked 2026-10-15

[16] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. 2023-01-26.

https://doi.org/10.6028/NIST.AI.100-1

Citation only; no Creative Commons reuse permission asserted. Original authors, agencies and rights holders retain applicable rights; source files are not redistributed.

Authoritative source / evidence | checked 2026-10-15

[17] John Barnard; Donald B. Rubin. Miscellanea. Small-sample degrees of freedom with multiple imputation. Biometrika, 86(4), 948-955. 1999-12-01.

https://doi.org/10.1093/biomet/86.4.948

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-15

This original primary-study protocol and completed methods investigation credit each cited author and data agency. BLS public-use source files were analyzed for a documented secondary audit and are not redistributed in this research package. The source studies are not republished as work by this manuscript author. The proposed new text license requires actual human author approval and the journal’s licensing decision; no source reuse license is inferred from free access.

A08 | PRIMARY-STUDY PROTOCOL AND REPRODUCIBLE METHODS RESEARCH

Accountable Intelligence at the Point of Decision: A Europe PMC Corpus Audit and Randomized Protocol for Evaluating Paper Agents in School-District Evidence Work

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-16 | Edition 1.1 | 4,191 article words

Primary-study protocol and reproducible methods; separate human editorial record.

Keywords: Movement Thinking, paper agents, school district governance, evidence fidelity, Europe PMC, blinded assessment, randomized protocol, provenance

Read online / article companion / citation exports

Contribution

An original human decision-support trial protocol, exact-version Europe PMC corpus audit and reproducible artificial-data design analysis that distinguish source availability from warranted educational judgment.

Prospective adult randomized benchmark; completed 36-record core/JATS source-feasibility audit and nine-scenario 10,000-replicate artificial-data design analysis. No participant/agent performance or educational effects reported.

Abstract

A paper agent changes how a school leader encounters research, but convenient access does not establish better judgment. This article develops a prospective randomized human study of an accountable paper-agent interface and reports two completed methodological analyses. The interface is evaluated against a source-readable active control using identical evidence packets, with human evidence-fidelity rather than agreement with an agent as the primary outcome. A bounded Europe PMC audit retrieves the first 36 records from a 98-hit leadership query, checks full-text JATS, article-level rights, identifiers and available status signals, and preserves source hashes. Twenty-three records pass the declared source-construction rules; thirteen remain excluded or unresolved. An indexed DOI and the downloaded JATS identify different F1000Research versions, illustrating why citation resolution alone is insufficient. The proposed study recruits 160 adult education professionals, randomizes within four strata, and administers twelve fictional district decisions. Blinded human raters score causal restraint, population and measure fidelity, citation entailment, uncertainty, rights and version interpretation, and bounded action. Artificial-data simulations show that design sensitivity depends strongly on within-person case correlation. No participant was recruited, no human or agent performance outcome was observed, and no educational or clinical benefit is inferred. The contribution is a reproducible design that makes accountable intelligence falsifiable at the point where a source becomes a decision.

1. From an intelligent library to accountable evidence work

School districts increasingly encounter research through search engines, summaries and conversational interfaces rather than through a linear reading of journal issues. An agent attached to a paper can help a superintendent find a methods passage or compare a claim with a limitation. The same interface can erase an important distinction: an observational association can become an intervention promise, a self-report measure can become an observed student outcome, and a readable paper can become a supposedly reusable asset. Institutional seriousness therefore requires evaluation of what a human can justify after using the interface, not simply whether an answer sounds scholarly.

Movement Thinking provides the intellectual starting point: educational work should remain open to revision, participation and changing conditions [1]. This study translates that orientation into a narrow governance hypothesis. A source-grounded interface may help professionals move between question, evidence, objection and revised judgment without losing attribution or decision responsibility. That is an original operational proposition, not a claim that the earlier philosophy already specified or validated this

experiment. The companion conceptual article, “The District as Accountable Intelligence,” remains a separate scholarly object; this protocol adds an empirical route for testing one of its claims.

The unit of practical concern is the decision brief. It records what is known, what remains uncertain, which evidence supports the recommendation and who must decide. A correct DOI is necessary bibliographic infrastructure but insufficient warrant. A paper about general school leadership cannot, merely through an interactive interface, establish the correct response to an individual child’s mental-health needs. The present benchmark is deliberately limited to evidence interpretation in fictional district situations, with no actual student, personnel, disciplinary or clinical decision delegated to software.

2. A falsifiable theory of accountable assistance

The proposed mechanism is reduced friction in inspecting evidence. The treatment makes the source passage, version and applicability limit reachable from an answer, then allows a participant to revise an initial brief. The alternative mechanism is automation deference: fluent language may increase confidence while weakening inspection. Both are plausible. The protocol therefore measures warranted reasoning, unjustified confidence and verification time separately. More clicks, more citations and higher satisfaction are not interchangeable indicators of better judgment.

The hypothesis is that assignment to the accountable interface improves the mean evidence-fidelity score across a fixed set of cases compared with an equally readable source-only interface. The causal target is the incremental effect of this specific interface, training and frozen evidence configuration on recruited adults completing these cases. It is not the effect of every AI model, a district-wide implementation programme, future employment, student attainment or clinical care. A positive benchmark result would justify a further field study rather than a claim of educational effectiveness.

The causal chain is assignment, inspection opportunity, human interpretation and decision brief. Intermediate interface events are measured as process variables. They are not randomized mediators and will not be assigned causal mediation interpretations merely because they correlate with the final score. A participant can reach the right decision without opening the suggested passage, or open it without understanding it. The final human response, assessed against independent criteria, remains the endpoint.

3. Why Europe PMC is an appropriate but bounded database

Europe PMC supplies a documented API for core publication metadata and open-access full-text XML [2]. Its open-access documentation expressly directs users to each article's terms; membership of the subset does not provide a uniform reuse license [3]. These features make it useful for testing the institutional problem at issue. A record contains multiple kinds of evidence - bibliographic identity, accessible text, permissions and linked publication-status information - that can agree or conflict.

The database is particularly useful for school mental health, implementation and wellbeing. It is not a complete index of education policy or school-administration scholarship. The audit is consequently a feasibility investigation for constructing a benchmark, not an exhaustive review of leadership research. The query, response, ordering, cut-off and exclusions are preserved so that selection cannot later be described as comprehensive. Subsequent trial construction requires a separate, independently reviewed case set; the present audit establishes a candidate pool rather than a validated reference standard.

The study does not merge Europe PMC records with TALIS respondents, SEDA districts or student files. Published studies may themselves use such databases, but this investigation treats those papers as scholarly sources with their original designs. The original researchers retain their authorship. Dr. Freedom Cheteni's contribution is the new evaluation protocol, corpus audit, instruments and analysis specification, not the empirical work reported in imported articles.

4. Completed corpus audit: retrieval and screening rules

The executed search combines the title/abstract phrases "school leadership" and "principal leadership," an initial publication-date window of 1 January 2017 through 15 October 2026, and Europe PMC's open-access and full-text indicators. It excludes nursing, medical-school and medical-education terms. The exact machine-readable expression and encoded API URL accompany the methods package. The response contains 98 hits. The audit intentionally stops at its first 36 relevance-ranked records, preserving the response hash and all selected identifiers. It does not inspect the remaining 62 hits or imply that their eligibility matches the audited subset.

For every selected record, the script retrieves JATS through the supported full-text API, computes SHA-256 over the received bytes, parses author contributions separately from editorial groups, and checks the article's own permissions. It records publication dates as supplied, including incomplete date components. Europe PMC's normalized

first-publication date is retained as an indexed field, not silently promoted into an exact version-release day. Core DOI and JATS article DOI are compared, while a PMCID plus byte hash identifies the precise retrieved snapshot.

A source passes construction rules only when the JATS identifies a research article, an article-level CC BY, CC BY-SA or CC0 version is identifiable, a nonempty body is present, the indexed DOI agrees with a JATS article DOI, no screened withdrawal or retraction signal is detected, and the topic fits the declared decision-support boundary. A data-only CC0 waiver cannot license the article. NC/ND licenses, unversioned rights statements and unresolved version disagreement remain outside the construction pool. No third-party full text or figures are redistributed with this article.

Three relevance or design exclusions are declared in code: university physical-education trainee monitoring, tele-delivered caregiver coaching, and a policy-review context excluded from the primary-research benchmark pool. These are explicit scope choices rather than claims that the excluded work lacks value. Reviews can inform the article but do not enter its research-source construction pool. Human case developers must still inspect the included sources for question suitability, contested interpretations and any restrictions attaching to particular third-party materials.

5. Completed corpus results and their epistemic limits

All 36 selected JATS documents were retrieved and parsed; the snapshot contains 36 distinct PMCIDs and 36 distinct nonempty indexed DOIs. Thirty-four documents carry a research-article JATS type, one a review-article type and one a systematic-review type. Twenty-nine have an identifiable article-level eligible license URL, four carry NC or ND restrictions, and three contain an insufficiently versioned license statement for the strict construction rule. These are document classifications, not judgments about the quality of the underlying science.

Combining design, relevance, rights, identifier and available status rules leaves 23 source-eligible candidates and 13 excluded or unresolved records. The categories overlap: the version-conflicted F1000Research item has identifiable CC BY 4.0 rights, yet fails identifier agreement. Its indexed DOI ends in .2 while the retrieved JATS article DOI ends in .3 [12]. The discrepancy does not establish misconduct or an invalid publication; it establishes that a user must resolve the version before quoting or linking a supposedly exact snapshot. The JATS version history identifies version 3 as 8 October 2026 and version 2 as 19 August 2026; the audit retains both identifiers rather than overwriting one with the other. A separate date discrepancy appears for the whole-school implementation paper: the indexed field is 1 January 2026 while the downloaded JATS publication date is 3

December 2025 [11]. The reference uses the document date and the audit preserves both, without disguising the disagreement.

No retraction, withdrawal or expression-of-concern signal was observed in the screened core publication types, title prefixes, available correction relations and JATS related-article fields. This is a limited, dated observation, not an independent certification of integrity. A later notice, an unindexed notice or a publisher correction can change the appropriate interpretation. Status must be checked again before trial freeze and before consequential deployment. The corpus audit reports availability and provenance; it does not measure retrieval quality, agent answer accuracy, human decision improvement or causal educational effects.

6. Source designs that make demanding test cases possible

Several sources show why the benchmark must examine inference rather than citation decoration. Lyon and colleagues' School Implementation Leadership Scale concerns measurement development in school implementation settings [4]. An agent should identify the measure and its setting rather than treat it as certification of a software platform. Låftman, Östberg and Modin's multilevel cyberbullying analysis supplies a different challenge: a relationship between leadership and reported cyberbullying does not license a claim that deploying an agent prevents bullying [5].

Bao's principal-inclusive-leadership analysis provides another opportunity to distinguish a statistical pathway from an intervention mechanism [6]. The benchmark can ask whether a district may promise a reduction in teacher turnover from a mediation result without a randomized local evaluation. Daly, Resnikoff and Litke's narrative review informs school mental-health leadership but remains a review, not a trial of a particular district programme [7]. It is cited for design interpretation and excluded from the primary-research construction pool, not represented as a failed source.

The HELM pilot provides an important contrast because it actually randomized schools to leadership strategies, including an attention comparator [8]. The benchmark must preserve what was assigned, where, and what outcomes were measured. Its NC license also illustrates that stronger causal design and broader reuse permission are different properties. Ertem's crisis-leadership article has NC-ND rights in the retrieved version [9]. Both can be cited and linked here without being copied into a supposedly unrestricted agent corpus. No abbreviated reading of an "open access" badge substitutes for that distinction.

7. Prospective participants, setting and recruitment

The planning sample is 160 consenting adults whose professional roles involve interpreting research for schools: superintendents and district leaders, principals, teachers and education-support professionals. Four strata cross leadership role versus practitioner/support role with lower versus higher self-reported research-agent experience, defined through a prespecified screening threshold. Each stratum recruits 40 participants before assignment. These numbers are a design target, not an enrolled cohort or evidence of institutional agreements.

Recruitment occurs outside supervisory evaluation and admission decisions. Participation cannot affect employment, professional standing, course access or publication eligibility. Invitations explain that the tasks are fictional and that software can be incomplete or wrong. Compensation is fixed for time rather than performance or endorsement. Eligibility requires sufficient English reading proficiency for the fixed materials; this restricts generalizability and does not establish an inherent language requirement for school leadership. Accessible formats, keyboard operation and planned breaks must be available.

An independent ethics determination, consent materials, data-management agreement, intervention freeze and dated registration precede recruitment. None is claimed to exist for this proposal. No minors are recruited, no actual student records are entered, and no prompt may contain identifiable information about staff or families. A participant can stop without having an incomplete response converted into an allegation of incompetence. Individual performance is not reported to an employer.

8. Randomization, intervention and a credible comparator

A separate allocation custodian assigns 20 of the 40 participants in each stratum to each condition after screening and baseline measures are locked. The allocation is generated using a recorded random procedure and concealed from recruiters until enrolment is complete. Participants cannot be blinded to a conversational interface; outcome raters, case-standard assessors and the primary analyst remain masked to arm labels until scoring and the analysis plan are locked.

Both interfaces display the same source packet, full bibliographic attribution, readable licensed text where permitted, title navigation, ordinary text search, notice and rights panels, equivalent typography and a decision-brief form. Training time and task time are equal. The control therefore represents competent digital reading rather than an intentionally frustrating PDF. The treatment adds a bounded question interface with passage-linked responses, visible uncertainty, exact-source access and a clear separation

between source statements and interface interpretation. Responses cannot change the supplied evidence or submit a real district action.

Before the trial, the treatment manifest must record code revision, retrieval settings, model and version if any, prompt rules, corpus hashes and session limits. This article does not invent a deployed model identifier or claim that the proposed interface has already passed evaluation. If a deterministic retrieval companion is used, that configuration is recorded explicitly. An evaluation of generative assistance requires a separate frozen configuration, retained input/output logs under consent, and cost, privacy and reproducibility arrangements. Model changes during the confirmatory study require an amendment or a separately analyzed cohort, not silent substitution [13] [14].

9. Twelve cases and an independent reference standard

The fixed benchmark contains twelve short fictional district decisions: three requiring causal and design interpretation, three requiring population and measurement transfer, three requiring uncertainty and source-entailment judgments, and three requiring rights, version or notice interpretation. Every case receives the same six-dimensional scoring rubric, so a category label does not exempt it from the other obligations. Case order is balanced independently of treatment assignment. Participants receive approximately six minutes per case, with training, baseline questions and breaks bringing the session to roughly 100 minutes.

Independent education researchers and an information specialist develop the case reference standard before seeing any treatment-generated answer. They identify acceptable conclusions, alternative defensible answers, source locations and errors that would materially change a decision. They rate the standard independently, resolve disagreements with written reasons and retain unresolved interpretations as explicit ranges. The standard is not an agent’s preferred response and does not reward agreement with Cheteni’s philosophy. Pilot participants and pilot raters are distinct from the confirmatory cohort.

Some errors are better tested with transparent fictional evidence than with a false claim about a real researcher. If a retraction-notice case cannot be built from a verified actual notice, the packet uses a wholly fictional article and labels the entire notice exercise synthetic. It never attaches an invented retraction to a real DOI. Similarly, a fictional local budget or school cannot be mistaken for collected district data. Case release is delayed until scoring is complete to reduce coaching; the protocol, categories, rubric and source rules can be public earlier.

10. Primary outcome and blinded scoring

The primary outcome is each participant's mean evidence-fidelity score across twelve cases, on a 0-24 scale. This is a newly specified, unvalidated rubric: pilot content review, inter-rater reliability and ceiling checks are required and will not establish universal validity. Six dimensions contribute 0-4 points each: design and causal restraint; population, setting and measurement fidelity; citation entailment and source location; uncertainty and appropriate abstention; rights, version and notice interpretation; and a bounded, accountable recommendation. A high score requires a defensible path from evidence to action, not merely polished prose. Correctly stating that the packet cannot resolve a question can earn full credit.

Two trained human raters score every final brief independently. The primary case score is the mean of their total scores; the participant score is the unweighted mean across cases. Disagreement is reported dimension by dimension and as an absolute-agreement reliability estimate, with uncertainty. A third assessor reviews prespecified severe-error disagreements for quality assurance but does not selectively replace the primary mean with a favorable score. If an instrument defect is found, the revision is documented before confirmatory scoring resumes.

A secondary sentinel-error outcome counts unsupported causal guarantees, fabricated source claims, misattributed authorship, false claims of permission and acceptance of an actual withdrawn finding without qualification. Severity definitions and denominators are frozen. The sentinel rate is not blended into the primary score after results are known. Other secondary outcomes include completion time, source verification time, confidence calibration, accessibility barriers, workload and participant-rated usefulness. Faster completion is beneficial only alongside adequate fidelity and an acceptable severe-error profile.

11. Statistical estimand and primary analysis

Let each person's potential twelve-case mean under the source-only interface and under the accountable agent be defined for the frozen source and case set. The primary estimand is the equal-stratum mean difference between assigned conditions among the recruited adults, regardless of actual agent use. With equal planned stratum sizes this is also the equally weighted participant effect. It does not treat the 1,920 case responses as 1,920 independent people.

For complete cohorts, the estimator averages the four within-stratum treatment-minus-control differences. Its variance sums the eight independent arm/stratum

sample-variance terms, each divided by its cell size and by sixteen. Confidence intervals and the two-sided test use Welch-Satterthwaite degrees of freedom calculated from those terms. This permits unequal outcome variances rather than assigning an unjustified pooled residual degree of freedom. The downloadable tool implements the complete-cohort calculation and verifies actual-design cell counts; it refuses a silently altered design.

The primary analysis is assignment-based, with no exclusion for ignoring the agent, disagreeing with it or producing a low score. A secondary crossed-effects model includes participant and case intercepts, fixed treatment and stratum terms, and a treatment-by-case random component if identifiable. Its purpose is to describe heterogeneity and sensitivity to case selection. With only twelve purposively constructed cases, a random-case model cannot establish representativeness of all educational decisions. Domain and experience interactions are exploratory, with effect estimates and uncertainty rather than selective significance claims.

12. Missing data, multiplicity and change control

An unanswered case during an otherwise completed timed task receives zero only when the predeclared task definition treats failure to produce a brief as observed performance. A lost submission, technical failure, withdrawn record or interrupted session is missing, not automatically a zero. The distinction is logged before arm labels are revealed. The confirmatory analysis uses fifty multiple imputations of missing case totals from a bounded predictive model including assignment, stratum, baseline evidence-literacy score, case identity and observed case responses. The analysis specification defines diagnostics, pooling and sensitivity departures; the supplied small Python estimator does not claim to implement that imputation model.

Sensitivity analyses shift imputed treatment values downward and control values upward by one and two rubric points, then report bounded worst-case contrasts. The report provides missingness by arm, reason and case and distinguishes retention effects from interpretation effects. Pooling follows repeated-imputation rules with a small-sample degree-of-freedom adjustment rather than presenting imputed records as observed people [15]. Complete-case results remain secondary when missingness is present.

There is one confirmatory outcome and one primary comparison. Secondary families are grouped before analysis; sentinel error and workload are explicitly evaluated as safety-relevant descriptive outcomes, not hidden when the primary score improves. Any multiplicity-adjusted exploratory tests are identified. The registration locks outcomes, code revision, case-standard hash and missing-data rules before recruitment. Changes

needed for accessibility or safety are permitted but documented with timing and reasons, allowing readers to distinguish correction from outcome-driven redesign.

13. Completed design-sensitivity analysis

The executed simulation uses 10,000 artificial datasets for each combination of 80, 160 or 240 adults and within-person case correlations of 0.10, 0.30 or 0.50. There are four equal strata, balanced assignment within each stratum and twelve equally weighted cases. Marginal case variance is standardized to one. The imposed treatment effect is 0.25 case standard deviations. The model is Gaussian, assumes complete outcomes and constant treatment effects, and includes neither ceiling effects nor case-specific treatment heterogeneity. These assumptions are inputs, not facts estimated from district professionals.

For the 160-person planning design, simulated power is 96.55%, 74.56% and 57.01% as within-person correlation rises from 0.10 through 0.30 to 0.50. Monte Carlo standard errors are approximately 0.18, 0.44 and 0.50 percentage points. Corresponding null rejection rates are 4.62%, 5.46% and 5.09%. These values show why counting twelve responses per adult as independent observations would exaggerate precision. Increasing the number of cases yields diminishing returns when stable person-level differences dominate.

The results do not establish that the proposed agent will produce the assumed effect or that 160 participants guarantee adequate power. A pilot should estimate score dispersion, ceiling frequency, case correlation and completion feasibility without testing the final efficacy claim. The investigator then chooses a funded confirmatory design before recruitment, reruns sensitivity calculations and registers any change. The public simulation reports seed, software versions, every scenario and Monte Carlo uncertainty; no simulated row is presented as an observed human response.

14. Adversarial interpretation, workload and institutional consequences

An interface should not receive credit for a correct reference beside an unsupported conclusion. Qian and colleagues' Shanghai analysis explicitly reports an exploratory moderation finding that does not reproduce under one scoring alternative [10]. A defensible benchmark response preserves that uncertainty rather than selecting whichever result supports an enthusiastic recommendation. Burns and colleagues' whole-school implementation study requires attention to its embedded case design and actual contributing schools [11]. Neither source becomes a universal district intervention result through citation alone.

The benchmark tests authority boundaries as well as evidence recall. When asked to approve an intervention, diagnose an individual, change a publication status or claim that AI completed peer review, a successful response explains the evidential and human process required. The participant's brief should likewise retain the named accountable professional. Software may draft or retrieve; it cannot bear scholarly responsibility or become the independent human peer whose review underwrites a journal decision.

A useful interface may still fail the institutional test if verification requires excessive effort or produces inaccessible interaction. Source-checking time, assistance requests and screen-reader or keyboard barriers are recorded separately from total time. A future deployment threshold must be set before a field trial and include severe-error bounds, maintenance cost and a correction route. This protocol does not convert a laboratory score into permission to automate personnel evaluation, admission, safeguarding or clinical action. Those decisions require distinct governance, evidence and authorization.

15. Limitations and conditions that would contradict the hypothesis

The corpus is a small relevance-ranked slice of a health-oriented scholarly index, not a probability sample of school research. Rights are interpreted conservatively from exact downloaded versions; unresolved wording may exclude material that a later permissions check would clear. Conversely, a clear article license does not remove restrictions on embedded third-party content. The audit is not a comprehensive retraction service, and its status observations expire as the scholarly record changes.

The human experiment is prospective. A fixed-case task cannot establish changes in school practice, equity, student achievement, wellbeing or clinical outcomes. English-language recruitment, volunteer participation, professional experience and a structured time limit

restrict transfer. Participants may recognize the software condition, and apparently neutral briefs may reveal its wording to raters. Masking success is assessed by asking raters to guess arm with confidence after scoring, without using those guesses to edit unfavorable results.

The hypothesis would be contradicted if the treatment failed to improve fidelity within practically relevant uncertainty, increased severe unsupported claims, raised confidence without accuracy, or saved time by discouraging necessary checking. A trade-off cannot be settled through a composite score chosen afterward. The strongest result would be improved warranted reasoning with usable verification, followed by independent replication. Even that result would support the evaluated configuration and case domain, not a claim that the institution possesses an infallible intelligence layer.

Method and data statement

This original article reports a completed database and full-text feasibility audit and a completed artificial-data design analysis, alongside a prospective primary human study. There were zero trial participants and no observed treatment outcomes. Europe PMC core metadata and 36 full-text JATS snapshots were inspected through documented API endpoints; raw third-party documents remain outside the public package. Public audit records preserve original authors, identifiers, dates, licenses, hashes, exclusions and access limitations. The methods package includes instruments, an analysis specification, executable audit and simulation code, results and tests. Exact byte identity supports reproducibility, not truth, permission or endorsement. Registration, ethics approval and final case validation remain prerequisites for collecting primary participant data [13].

AI, authorship and editorial disclosure

ChatGPT assisted with research organization, code, source inspection and drafting on 15 October 2026 [16]. The tools executed real downloads, metadata parsing and explicitly artificial simulations; they did not recruit participants or supply a human peer-review report. Dr. Freedom Cheteni is the named human author, with responsibility for final approval and scholarly accountability. Source authors retain their contributions, and neither Europe PMC nor the source institutions are coauthors or endorsers of this protocol. This is a new companion manuscript to A03. The earlier conceptual article’s publisher-confirmed review does not constitute review, acceptance or authorization of this new protocol. No acceptance date, publication date, registration identifier, ethics approval or DOI is invented. The Editor in Chief’s authorship requires independent human handling of any subsequent editorial decision.

Research materials and reproducibility

Inspect the prospective study specification, original draft instruments, executed public-file audit, design simulations and documented tools. No local participant trial outcomes are reported.

Complete research package · ZIP

Methods guide

Study specification

Assessment instruments and scoring manual

Statistical analysis specification

Executable audit and design tools · Python

Completed corpus audit

Executed design simulations

Granular source-verification register

Reproducibility tests

Read the corresponding conceptual foundation

Evidence and limits

Claim	Evidence	Limit
A paper-agent interface can be evaluated without delegating human authority.	Four-stratum adult randomization, source-readable active control and independently blinded human evidence-fidelity scoring. [1, 13, 14]	Prospective design; no human performance effect observed.
Open access and exact-version reuse eligibility are distinct.	Executed audit:36 JATS bodies, 29 eligible article-license URLs, 4 restricted and 3 unresolved rights records; 23 pass all construction rules. [2, 3, 8, 9]	A bounded 36 of 98 query slice, not a systematic review or integrity certification.
A valid identifier can still point to a different source version.	Observed F1000Research indexed DOI .2 versus downloaded JATS article DOI .3; both retained. [12]	Version conflict requires human resolution; no allegation of misconduct or source invalidity.
Twelve responses per adult do not supply twelve independent participants.	Executed simulations under three within-person case correlations, with 10,000 replicates/scenario. [13, 15]	Artificial Gaussian assumptions; no observed agent effect, dropout or case-treatment heterogeneity.

References and rights

[1] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only; no republication rights asserted. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[2] Europe PMC. Europe PMC RESTful Web Service. EMBL-EBI Europe PMC. Accessed 2026-10-15.

https://europepmc.org/RestfulWebService

Citation only; no republication rights asserted. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[3] Europe PMC. Open access subset: copyright and web services. EMBL-EBI Europe PMC. Accessed 2026-10-15.

https://europepmc.org/downloads/openaccess

Citation only; no republication rights asserted. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[4] Aaron R Lyon; Catherine M Corbin; Eric C Brown; Mark G Ehrhart; Jill Locke; Chayna Davis; Elissa Picozzi; Gregory A Aarons; Clayton R Cook. Leading the charge in the education sector: development and validation of the School Implementation Leadership Scale (SILS). Implementation Science : IS. 2022-07-19.

https://doi.org/10.1186/s13012-022-01222-7

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[5] Sara B Låftman; Viveca Östberg; Bitte Modin. School Leadership and Cyberbullying - A Multilevel Analysis. International Journal of Environmental Research and Public Health. 2017-10-15.

https://doi.org/10.3390/ijerph14101226

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[6] Yuangen Bao. A dual mediation model of the association between principal inclusive leadership and teacher innovative behavior. Scientific Reports. 2024-05-29.

https://doi.org/10.1038/s41598-024-63332-0

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[7] Brian P Daly; Annie Resnikoff; Shannon Litke. Effective School Leadership for Supporting Students’ Mental Health: Findings from a Narrative Literature Review. Behavioral Sciences. 2025-01-01.

https://doi.org/10.3390/bs15010036

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[8] Jill Locke; Catherine M Corbin; Roger Goosey; Vaughan K Collins; Mark G Ehrhart; Kurt Hatch; Christine Espeland; Aaron R Lyon. Not getting better but not getting worse: A cluster randomized controlled pilot trial of a leadership implementation strategy. Implementation Research and Practice. 2025-01-29.

https://doi.org/10.1177/26334895241312405

Citation only; restricted or conflicting source rights statements. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[9] Hasan Yücel Ertem. School leadership fostering mental health in the times of crisis: synthesis of school principals' views and PISA 2022. BMC Psychology. 2024-11-26.

https://doi.org/10.1186/s40359-024-02195-6

Citation only; restricted or conflicting source rights statements. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[10] Mengting Qian; Chunshun Yan; Xinyi Wang. Perceived school-wide teacher-student relations and teacher job satisfaction: An affective events theory analysis of Shanghai teachers. PLOS One. 2026-09-03.

https://doi.org/10.1371/journal.pone.0357279

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[11] Sharyn Burns; Roisin Glasgow-Collins; Hanna Saltis; Jacqueline Hendriks. Implementation of a Whole-School Relationship and Sexuality Education Project in Western Australian Schools: Evaluation Data From a Multiple, Embedded Case Study. Health Promotion Journal of Australia. 2025-12-03.

https://doi.org/10.1002/hpja.70138

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[12] Yihenew Belay; Mateb Tafere. Leading beyond the Job Description: Investigating the Association between Principals' Leadership Styles and Teachers' Organizational Citizenship Behaviors in North Wollo Secondary Schools. F1000Research. 2026-10-08 (downloaded JATS version 3).

https://doi.org/10.12688/f1000research.183874.3

CC BY 4.0. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[13] An-Wen Chan; Isabelle Boutron; Sally Hopewell; David Moher; Kenneth F Schulz; Gary S Collins; Ruth Tunn; Rakesh Aggarwal; Michael Berkwits; Jesse A Berlin; Nita Bhandari; Nancy J Butcher; Marion K Campbell; Runcie C W Chidebe; Diana R Elbourne; Andrew J Farmer; Dean A Fergusson; Robert M Golub; Steven N Goodman; Tammy C Hoffmann; John P A Ioannidis; Brennan C Kahan; Rachel L Knowles; Sarah E Lamb; Steff Lewis; Elizabeth Loder; Martin Offringa; Philippe Ravaud; Dawn P Richards; Frank W Rockhold; David L Schriger; Nandi L Siegfried; Sophie Staniszewska; Rod S Taylor; Lehana Thabane; David J Torgerson; Sunita Vohra; Ian R White; Asbjørn Hróbjartsson. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. The BMJ. 2025-04-28.

https://doi.org/10.1136/bmj-2024-081477

Citation only; restricted or conflicting source rights statements. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[14] Tammy C Hoffmann; Paul P Glasziou; Isabelle Boutron; Ruairidh Milne; Rafael Perera; David Moher; Douglas G Altman; Virginia Barbour; Helen Macdonald; Marie Johnston; Sarah E Lamb; Mary Dixon-Woods; Peter McCulloch; Jeremy C Wyatt; An-Wen Chan; Susan Michie. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014-03-07.

https://doi.org/10.1136/bmj.g1687

Citation only; no republication rights asserted. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[15] John Barnard; Donald B. Rubin. Small-sample degrees of freedom with multiple imputation. Biometrika. 1999-12.

https://doi.org/10.1093/biomet/86.4.948

Citation only; no republication rights asserted. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

[16] OpenAI. ChatGPT [AI-assisted drafting and coding software]. Software/tool citation, not an author or peer reviewer. Accessed 2026-10-15.

https://chatgpt.com/

Tool use disclosed; no exact model version asserted. Original authors and rights holders retain their contributions; no source text or figures reproduced.

Authoritative source / evidence | checked 2026-10-15

Original primary-study protocol and reproducible corpus/design methods by Dr. Freedom Cheteni. Original source authors, titles, venues, dates, versions and rights remain distinct. No third-party full text, figure, participant record or reviewer report is redistributed. New-manuscript licensing and author approval are separate from source permissions.

A09 | PRIMARY-STUDY PROTOCOL AND REPRODUCIBLE METHODS RESEARCH

From Presence to Revisable Explanation: A Factorial Primary-Study Protocol for Spatial Transfer with NASA Exoplanet Data

Agathe Bertin | First author

Hasso Plattner Institute of Design; Santa Clara University

Dr. Freedom Cheteni | Second author · Principal Investigator

ORCID 0009-0006-3507-0094

Prepared 2026-10-16 | Edition 1.1 | 4,529 article words

Primary-study protocol and reproducible methods; separate human editorial record.

Keywords: Movement Thinking, spatial transfer, factorial trial, immersive learning, exoplanet data, evidence revision, accessibility

Read online / article companion / citation exports

Contribution

A prospective blocked factorial test of evidence-revision instruction across immersive and desktop offers, with an executed NASA stimulus audit and reproducible interaction-power analysis.

Original prospective study specification; executed public astronomical-file feasibility audit and artificial-data simulations; no participant outcomes.

Abstract

Immersive school environments can increase presence without establishing that learners can explain unfamiliar phenomena after the display is removed. This article converts Movement Thinking's prediction-action-revision proposition into a prospective, blocked factorial study of spatial transfer. It also reports two completed methodological investigations: a reproducible NASA Exoplanet Archive stimulus-feasibility audit and an artificial-data sensitivity analysis. The proposed study recruits one science instructor in each of 32 schools, with four course sections per instructor and a fixed baseline cohort of 24 learners per section. Within each instructor block, sections are assigned to immersive or desktop representation crossed with explicit evidence-revision or time-matched guided comparison. The primary estimand is the display-by-instruction interaction on a delayed, blinded, 20-point transfer assessment. Astronomical records supply authentic task parameters, never student outcomes. A frozen public-query response contains 6,451 default parameter rows; sequential restrictions identify 1,692 core stimulus candidates and 1,220 with the specified two-sided uncertainty information. The audit detects candidate parameter solutions, projected-separation hazards and missing eccentricities that a superficial download would conceal. Ten thousand artificial replicates per design scenario expose the cost of powering an interaction rather than a main effect. No learner has been recruited or evaluated in this work. The contribution is an inspectable route from spatial-learning theory to primary evidence: matched content, a defensible contrast, explicit model assumptions, accessible participation, original scoring instruments and reproducible methods.

1. The superintendent's question is what survives the display

A learner who can navigate a spectacular planetary scene may remain unable to explain why an eccentric orbit does not produce equal distances in equal times. A learner who studies a quiet diagram may develop that explanation well. District leaders need evidence that distinguishes those possibilities before treating visual engagement as a return on investment. The relevant public obligation concerns the durability and portability of understanding, together with the cost and accessibility of the conditions that support it.

The companion conceptual article, Beyond the Rendered Classroom, proposes that movement between prediction, action, feedback and revision is an educational mechanism requiring investigation. Cheteni's philosophical account supplies its provenance [1]. The present article contributes a different object: a primary-study protocol and completed methodological analyses that make a particular test feasible. It does not reclassify the

conceptual argument as an experiment or claim that a compelling philosophy establishes an intervention effect.

The research problem is therefore conditional. Does an explicit evidence-revision method change the advantage, disadvantage or neutrality of offering immersion when disciplinary information and learning time are held comparable? A finding that revision instruction works equally well on both displays would matter. So would an interaction favoring desktop delivery, or greater discomfort without a transfer benefit. The protocol makes these outcomes interpretable rather than defining success as whatever a technology happens to improve.

2. Theory that permits an unfavorable finding

The Cognitive Affective Model of Immersive Learning distinguishes technological features from psychological processes and identifies plausible routes through which presence and agency may support or hinder learning [2]. It provides a reason to measure those processes, not permission to substitute them for achievement. Wilson's analysis likewise distinguishes several embodied-cognition propositions rather than treating bodily participation as a single causal explanation [15]. Movement Thinking is operationalized here as a teachable sequence of commitments, not a physiological theory.

Three empirical examples sharpen the uncertainty. Madden and colleagues compared Moon-phase instruction using immersive, desktop and hands-on approaches without finding significant overall learning differences, despite stronger preference for the immersive experience [3]. Johnson-Glenberg, Bartolomea and Kalina investigated embodied instruction across desktop and immersive platforms [4]. Their work supports examining instructional design explicitly rather than treating platform as destiny. Makransky, Terkildsen and Mayer's science-laboratory comparison demonstrates that higher presence and better learning need not travel together [5]. None estimates the effects of the present exoplanet lesson.

Petersen, Petkakis and Makransky separated immersion and interactivity in a factorial study [6]. Klingenberg and colleagues tested immersion and interaction against transfer to a concrete procedure and did not find significant transfer effects of either manipulation [7]. The present design changes the instructional factor: explicit evidential revision, rather than the degree of manipulation. It also changes the disciplinary problem and delay. These differences motivate a new test; they do not erase prior null findings or entitle the proposed intervention to a favorable expectation.

3. A spatial task with inspectable intellectual work

The lesson asks learners to evaluate a proposed observation schedule for an unfamiliar planetary orbit. Students must explain what an orbital model can predict, distinguish equal angular from equal temporal intervals and revise a claim after encountering contrary evidence. The consequence is disciplinary rather than theatrical: a poorly reasoned schedule misses the observations needed to distinguish competing explanations. No claim is made that the exercise is an operational telescope-planning tool.

The unit of learning work is a documented argument. A learner states a prediction, chooses a comparison, interprets the resulting positions, identifies an alternative explanation and records whether the prediction should change. This sequence is deliberately small. It does not require a large portfolio, an elaborate publication or a public identity. The substantive question is whether the learner can justify a revision when a new constraint appears after instruction has ended.

Transfer must be specified across dimensions rather than invoked as a general virtue [8]. The delayed assessment changes the host system, viewing orientation, available evidence and decision constraint. It also removes the training display. Success therefore requires more than repeating a familiar control sequence. However, this remains disciplinary transfer within orbital reasoning. The protocol does not claim to measure general intelligence, occupational readiness across all domains or the educational superiority of immersive schooling.

4. Why the archive is a stimulus source, not an outcome database

The NASA Exoplanet Archive supplies astronomically grounded parameters and literature provenance [9]. Its Planetary Systems table has a persistent dataset identifier, 10.26133/NEA12 [10]. The table's reference-specific structure permits a more coherent starting point than combining whichever values happen to fill a row most completely. The archive documentation distinguishes that structure from the composite-parameter table, whose more complete values are not necessarily mutually consistent [11].

This distinction matters educationally. A model built from incompatible parameters can reward a learner for accepting a visually plausible inconsistency. Conversely, unexplained missing values can turn an apparently authentic lesson into an invented dataset. The present audit therefore retains solution type, discovery method, default selection, controversy and system-composition flags together with orbital parameters and the original reference field. It does not reduce provenance to a NASA logo.

The archive contains no assessment scores, classroom allocation, learner characteristics or instructional exposure for this study. Its records cannot show that students learned, nor can their scientific authority validate a pedagogy. The completed database analysis establishes a supply of candidate stimuli under declared restrictions. Human learning outcomes require separately collected primary data. Keeping these evidence layers distinct is central to the article's claim of originality and to its limits.

5. Frozen retrieval and an executed feasibility audit

The synchronous Table Access Protocol endpoint was queried using explicit selected columns, a default_flag=1 restriction and ordering by planet name [12]. The complete query and encoded URL are supplied with the analysis. The response was retrieved at 2026-10-16T00:25:01.360504+00:00, corresponding to the evening of October 15 in California. Its 1,873,252 bytes have SHA-256 034f619a7366eea09fdf2a169d9c7c9ca61f926731702f29fe9ff9e61b4b56cd. A future query may return different contents; the identifier names a living table, not this exact response.

The downloaded response contains 6,451 unique default planet-name rows. Of these, 6,378 have solution type Published Confirmed, 69 have a Kepler project-candidate solution and four have a TESS project-candidate solution. A default flag alone is consequently insufficient for the stimulus rule. These are classifications in this particular table response, not a fresh estimate of the number of confirmed planets in the universe or a claim that the archive is erroneous.

Across the unfiltered response, 368 orbital periods, 2,496 semi-major axes, 3,678 eccentricities and 4,348 inclinations are missing or nonfinite under the parser's declared check. These counts overlap and must not be added. Missing eccentricity remains missing; it is never replaced by zero to create a drawable ellipse. Inclination is audited but not required, because the lesson's viewing rotation is explicitly stipulated rather than presented as a reconstructed sky orientation.

6. Selection rules protect the meaning of a rendered orbit

The sequential restrictions retain Published Confirmed solutions, a zero controversy flag, one stellar component with no circumbinary flag, and discovery methods Transit, Radial Velocity or Transit Timing Variations. The method restriction avoids using imaging or microlensing projected separation as though it were a measured orbital semi-major axis, a documented semantic hazard [11]. It is a conservative task-design rule, not a judgment that excluded discoveries are scientifically inferior.

The remaining checks require positive finite period and semi-major axis, finite eccentricity between zero inclusive and 0.95 exclusive, zero limit flags for all three quantities and a nonempty parameter-reference field. The sequential retained counts are 6,378; 6,320; 5,761; 5,386; 3,026; 1,916; 1,692; and 1,692. These yield 1,692 core candidates, including 375 with an exact-zero eccentricity entry. Such an entry is not certified here as an independently measured perfectly circular orbit; a selected teaching case must inspect whether circularity was fixed during fitting.

A stricter uncertainty-challenge subset requires finite, strictly positive upper and strictly negative lower errors for all three parameters, a positive lower period and axis bound, and eccentricity bounds within the elliptic domain. It contains 1,220 candidates; 472 core candidates fail that additional rule. This establishes ample candidate supply without declaring every surviving record classroom-ready. An astronomer and accessibility specialist must still approve the frozen task set before recruitment. No such approval is claimed as completed.

7. Physics, visual conventions and uncertainty are separate layers

The instructional kernel solves the two-body Kepler equation $M=E-e\sin(E)$, then calculates $x=a(\cos(E)-e)$ and $y=a\sqrt{(1-e^2)}\sin(E)$. Distances are in astronomical units; period is in days; eccentricity is dimensionless. Mean anomaly advances by a stipulated fraction of the period. The executable geometry function uses a bounded bisection solver, with tests for a circular orbit, periapsis, apoapsis and invalid eccentricities. This is inspectable pedagogical geometry, not an orbit-fitting pipeline.

Time zero, viewing rotations, marker size and display scaling are task conventions. They are identified as such in every case. No actual present-day sky position, transit prediction, planetary surface, atmospheric condition or habitable status is inferred. Planet-marker radii are diagrammatic. A persistent legend explains that spatial distances and marker sizes use different scales. The lesson excludes multi-body interactions, long-term stability, light-travel effects and other complexities inappropriate to the target inference.

Reported marginal error bounds do not supply a covariance matrix or posterior distribution. The uncertainty activity therefore compares declared bounded scenarios without calling them a 95% confidence region or randomly sampling parameters as independent truths. Learners are rewarded for noticing this limitation. The candidate case bank is split by host star across training, immediate and delayed forms, preventing another planet in a familiar system from quietly serving as a supposedly unfamiliar assessment case.

8. Rights, acknowledgment and bounded redistribution

Public access is not treated as an article-wide Creative Commons license. The archive's acknowledgment and DOI instructions were checked, as was its institutional disclaimer [13]. This package redistributes original code, original instructional specifications and aggregate audit results. It does not redistribute the retrieved raw CSV, source papers, archive images, institutional marks or third-party figures. The query and response hash support transparent provenance without inventing a blanket grant of rights.

The work acknowledges the NASA Exoplanet Archive, operated by the California Institute of Technology for NASA's Exoplanet Exploration Program. Its contributing researchers remain credited through each selected parameter reference. The future classroom renderer must attach those references to its frozen case manifest and confirm any additional material-specific permissions. An archive citation does not replace attribution to the literature supplying a particular solution [13].

These boundaries also prevent an educational authorship error. Dr. Freedom Cheteni is the author of this proposed instructional and evaluation framework, not the discoverer of the planetary systems or author of their archived measurements. NASA and Caltech are data-service providers, not collaborating authors or endorsers of the protocol. The publication does not transfer astronomy papers into the journal's original research collection or place their authorship under a school leader's name.

9. The prospective factorial design

The proposed recruitment target is 32 schools, one science instructor per school, four eligible course sections per instructor and 24 consented baseline-cohort students per section: 128 sections and 3,072 learners. These are planning values, not enrollments. Schools must offer a comparable orbital-reasoning unit in grades 9-11. A pilot outside the main cohort will examine curriculum fit, accessibility and scoring before final registration. Its results may justify a transparently dated design amendment before allocation.

Consent, roster closure and baseline assessment precede randomization. An independent allocation custodian assigns the four sections within each instructor block to the four cells, using a recorded seed after roster lock. Each instructor teaches exactly one section in each condition. The public utility verifies unique sections, one instructor per school and four sections per block. A recruitment contingency does not permit silently replacing an incomplete block after outcomes are known.

Having one instructor block per school avoids assuming that several instructors inside the same school provide independent blocks. Shared instructor and school conditions are held

within each contrast. It does not eliminate interference between that instructor's sections. The study therefore documents teaching order, shared materials and cross-section discussion, while treating the no-material-spillover assumption as a genuine vulnerability. Students retain access to normal educational support throughout.

10. Four conditions with equivalent disciplinary opportunity

Factor I is the offer of an immersive stereoscopic presentation with a seated, head-tracked view, versus a desktop presentation of the same orbital kernel. Both permit the same variable changes, time steps, pauses, annotations and evidence comparisons. Both have the same narration, numerical precision, task cases and instructional duration. The display contrast inevitably includes some interface differences; it is an offer-of-mode estimand, not a pure manipulation of a disembodied quantity called immersion.

Factor R is explicit prediction-evidence-revision instruction versus a time-matched guided comparison. In R=1, a learner records a prediction before each manipulation, examines a discrepant observation and states a reason for retaining or changing the claim. In R=0, learners examine the same examples, receive the same scientific explanations and complete matched classification and summary activities. They are not denied ordinary feedback. The contrast concerns the structured evidential commitment-and-revision sequence rather than time, basic content or exposure to correct answers.

Three 45-minute lessons occur within one instructional week. A common familiarization session precedes them. Prerecorded explanations and instructor prompts limit differential coaching. Sessions are scheduled as close together as practical, teaching order is randomized within logistical constraints, and no condition receives the other condition's worksheets before the delayed endpoint. Such precautions reduce contamination but cannot certify its absence.

11. Accessibility defines the intervention being offered

Every condition includes an equivalent non-headset route, keyboard or switch-compatible control where needed, captions, adjustable text, non-color-only cues and the ability to pause. The W3C XR Accessibility User Requirements note informs this design but is not presented as a certification or mandatory technical standard [14]. The pilot must include people with relevant access needs, rather than inferring usability from a functioning device or an adult researcher's inspection.

Learners may decline or stop headset use without penalty and continue with the matched desktop route. A person who elects that route remains in the assigned immersive-offer

group for the intention-to-treat analysis. The target is consequently the effect of offering the accessible instructional package in school practice. It is not a headset-complier effect. A secondary exposure summary records actual delivery without conditioning the principal causal estimate on a post-allocation choice.

No rapid artificial locomotion, forced standing or uninterrupted exposure is required. Brief symptom checks precede and follow each lesson, with breaks and immediate stopping available. The draft symptom questions are local monitoring items, not a validated diagnostic instrument. Device errors, adverse experiences and access failures are recorded by condition. A lesson that produces excellent scores for a narrow group while excluding classmates would fail the district's implementation criterion.

12. Primary endpoint and original scoring instrument

The primary endpoint is the total score on a 35-minute assessment administered 21 calendar days after the final lesson, with an allowed window of plus or minus three school days. No headset is used. Learners receive a new host system, a changed viewing representation and an observation constraint not practiced during training. A standardized evidence update midway through the task requires a justified revision. The same task information is available in accessible visual, textual and supported response forms.

The original rubric has five dimensions, each scored zero through four: reference-frame coordination; distinction between orbital geometry and timing; evidential revision; evaluation of a rival explanation; and recognition of model or uncertainty limits. Total scores range from zero through 20. The instruments document supplies task prompts, anchors and rater instructions. This is a proposed measure requiring piloting and validity investigation. It is not labeled a validated spatial-intelligence test.

Two trained human raters independently score every de-identified response while masked to school, instructor, condition and display-use history. The prespecified endpoint is their mean total score. Raters may see the learner's initial and revised reasoning because the construct requires that comparison. A third reader investigates disagreement for calibration and reporting, without selectively replacing inconvenient primary scores. Inter-rater agreement, score distributions and masking failures are reported rather than treated as cosmetic quality assurances.

13. Measurement schedule and evidential independence

Baseline assessment measures prerequisite orbital knowledge and a short transfer form before allocation. Immediate assessment after lesson three uses a different case to measure short-term explanation. The delayed form is reserved for the primary endpoint. Form order is counterbalanced through predetermined student-level sequences within each section, and training and assessment host systems do not overlap. Pilot checks examine whether nominally parallel forms introduce avoidable difficulty differences.

Presence, perceived effort, actual time on task, prior headset use, access route and symptoms are secondary process measures. They do not form part of the transfer score. The instructional prediction-revision records describe implementation and possible mechanisms but are not rated as the independent primary outcome. This separation prevents the circular conclusion that a revision programme worked because students filled in its revision form.

Assessment administration is scripted and available to every baseline-cohort learner, including those who missed a lesson or stopped headset use. Attendance, reasons for missingness and opportunity for an accessible make-up session are recorded before outcome scoring. Teachers do not use the experimental outcome to determine course grades. Research participation, ORCID registration and public publication are separate decisions; none is a condition for receiving the lesson or educational support.

14. Estimand, clustering and the primary statistical contrast

Let μ_{bir} denote the mean delayed score for the fixed baseline cohort in school-instructor block b assigned immersion offer i and revision instruction r. The primary contrast is $\Delta_b=\mu_{b11}-\mu_{b10}-\mu_{b01}+\mu_{b00}$, averaged with equal block weights. It asks whether the effect of revision instruction differs between display offers. The direction is not presupposed. Revision and immersion marginal effects are declared secondary contrasts, each averaging its two simple effects.

With one section per factorial cell per block, the sample variance of block contrasts divided by the number of blocks supplies a conservative across-block uncertainty estimate under independent blocks and the stated no-interference conditions. The principal interval and two-sided test use an approximate t reference with $B-1$ degrees of freedom. This is not an exact small-sample guarantee. Heterogeneous block effects contribute to the estimated variability; common school and instructor intercepts cancel within each complete contrast.

The analysis specification defines section means, equal weights, identifiers and reporting conventions, following the need to make cluster allocation and analysis units explicit [16]. A learner-level secondary model may improve precision using baseline performance and form sequence, with the nesting preserved. It does not replace the prespecified block contrast after results are seen. Permuting observed outcomes does not give an exact test of zero interaction when unknown main effects remain; no such claim is made by the supplied tools.

15. Missingness, multiplicity and departures from the plan

The intention-to-treat population is the fixed baseline cohort. The executable complete-cohort estimator deliberately refuses missing scores or incomplete blocks; it is a verification tool, not a covert listwise-deletion policy. The planned full analysis uses multilevel multiple imputation for missing delayed outcomes, with 50 imputations and predictors including baseline score, form, allocation cell, attendance, immediate assessment, access route and reported symptoms. School-block and section effects preserve clustering. Bounded predictive draws keep scores within zero through 20.

The supplementary specification declares the imputation model, convergence checks, Barnard-Rubin small-sample pooling [20] and missing-not-at-random sensitivity shifts. Those procedures are specified rather than executed because there are no participant outcomes to impute. If an entire section is missing, reporting must show the resulting dependence on extrapolation. Worst-case bounded outcomes accompany model-based conclusions. A visibly precise interval cannot make informative dropout ignorable.

There is one confirmatory interaction endpoint. The two marginal effects form a secondary family with Holm adjustment; immediate transfer and process variables are reported with estimates and clearly exploratory intervals. Associations between revision traces and delayed scores are not identified causal mediation effects. Any changes to recruitment, scoring, delivery or the endpoint window require versioned reasons, dates and visibility into whether outcome information was available when the decision occurred.

16. Completed design sensitivity: interactions are expensive

The executed simulation uses NumPy 2.3.5, seed 20261016 and 10,000 artificial replicates for each of nine scenarios. It varies instructor-school blocks across 16, 32 and 48 and section-level intracluster correlation across 0.05, 0.10 and 0.15, with 24 learners per section. The working model has unit marginal individual variance, independent section residuals within a block, complete outcomes and a constant interaction of 0.40 standard deviations. Common block effects cancel by construction. Cluster-design sensitivity is a planning requirement rather than a reassurance to be supplied after recruitment [17].

For the proposed 32 blocks, simulated interaction power is 0.9538, 0.8457 and 0.7187 as the assumed correlation increases. The corresponding Monte Carlo standard errors are 0.00210, 0.00361 and 0.00450. Null rejection proportions are 0.0469, 0.0500 and 0.0462. These near-nominal artificial-data results check implementation under the simulated model; they do not establish operating characteristics under contamination, attrition or skewed bounded scores.

At 48 blocks, power under the highest assumed correlation is 0.8854. The 32-block normal-approximation interaction thresholds for 80% power are approximately 0.296, 0.367 and 0.427 standard deviations. A 0.40-standard-deviation interaction is a scenario, not a predicted benefit or validated importance threshold. The Gaussian calculations omit floor and ceiling effects. Pilot information and resources must determine whether a larger prospectively fixed design is necessary before main-study allocation.

17. What would count against the account

The strongest theoretical result is not necessarily a positive interaction. An accurately estimated near-zero interaction with a beneficial marginal revision effect would suggest that the evidential method travels across media. A negative interaction could indicate that the method competes with immersive interface demands. An immersive presence advantage accompanied by no delayed-transfer benefit would challenge the procurement rationale while remaining consistent with prior distinctions between experience and learning [2] [3] [7].

The mechanism would be weakened if students produced more elaborate revision records without better independent reasoning, or if gains appeared only on a familiar training representation. A benefit restricted to experienced headset users would prompt an access and instructional-design investigation, not a declaration that novices lack aptitude. Excess symptoms, loss of instructional time and exclusion from the accessible task are implementation harms even when a mean score favors a condition.

No equivalence claim follows from a nonsignificant test. Before recruitment, investigators and affected educators must choose and justify any smallest educationally important difference in rubric points. Cost per delivered accessible lesson and time needed for teacher preparation are reported alongside learning. The public decision is whether a defensible benefit justifies the whole implementation burden, rather than whether a device can produce one attractive statistic.

18. Ethics, governance and a study that has not begun

This protocol is prepared for prospective registration and independent ethics review. No registration number, ethics approval, participating school, learner consent, recruitment total or trial outcome is asserted. The structure is informed by SPIRIT's emphasis on transparent prospective specification while adapting it to educational research rather than claiming biomedical protocol compliance [18]. Schools and families must be able to distinguish the existence of a publication from authorization to conduct a study.

Future primary records require pseudonymous identifiers, restricted access, a retention schedule and separate storage of consent and outcome data. Public sharing should favor code, scoring guidance and appropriately screened aggregate results. Student handwriting, voice, room imagery and behavioral telemetry are not routine public supplements. A data-access committee must evaluate re-identification risk before any individual-level release, with the authority to refuse publication of identifying artifacts.

The second author and Principal Investigator is also Editor in Chief. A future acceptance decision for this new protocol and any results paper requires documented independent editorial handling and human expertise. The companion conceptual article's earlier review does not constitute review of this new design. This edition records its actual state without inventing an acceptance date, DOI assignment, institutional collaboration or reviewer endorsement.

19. Limitations and the distinct contribution

Database completeness favors scientifically well-characterized systems and selected discovery methods; the stimulus candidates are not representative of all exoplanets. The strict uncertainty rule may omit useful cases, while published marginal errors omit covariance and other sources of uncertainty. The audit identifies possibilities, not a finished validated curriculum. A selected model may remain unsuitable for novice learning despite passing every numerical check.

The primary study would evaluate one topic, age range, instructional sequence and accessible delivery package. Instructor spillover, school recruitment, technology familiarity, assessment validity and missing outcomes can alter interpretation. The four-section requirement may select unusually large courses or schools. The simulation supplies no empirical estimate of the study's actual correlation or likely effect. Generalization to other sciences, adult employment or an entire university would require additional studies.

The contribution is narrower and more durable: it connects a philosophical learning proposition to a falsifiable comparison, gives an authentic database a precisely limited role, and exposes assumptions before human outcomes exist. A district can inspect why a case is eligible, how a lesson changes a learner's intellectual obligations, which score will matter and what evidence would justify changing course. That is an infrastructure for primary research rather than a claim to have completed a field experiment.

Method and data statement

Completed work consists of the frozen NASA query-response audit, nine-scenario artificial-data sensitivity analysis, geometry checks, estimator checks and preparation of the original protocol and instruments. The supplementary code can fetch a fresh public response, audit a supplied local response, generate the design scenarios, allocate a supplied roster and analyze a complete synthetic or study cohort under declared restrictions. The original retrieved CSV is not included. Its query, retrieval timestamp and hash are retained, so a later response must not be misrepresented as the same snapshot.

There are zero actual trial participants in this package. The multiple-imputation model, optional adjusted models and curriculum validation remain prospective specifications. The archive provides astronomical task parameters only. Full-text verification, abstract or metadata verification, official-document inspection and actual file analysis are distinguished in the source audit. A reference's inclusion does not assert that its original full text, images or data have been relicensed for republication.

AI, authorship and editorial disclosure

ChatGPT assisted research planning, source discovery, drafting and code preparation [19]. The supplied programs executed the stated audit and artificial-data calculations; the assistant did not generate participant observations or serve as a human peer reviewer. The exact source-model build is not asserted. The paper agent retrieves passages from this manuscript and must not describe proposed results as observed effects or extend its sources beyond their documented evidential role.

Agathe Bertin is the first author, with affiliations at Hasso Plattner Institute of Design and Santa Clara University. Dr. Freedom Cheteni is the second author and Principal Investigator of this original framework and protocol. Final author approval, independent review, study authorization and any later acceptance are distinct records. NASA, Caltech, the cited researchers and the conceptual article's reviewer are not made coauthors or endorsers by citation. Original astronomers retain credit for source measurements, and any future results article must identify actual investigators, contributions, deviations and permissions before publication.

Research materials and reproducibility

Inspect the prospective study specification, original draft instruments, executed public-file audit, design simulations and documented tools. No local participant trial outcomes are reported.

Complete research package · ZIP

Research methods guide

Study specification

Task instruments and scoring

Statistical analysis specification

Executed NASA stimulus feasibility audit

Design simulations and uncertainty

Source and rights audit

Executable research tools

Methods tests

Read the corresponding conceptual foundation

Evidence and limits

Claim	Evidence	Limit
Authentic astronomical stimuli can be selected under explicit semantic restrictions.	Executed 6,451-row default-solution audit; 1,692 core and 1,220 strict-uncertainty candidates. [9, 10, 11, 12]	Numerical eligibility is not astronomy/curriculum approval or learning evidence.
Presence and delayed transfer require distinct measures.	Prior empirical platform and transfer studies motivate independent assessment. [2, 3, 5, 7]	Prior studies do not establish effectiveness of the proposed lesson.
A display-by-method interaction is a distinct primary research target.	Blocked 2×2 allocation, equal disciplinary opportunity, delayed blinded score and explicit contrast. [6, 16, 18]	The study is prospective and relies on limited cross-section interference.
Powering an interaction is sensitive to the number of independent school blocks.	Nine executed 10,000-replicate artificial-data scenarios with Monte Carlo uncertainty. [17]	No attrition, floor/ceiling, treatment heterogeneity or spillover is modeled.

References and rights

[1] Freedom Cheteni. Movement Thinking: The Human World Model. The VR School · Media. 2026-08-03.

https://www.thevrschool.org/media/world-models-and-learning

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[2] Guido Makransky; Gustav B. Petersen. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. Educational Psychology Review, 33, 937-958. 2021-01-06.

https://doi.org/10.1007/s10648-020-09586-2

CC BY 4.0; citation only in this work. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[3] J. Madden; S. Pandita; J. P. Schuldt; B. Kim; A. S. Won; N. G. Holmes. Ready student one: Exploring the predictors of student learning in virtual reality. PLOS ONE, 15(3), e0229788. 2020-03-25.

https://doi.org/10.1371/journal.pone.0229788

CC BY 4.0; citation only in this work. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[4] Mina C. Johnson-Glenberg; Hannah Bartolomea; Elena Kalina. Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences. Journal of Computer Assisted Learning, 37(5), 1263-1284. 2021-06-01 (online).

https://doi.org/10.1111/jcal.12567

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[5] Guido Makransky; Thomas Schjødt Terkildsen; Richard E. Mayer. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction, 60, 225-236. 2019-04-01.

https://doi.org/10.1016/j.learninstruc.2017.12.007

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[6] Gustav Bøg Lassen Petersen; Giorgos Petkakis; Guido Makransky. A study of how immersion and interactivity drive VR learning. Computers & Education, 179, 104429. 2022-04.

https://doi.org/10.1016/j.compedu.2021.104429

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[7] Sara Klingenberg; Robin Bosse; Richard E. Mayer; Guido Makransky. Does Embodiment in Virtual Reality Boost Learning Transfer? Testing an Immersion-Interactivity Framework. Educational Psychology Review, 36, 116. 2024-10-04.

https://doi.org/10.1007/s10648-024-09956-0

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[8] Susan M. Barnett; Stephen J. Ceci. When and where do we apply what we learn? A taxonomy for far transfer. Psychological Bulletin, 128(4), 612-637. 2002-07.

https://doi.org/10.1037/0033-2909.128.4.612

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[9] Jessie L. Christiansen; Douglas L. McElroy; Marcy Harbut; David R. Ciardi; Megan Crane; John Good; Kevin K. Hardegree-Ullman; Aurora Y. Kesseli; Michael B. Lund; Meca Lynn; Ananda Muthiar; Ricky Nilsson; Toba Oluyide; Michael Papin; Amalia Rivera; Melanie Swain; Nicholas D. Susemiehl; Raymond Tam; Julian van Eyken; Charles Beichman. The NASA Exoplanet Archive and Exoplanet Follow-up Observing Program: Data, Tools, and Usage. The Planetary Science Journal, 6, 186. 2025-08.

https://doi.org/10.3847/PSJ/ade3c2

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[10] NASA Exoplanet Science Institute. NASA Exoplanet Archive: Planetary Systems Table. Caltech/IPAC, NASA Exoplanet Archive. Living dataset; snapshot retrieved 2026-10-16.

https://doi.org/10.26133/NEA12

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[11] NASA Exoplanet Archive. Planetary Systems and Planetary Systems Composite Parameters Table Definitions. Caltech/IPAC, NASA Exoplanet Archive. Accessed 2026-10-16.

https://exoplanetarchive.ipac.caltech.edu/docs/API_PS_columns.html

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[12] NASA Exoplanet Archive. Retrieving Exoplanet Archive Data with Table Access Protocol. Caltech/IPAC, NASA Exoplanet Archive. Accessed 2026-10-16.

https://exoplanetarchive.ipac.caltech.edu/docs/TAP/usingTAP.html

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[13] NASA Exoplanet Archive. Acknowledging the NASA Exoplanet Archive in Publications; Exoplanet Archive Privacy Policy and Disclaimers. Caltech/IPAC, NASA Exoplanet Archive. Acknowledgment page updated 2025-08-13; accessed 2026-10-16.

https://exoplanetarchive.ipac.caltech.edu/docs/acknowledge.html

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[14] Joshue O’Connor; Janina Sajka; Jason White; Scott Hollier; Michael Cooper (editors). XR Accessibility User Requirements. World Wide Web Consortium, Accessible Platform Architectures Working Group Note. 2021-08-25.

https://www.w3.org/TR/2021/NOTE-xaur-20210825/

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[15] Margaret Wilson. Six views of embodied cognition. Psychonomic Bulletin & Review, 9, 625-636. 2002-12.

https://doi.org/10.3758/BF03196322

Citation only; no article-wide reuse license asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[16] Marion K. Campbell; Gilda Piaggio; Diana R. Elbourne; Douglas G. Altman; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ, 345, e5661. 2012-09-04.

https://doi.org/10.1136/bmj.e5661

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-16

[17] Clare Rutterford; Andrew Copas; Sandra Eldridge. Methods for sample size determination in cluster randomized trials. International Journal of Epidemiology, 44(3), 1051-1067. 2015.

https://doi.org/10.1093/ije/dyv113

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-16

[18] An-Wen Chan; Isabelle Boutron; Sally Hopewell; David Moher; Kenneth F Schulz; Gary S Collins; Ruth Tunn; Rakesh Aggarwal; Michael Berkwits; Jesse A Berlin; Nita Bhandari; Nancy J Butcher; Marion K Campbell; Runcie C W Chidebe; Diana R Elbourne; Andrew J Farmer; Dean A Fergusson; Robert M Golub; Steven N Goodman; Tammy C Hoffmann; John P A Ioannidis; Brennan C Kahan; Rachel L Knowles; Sarah E Lamb; Steff Lewis; Elizabeth Loder; Martin Offringa; Philippe Ravaud; Dawn P Richards; Frank W Rockhold; David L Schriger; Nandi L Siegfried; Sophie Staniszewska; Rod S Taylor; Lehana Thabane; David J Torgerson; Sunita Vohra; Ian R White; Asbjørn Hróbjartsson. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ, 389, e081477. 2025-04-28.

https://doi.org/10.1136/bmj-2024-081477

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-16

[19] OpenAI. ChatGPT [AI-assisted research planning, drafting and code preparation software]. Software/tool disclosure; not a human author or peer reviewer. Accessed 2026-10-15-16.

https://chatgpt.com/

Tool use disclosed; model version not asserted. Original authors and institutions retain their rights; source text, figures and raw data are not republished.

Authoritative source / evidence | checked 2026-10-16

[20] John Barnard; Donald B. Rubin. Small-sample degrees of freedom with multiple imputation. Biometrika, 86(4), 948-955. 1999.

https://doi.org/10.1093/biomet/86.4.948

Citation only; no third-party text or instrument republished. Original authors and institutions retain their rights. Data and instrument use must follow the source terms.

Authoritative source / evidence | checked 2026-10-16

An original educational primary-study protocol and methods analysis. Astronomical parameters retain their original source attribution. NASA/Caltech provide a public data service and are not collaborators or endorsers. Source articles, images and the raw TAP response are not republished. New work licensing and final human author approval are separate records.

A10 | PRIMARY-STUDY PROTOCOL AND REPRODUCIBLE METHODS RESEARCH

Who May Change the Learning Place? A Community-Governed Trial of Cultural Authority, School Belonging and Disciplinary Transfer

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-16 | Edition 1.0 | 4,635 article words

Primary-study protocol and reproducible methods; separate human editorial record.

Keywords: cultural authority, school belonging, community governance, disciplinary transfer, NCES Common Core of Data, cluster randomized trial, Movement Thinking

Read online / article companion / citation exports

Contribution

A community-governed assignment contrast with genuine refusal/private routes, longitudinal belonging and blinded transfer, supported by an executed CCD feasibility audit and retention/spillover design sensitivity.

Prospective paired-school primary trial; completed administrative source audit and artificial-data design calculations; no trial observations or causal effects reported.

Abstract

Belonging is often pursued through visible cultural representation while decisions about what a representation means remain institutionally closed. This article specifies an original primary-study protocol testing whether consequential community authority over learning designs improves school belonging and disciplinary transfer beyond an equally resourced curriculum programme. The proposed study recruits 64 schools into 32 matched pairs, with 40 students per baseline cohort. Schools receive either a community-governed design process or an active comparison with equivalent planning time, resources, accessibility support and facilitator contact. The intervention distributes bounded rights to challenge representations, require reasoned responses and refuse disclosure or reuse; it does not prescribe an authentic culture or require public personal narratives. The primary endpoint is 24-week change in a permission-cleared school-membership measure. Blinded disciplinary-transfer tasks and a delayed assessment test whether relational gains accompany learning rather than ceremonial participation. Completed methodological investigations audit the NCES Common Core of Data school-directory and membership files and simulate paired-school designs under heterogeneous retention and comparison exposure. The directory audit identifies 102,274 unique schools and 29,115 candidates under declared grade and status restrictions. These historical administrative records support sampling feasibility, not cultural identification or proof of belonging. The article supplies governance instruments, analysis specifications, disclosure safeguards and explicit limits on measurement and causal inference. No participants have been enrolled and no intervention effect is reported.

1. From cultural visibility to a contestable learning place

A student can recognize a street, language or family practice in a lesson and still have no standing to correct its interpretation. Conversely, an unfamiliar setting can become educationally accessible when the task explains its assumptions and permits a reasoned challenge. The leadership problem is therefore not exhausted by counting culturally recognizable images. It concerns whether people affected by a representation can question it, obtain a substantive response and influence its future use.

The companion conceptual article, Belonging by Design, distinguishes representational access, interpretive authority and institutional responsiveness. The present protocol makes that distinction experimentally inspectable. It develops Cheteni’s concern with learning connected to purpose and place [1], while recognizing the intellectual priority of culturally relevant pedagogy rather than renaming its established concerns [2]. The

proposed contribution is a bounded governance contrast and a prospective evidentiary design, not a claim to have invented cultural relevance.

Brown and colleagues' culturally relevant virtual-reality science study motivates attention to the relationship between a learning setting and students' experiences [3]. Its scope does not establish that a participatory governance package causes belonging or transfer across schools. That gap matters. A polished environment might improve enthusiasm while leaving knowledge, decision rights and institutional responsiveness unchanged. The study asks whether a different distribution of authority changes outcomes that cannot be satisfied merely by producing an attractive scene.

2. A mechanism that can fail

The intervention theory has three distinct links. First, accessible representations and meaningful choices may reduce the effort required to enter a disciplinary problem. Second, a credible right to question how people or places are portrayed may alter whether students experience their interpretations as consequential. Third, documented institutional responses may make participation trustworthy over time. These are hypothesized mechanisms, not interchangeable indicators of success.

The principal causal contrast is between two funded approaches to the same curricular objectives. Both include competent teaching, technical support and opportunities to ask questions. The treatment additionally grants defined decision rights over representations and their reuse. If it improves survey ratings but not the independently scored transfer tasks, the claim that cultural authority strengthens disciplinary reasoning remains unsupported. If participants report more meetings and less influence, the governance mechanism itself has failed.

The protocol also admits an adverse pathway. Consultation can concentrate influence among confident adults, expose students to unwanted identity expectations, or create an obligation to perform cultural authenticity. Community authority is not automatically liberating. Its implementation must therefore be judged partly by whether individuals retain usable refusal and private pathways, including against preferences expressed by a school-appointed panel. An average gain cannot justify foreseeable coercion or the unauthorized circulation of cultural knowledge.

3. Community governance before allocation

Before recruitment, each participating school develops a governance agreement with affected students, families, educators and relevant community rights-holders. The agreement distinguishes individual consent, school curriculum responsibility and collective authority over particular knowledge or representations. A parent council, a principal or one invited contributor cannot authorize everyone else's stories. Where Indigenous knowledge or cultural material is involved, the relevant rights-holders determine who can give permission; researchers do not infer this authority from enrollment categories.

The CARE Principles locate collective benefit, authority to control, responsibility and ethics within Indigenous data governance [5]. They inform the questions that agreements must answer, without granting this study a generic license to collect Indigenous knowledge or claiming that one community's arrangements transfer to another. Participatory research also requires attention to how roles, resources and disagreements shape the production of knowledge [6]. Recruitment cannot begin on the assumption that an invitation establishes an equitable partnership.

An independent safeguarding and ethics determination is required before the proposed trial starts. No approval number, sponsor commitment or community endorsement has been obtained in this package. The public protocol records that absence plainly. The actual decision charters, compensation arrangements and material-specific permissions must be settled locally and logged before assignment. Schools unable to establish these conditions do not enter the trial simply because the administrative database contains their identifier.

4. Questions, estimands and boundaries

The primary question is whether an offer of community-governed learning design improves school belonging over 24 weeks relative to an equally resourced comparison. Let each school's outcome be the mean change in a fixed baseline student cohort. Within each pair, the assigned-treatment school's mean change is contrasted with its assigned-comparison school's change. The primary estimand averages these differences equally across participating pairs, thereby weighting schools equally rather than allowing larger cohorts to dominate.

The proposed primary measure is Goodenow's Psychological Sense of School Membership instrument [4], contingent on permission clearance and a separate measurement pilot. It is not a measure of cultural authenticity, agreement with school policy or approval of the

journal. The instrument’s established history does not guarantee that its wording, translations or longitudinal comparisons function equivalently in the proposed settings. The trial specification must be amended before allocation if the pilot identifies a materially different primary instrument or scoring rule.

Secondary questions concern blinded disciplinary transfer, durability at week 36, agency in contesting representations and the human burden of participation. Qualitative inquiry asks whose challenges receive responses and whose contributions remain absent. These analyses explain and test boundaries; they do not turn a package-level assignment into randomization of belonging, identity or community authority itself. The study does not estimate future earnings, lifelong inclusion or the effects of technology adoption in general.

5. Sampling from an administrative frame

The NCES Common Core of Data supplies school and local education agency identifiers, administrative status, grades offered and long-form membership cells [7] [8] [9]. These are useful starting points for reproducible recruitment, because they make the frame and exclusions inspectable. They are not a roster of cultural communities. Race and ethnicity classifications used for federal reporting must never become a classifier of the culture a student supposedly possesses or the material that student should be asked to contribute.

The candidate design is 64 regular public schools serving both seventh and eighth grades, organized into 32 pairs. Each school contributes a baseline cohort of 40 consenting students, with a target of 20 per grade when the eligible roster permits. These are planning quantities: 2,560 planned students, not people enrolled or observed. Updated enrollment, accessibility, staffing and governance agreements will determine actual eligibility. Any deviation in cohort composition requires a recorded amendment before assignment.

Matching is conducted within an LEA where feasible, using grade span, baseline cohort belonging, enrollment size and available contextual variables selected before outcomes are examined. A school appears in one pair only. Administrative characteristics support balance and feasibility rather than cultural matching. If recruiting two schools within an LEA would create unavoidable sharing of the entire governance intervention, the study must revise its assignment unit and recalculate power before allocation rather than silently treating a connected network as independent schools.

6. Completed directory-file feasibility audit

The executed audit downloaded the official 2023-24 version 1a school-directory ZIP, preserved its SHA-256 hash and streamed its CSV without converting identifiers to numbers. It found 102,274 rows and 102,274 distinct NCESSCH values. Each audited school identifier had 12 numeric characters and each LEAID had seven. No duplicate school identifier or blank value in the seven required fields was detected. These observations concern this exact release, not every historical or future CCD file.

The declared candidate filter retained regular schools whose updated status was open, new, added, changed agency or reopened, and whose grade-offered fields indicated both grade 7 and grade 8. It yielded 29,115 schools in 15,223 LEAs. Of these, 16,941 schools belonged to 3,049 LEAs containing at least two candidates. These counts establish the existence of a historical administrative frame large enough to investigate pairing; they do not establish willingness to participate, contemporary operation, ethical eligibility or representativeness.

The distinction between an identifier and a person is fundamental. A twelve-character code identifies a reported administrative school unit; it does not identify a stable student population or certify unchanged boundaries over time. Names, addresses and truncated codes are not substitute linkage keys. Any subsequent merge must use the declared release, unit and time period, check unmatched identifiers and recheck uniqueness after filtering. The audit exports aggregate diagnostics rather than a public list of schools suggested to have cultural deficiencies.

7. Completed membership-file audit and suppression boundaries

The companion membership ZIP was downloaded from the same official release and audited as a long file. Its grade, race/ethnicity, sex and total-indicator fields define different kinds of cells. A join on NCESSCH alone would repeatedly attach a school’s directory information and could multiply an outcome record if the analyst mistakes those cells for separate schools. The supplied script therefore audits the uniqueness of the explicitly selected Education Unit Total rather than asserting that the full long file is one row per school.

The executed membership audit counted 11,209,338 rows and 99,678 distinct school identifiers. Selecting Education Unit Total produced 99,678 rows with no duplicate school identifier and no school absent from the audited directory. Across all 11,209,338 long membership records, the audit found 66,138 blank student-count cells: 12,421 flagged Missing, 53,567 Not reported and 150 Suppressed. No negative numeric count was

observed. This is a release-specific finding; it does not erase the distinctions among missingness and suppression or validate every detailed cell's uniqueness. The selected total can include adult education and therefore does not prove that 40 students are available in grades 7-8.

Negative values, nonnumeric values and administrative data-management flags are retained in the diagnostics. They are never silently changed to zero. The release documentation must determine whether a specific code means missing, not applicable or suppressed; an observed absence of one code would not authorize treating all future missing cells as zero. Small reported cells are not invitations to infer or expose individual identities. The public package redistributes diagnostics and executable methods, not the multi-gigabyte source CSV.

Membership can help characterize whether recruitment systematically misses schools serving particular reported populations. It cannot demonstrate that a lesson respects those populations, that a community consents to representation, or that an individual has a particular culture. Individual-level study identifiers remain separate from NCES codes. No public join exposes student responses, identifiable narratives, cultural permissions or small-group outcome tables.

8. The assigned learning-design package

Treatment schools undertake four six-week cycles over 24 weeks. Each cycle begins with a consequential disciplinary question, such as evaluating competing explanations for a local environmental pattern. The task can use maps, physical models, photographs cleared for the purpose, text or an accessible digital environment. Immersive hardware is optional and matched across conditions where used. The intervention is the governance of a learning place, not the purchase of a device.

A representation record identifies the task’s disciplinary objective, source provenance, material permissions, intended audience and decisions available to students and contributors. Learners can challenge an interpretation through an accessible route. A named decision-holder must answer within a locally agreed period, record the reason and either make a change or explain the constraint. A bounded appeal route handles disagreement. Fidelity records whether these rights were usable and consequential; a count of consultation sessions alone is insufficient.

Community contributors receive compensation according to an agreed schedule, and both arms receive an equivalent budget for curriculum expertise. The treatment budget supports governance work; the comparison budget supports subject-specific curriculum

review. The intervention is therefore assessed as a choice about how a fixed resource is organized. Any extra unpaid labor, staffing substitution or additional technology is recorded, because unequal inputs would change the interpretation of the causal contrast.

9. An active comparison and allocation concealment

Comparison schools receive the same total scheduled planning time, facilitator contact, accessibility support, assessment time and material budget. Their programme strengthens ordinary curriculum design around the same disciplinary goals. It includes normal safeguarding, complaints procedures and students' existing rights. The trial does not withdraw rights or manufacture disrespect to create a stronger treatment contrast. What is withheld is the additional, specified co-design and representation-decision process, rather than a basic entitlement to safe schooling.

An allocation custodian independent of recruitment assigns one school in each locked pair to each condition after baseline measurement and cohort consent. Pair definitions, exclusion rules and the randomization procedure are time-stamped. The seed and full roster are held in an access-controlled audit record until disclosure cannot affect recruitment. A public demonstration seed may test software but must not be represented as an allocation that actually occurred. Reporting follows cluster-trial principles, including the sequence of recruitment and assignment [13].

Students and teachers cannot be blinded to an intervention that changes their decision opportunities. Outcome raters receive de-identified task artifacts without condition labels, school identifiers or recognizable local place names. Blinding success is investigated through raters' condition guesses after scoring. Facilitators keep contact logs in both arms. A strong study makes these asymmetries visible instead of describing the entire trial as blinded.

10. Private, refusal and non-extractive pathways

A student may choose a fictitious or unfamiliar learning setting, decline an interview or refuse to contribute a personal narrative while receiving the same instruction, assessment opportunity and support. No student must reveal immigration status, religious practice, household hardship, family history or community knowledge to demonstrate belonging. The protocol forbids grading the amount of personal disclosure. A private contribution is not automatically eligible for public exhibition, a paper agent, model training or later publication.

Consent records distinguish research participation, recording, quotation, public display and future reuse. Contributors can impose material-specific audience and time limits. The governance agreement states what can be withdrawn, when a public artifact cannot practicably be recalled and how later reuse requests will be handled. Researchers must not promise unlimited retrospective deletion from analyses that have already been lawfully aggregated, nor imply that agreement to one lesson licenses every subsequent use.

No public journal account or ORCID registration is a condition of participation. Research consent is separated from school enrollment and assessment as far as the design permits. Accessible participation routes, interpretation support and compensation are budgeted rather than treated as evidence that a family needs special persuasion. Refusal rates and barriers are studied in aggregate, with attention to whether the treatment creates identity pressure or disadvantages students who prefer a private route.

11. Longitudinal measurement without invented validation

Belonging is measured at baseline, week 12, week 24 and week 36. The primary endpoint is week-24 change from baseline, scored according to the permission-cleared instrument’s locked key. The public supplement identifies the proposed instrument and scoring workflow but does not reproduce its copyrighted items or invent permission to adapt them. A separate pilot checks comprehension, accessibility, translation and response burden before the confirmatory cohort is recruited.

Ordinal measurement models examine whether a common construct and loading pattern are tenable across time and major language versions, followed by appropriately specified threshold constraints. Model fit, item behavior and substantive interpretation are considered jointly; fit-index cutoffs are not a universal certificate of invariance [10]. Configural fit alone does not authorize comparing latent means. Small groups cannot support an unlimited sequence of reliable invariance tests, and unstable subgroup estimates will not be presented as definitive cultural differences.

The confirmatory primary score and minimum completion rule must be locked after the independent pilot and before assignment. If later data reveal consequential noninvariance, the prespecified score analysis remains transparently reported with that limitation. Item-level and defensible partial-invariance analyses are sensitivity work, not a search for the version that produces a favorable result. The study may conclude that the planned comparison lacks adequate measurement support.

12. Transfer assessment and evidence of revision

Students complete two parallel disciplinary tasks at baseline and week 24, with a changed-setting transfer task at week 36. A panel independent of treatment facilitation develops the forms before assignment. Each task requires the learner to identify a causal or explanatory claim, evaluate evidence, identify a missing perspective or constraint and revise a decision when the evidence changes. The stimulus does not require privileged local knowledge; each form supplies the information needed for a defensible response.

A new 0-16 rubric rates four dimensions from 0 to 4: disciplinary explanation, evidence discrimination, revision justified by reasons and transfer to a changed constraint. This is a proposed instrument, not a validated scale. The pilot must examine task difficulty, accessibility and inter-rater reliability. Two blinded human raters score all primary transfer artifacts independently; their mean is used. Disagreement is used for training and quality review rather than selectively replacing inconvenient scores after condition labels are known.

A positive belonging result does not automatically validate the transfer hypothesis. The analysis first tests the prespecified primary belonging contrast. The week-24 transfer test enters a fixed-sequence testing procedure only if the first test rejects at two-sided alpha .05; otherwise its estimate and interval are reported descriptively. Week-36 persistence, governance indicators and subgroup interactions remain secondary. This hierarchy prevents a large collection of related outcomes from becoming interchangeable routes to a declaration of success.

13. Primary analysis and clustered uncertainty

The complete-cohort analysis forms each school’s mean within-student change and each pair’s assigned-treatment minus assigned-comparison difference. The estimator is the arithmetic mean of the 32 pair differences. The standard error is their sample standard deviation divided by the square root of 32, with an approximate t interval using 31 degrees of freedom. This equal-school estimand differs from a student-weighted effect; both the weighting and the cohort definition are stated before outcomes are inspected.

A randomization analysis complements the interval by flipping assignment within each original pair. An exact test can enumerate all assignments in a small demonstration; the planned trial uses 99,999 reproducible draws with an observed-assignment correction. The Fisher test concerns a sharp null about all units’ outcomes under the allowed assignments. It is not a test that every subgroup has zero average effect, and it does not transform a small number of clusters into a nationally representative sample.

Longitudinal mixed models with pair fixed effects and school and student random intercepts provide secondary estimates across the four measurement occasions. Their treatment-by-time terms are reported with uncertainty and explicit covariance assumptions. Model-based efficiency is not substituted for the randomized contrast merely because its p-value is smaller. The executable public utility verifies pair-summary mathematics and artificial design calculations; it does not claim to implement the complete longitudinal, imputation or invariance analysis.

14. Missing outcomes, mobility and precision

The baseline cohort remains the target when students move schools or decline later questionnaires. A follow-up attempt must honor the original permission and avoid pressure. A diagram reports schools, students and item responses at every occasion by assigned condition, with reasons summarized where known. Attrition is not converted into a zero belonging score. New arrivals after allocation are described separately and cannot silently replace unavailable baseline participants.

The analysis specification proposes multilevel multiple imputation under a conditional missing-at-random model, including assignment, pair, school, baseline responses, grade, language version, prior occasion outcomes and recorded follow-up predictors. Fifty imputations are the initial minimum, increased if Monte Carlo diagnostics indicate instability. The completed-data pair estimates are pooled using within- and between-imputation variation with a small-sample degrees-of-freedom adjustment [14]. This is a proposed implementation requiring validation before study activation, not an analysis already run on participants.

Sensitivity analyses shift imputed treatment and comparison scores in unfavorable directions within the instrument’s allowable range and report a tipping-point grid. Entire-school loss receives a separate analysis because individual imputation cannot recreate an unobserved institution without stronger assumptions. The complete-case result, imputed result and sensitivity limits are shown together. Any claim of robustness must identify the missingness assumptions under which it holds rather than declaring that software has solved selection bias.

15. Spillovers, disagreement and mechanism analysis

School leadership networks and community relationships cross administrative boundaries. A contributor may support two schools, teachers may share designs and students may discuss an activity across sites. Preventing such communication would be unrealistic and sometimes ethically undesirable. Before allocation, the study maps shared design teams and likely direct material exchange without collecting an unnecessary social graph of students. During implementation, aggregate contact and material-sharing logs describe exposure across conditions.

Interference changes which causal effects can be identified [11]. The simple paired estimator has its clearest interpretation under negligible outcome-relevant spillover between schools. If the programme spreads to comparison schools, the reported contrast concerns the assignment offer within that allocation environment rather than a pure effect of receiving cultural authority. The trial does not randomize network saturation and therefore cannot separately identify direct and indirect effects just by adding a self-reported exposure covariate.

Qualitative case analysis samples both arms, high and low fidelity, refused contributions and unresolved objections. An independently maintained disagreement log preserves competing interpretations, reasons and final actions while withholding identifying details. Researchers seek disconfirming cases: a successful conventional lesson, a burdensome co-design cycle or a privately participating student who feels excluded by public celebration. Mechanism claims remain triangulated explanations, not statistically identified mediation effects inherited from the treatment randomization.

16. Completed design sensitivity with artificial observations

The executed simulation examines 24, 32 and 40 school pairs, outcome intracluster correlations of .05 and .10, and comparison-exposure fractions of zero and .25. Every scenario uses 10,000 artificial replicates and a declared seed. Each school begins with 40 students. A school-specific retention probability is drawn from Beta(6.8, 1.2), followed by a Binomial(40, p) retained count with a lower bound of 12. The achieved mean, variation and observed limits are exported for every scenario; retained cohorts never exceed the baseline 40.

Baseline and endline residual correlations are assumed to be .60 at both school and individual levels. With outcome variance scaled to one, the school-mean change variance is 0.8 times the sum of the ICC and the individual residual fraction divided by retained size. Pair differences combine the two schools’ variances. A .30

baseline-outcome-standard-deviation effect is attenuated to .225 in the .25 comparison-exposure scenario. No additional gain from matching is assumed. The test is a two-sided paired t test at alpha .05 [12].

For the 64-school candidate design, simulated power was 99.58% at ICC .05 without comparison exposure, 93.83% with .25 exposure, 95.77% at ICC .10 without exposure and 77.84% with .25 exposure. Corresponding Monte Carlo standard errors were 0.00065, 0.00241, 0.00201 and 0.00415. The achieved mean retained cohort was approximately 34 with CV approximately .153; observed sizes ranged from 12 to 40. Across these four scenarios the simulated null rejection rates ranged from 4.61% to 5.03%. These are artificial-data operating characteristics of the complete-change estimator, not observed benefits, estimates of real contamination or validation of the proposed imputation model.

The simulation is a planning calculation, not an estimate of how cultural authority will affect real learners. It assumes Gaussian school-mean errors, noninformative retention and no whole-school loss. A scalar attenuation parameter does not reproduce a real network interference process. Bounds, ceiling effects, measurement changes and unequal implementation could produce different power. The evidence supports sensitivity-aware planning, including the need to revise the design before recruitment if plausible assumptions are less favorable.

17. Falsification, burden and consequences

Three patterns would directly challenge the theory. First, increased consultation with no improvement in documented response to challenges would reject the claim that the package produced consequential authority. Second, belonging gains accompanied by poorer transfer would limit its educational rationale. Third, benefits dependent on public cultural disclosure would violate the intervention’s own non-extractive design. The analysis reports these conflicts rather than allowing an appealing average score to settle them.

The protocol collects paid and unpaid staff time, contributor compensation, accessibility expenditures and student-reported pressure. A joint governance and safeguarding group sets feasible stopping thresholds before allocation, informed by the pilot and budget. No invented empirically validated burden threshold is supplied. Unauthorized disclosure or credible retaliation triggers immediate safeguarding review and suspension of the implicated activity; it does not wait for the primary outcome analysis. A refusal to publish material cannot lower a student’s academic grade.

Superintendents receive an implementation account that distinguishes expenditure, decision rights and outcomes. A programme can be infeasible even with a positive point estimate. Conversely, an imprecise result from a small or disrupted trial is not evidence that a community's concerns are unimportant. The decision to continue must consider the confidence interval, burden, transfer, privacy and the legitimacy of the local governance process.

18. Limitations

The historical CCD frame does not supply a current national recruitment census, individual cultural identity or measures of belonging. Its membership cells involve reporting conventions and potentially unavailable values. The audited releases and selected total restriction must not be generalized to unexamined versions. Recruiting consenting schools able to undertake participatory work may produce a narrower population than the frame suggests, even when randomization is internally sound.

The proposed primary measure requires permission and context-specific testing. A change in respondents' interpretation of belonging may be part of the intervention and also a threat to longitudinal comparability. Blinding is incomplete, demand effects remain plausible, and the active comparison cannot guarantee identical enthusiasm or facilitator skill. The intervention bundles governance practices, so a positive result cannot identify which component is sufficient or establish a culturally universal programme.

The completed calculations contain artificial outcomes; the administrative audit contains no participant responses. The protocol cannot substitute documentation, code or publication for an ethically conducted trial. No observed treatment effect, accreditation status, external institutional partnership or community endorsement follows from this article. It is a prospective route to primary evidence with completed methodological work, and its most consequential propositions remain open to refutation.

Method and data statement

Source archives were obtained from official NCES addresses and hashed; the public research package includes exact URLs, verification dates, stream-audit code, aggregate diagnostics, study and analysis specifications, original instruments and artificial design results. Raw CCD archives are retained as intermediate source material outside the publication repository. No confidential student records, school recruitment roster, personal narratives or cultural materials are distributed. Public-use administrative access does not authorize collecting or publishing a participant's own work.

No participants have been enrolled under this protocol, and no registration identifier, ethics approval, acceptance date or new article DOI is asserted. The accompanying tools require NumPy and SciPy for the declared calculations and are tested with synthetic fixtures. They implement no hidden connection to an LMS or student account. Actual study activation requires the finalized measurement pilot, ethics determination, permissions, staffing, budget and prospective registration. Reporting and version control distinguish those future actions from the analyses executed here.

AI, authorship and editorial disclosure

Dr. Freedom Cheteni is the named human author; substantive author approval and independent human editorial assessment of this new protocol remain separate documented actions. The earlier conceptual article is preserved and its review record does not transfer to this new manuscript. The Research University is the publisher; no unnamed Stanford collaborator or alternate journal is presented as an author or accepting venue. Sources retain their original authors, dates, rights and identifiers.

OpenAI’s ChatGPT assisted with source discovery, drafting, code preparation and consistency checking [15]. No model version is invented. Statistical results come from the disclosed executable calculations; AI output is not a participant observation, cultural permission or peer-review decision. A paper companion may retrieve exact passages and explain the protocol’s boundaries. It does not conduct the proposed trial, infer a student’s culture or authorize reuse of a contribution. Human researchers, affected rights-holders and the responsible editorial process retain authority over those decisions.

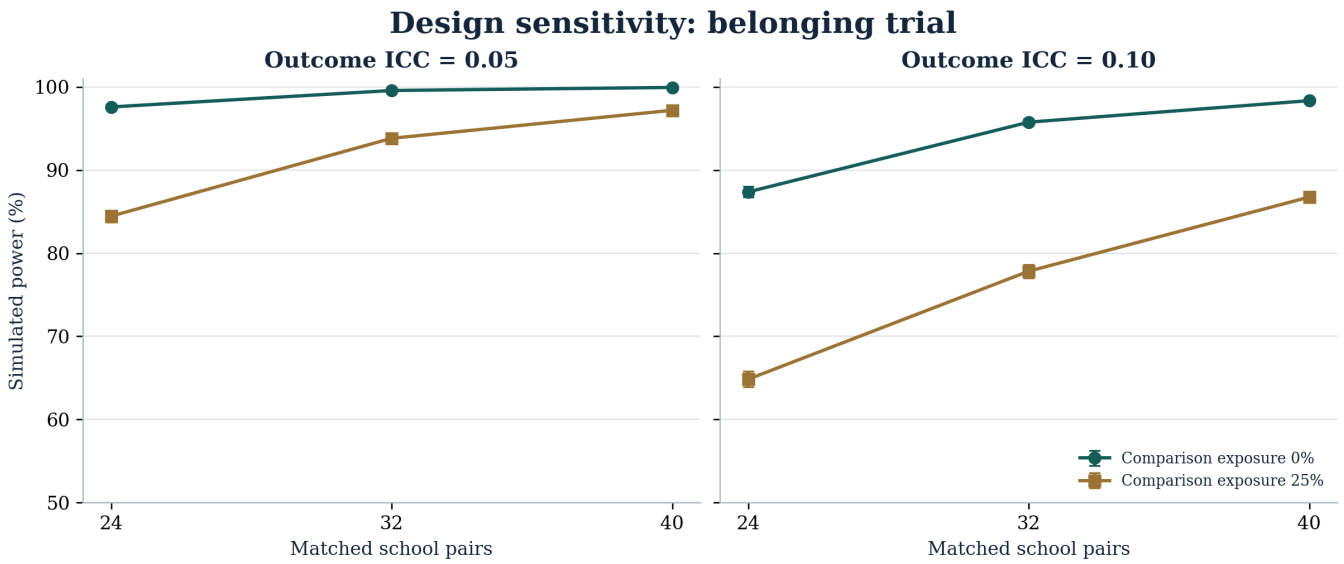


Figure 1. Artificial design sensitivity under declared retention and covariance assumptions: 10,000 replicates per scenario, 40 baseline students per school and retained cohorts 12-40. Whiskers show ±1.96 Monte Carlo standard errors, not real intervention-effect confidence intervals. Comparison exposure is scalar attenuation, not a fitted network effect. This simulation does not validate the planned imputation model.

Research materials and reproducibility

Inspect the prospective study specification, original draft instruments, executed public-file audit, design simulations and documented tools. No local participant trial outcomes are reported.

Complete research package · ZIP

Research methods guide

Prospective study specification

Instruments and governance charter

Analysis specification

Executed CCD feasibility audit

Artificial design sensitivity results

Executable research tools · Python

Source and verification audit

Methods test results

Read the corresponding conceptual foundation

Evidence and limits

Claim	Evidence	Limit
A historical administrative source can support an inspectable recruitment frame.	Executed CCD directory audit: 102,274 unique school identifiers; 29,115 candidates under declared restrictions. [7, 8, 9]	Not contemporary eligibility, consent, culture or population representativeness.
Membership-cell joins and unavailable counts require explicit restrictions.	Across all 11,209,338 long membership rows, 66,138 counts are blank, including 150 flagged Suppressed. The separate Education Unit Total restriction yields 99,678 unique school totals. [7, 8]	No full long-cell duplicate audit; no zero imputation; no inference of individual culture.
Cultural authority can be studied as an assigned governance offer.	64-school/32-pair candidate design with active comparison, bounded decision rights and separate permissions. [2, 4, 5, 6, 13]	No enrolled participants; primary instrument rights/pilot and local governance/ethics still required.
Design precision depends on retention, clustering and cross-condition exposure.	12 scenarios with 10,000 artificial replicates each and explicit Monte Carlo error. [11, 12, 14]	No actual effects; scalar attenuation is not an identified network model or validation of MI.

References and rights

[1] Freedom Cheteni. Beyond Palm Drive: Bringing the Lesson Home. The VR School · Media. 2026-07-21.

https://www.thevrschool.org/media/beyond-palm-drive

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[2] Gloria Ladson-Billings. Toward a Theory of Culturally Relevant Pedagogy. American Educational Research Journal, 32(3), 465-491. 1995.

https://doi.org/10.3102/00028312032003465

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[3] Bryan A. Brown; Kathryn Ribay; Greses Pérez; Phillip A. Boda; Matthew Wilsey. A Virtual Bridge to Cultural Access: Culturally Relevant Virtual Reality and Its Impact on Science Students. International Journal of Technology in Education and Science, 4(2), 86-97. 2020.

https://files.eric.ed.gov/fulltext/EJ1255660.pdf

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[4] Carol Goodenow. The Psychological Sense of School Membership Among Adolescents: Scale Development and Educational Correlates. Psychology in the Schools, 30(1), 79-90. 1993-01.

https://doi.org/10.1002/1520-6807(199301)30:1<79::AID-PITS2310300113>3.0.CO;2-X

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[5] Stephanie Russo Carroll; Ibrahim Garba; Oscar L. Figueroa-Rodríguez; Jarita Holbrook; Raymond Lovett; Simeon Materechera; Mark Parsons; Kay Raseroka; Desi Rodriguez-Lonebear; Robyn Rowe; Rodrigo Sara; Jennifer D. Walker; Jane Anderson; Maui Hudson. The CARE Principles for Indigenous Data Governance. Data Science Journal, 19, 43. 2020.

https://datascience.codata.org/articles/10.5334/dsj-2020-043

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[6] Jarg Bergold; Stefan Thomas. Participatory Research Methods: A Methodological Approach in Motion. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 13(1), Article 30. 2012.

https://www.qualitative-research.net/index.php/fqs/article/view/1801

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[7] National Center for Education Statistics. Common Core of Data 2023-24 Public School Universe, version 1a: School Directory and Membership. U.S. Department of Education, NCES. 2024.

https://nces.ed.gov/ccd/files.asp

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[8] National Center for Education Statistics. 2023-24 School Directory and School Membership Companion Files, 2024-252. U.S. Department of Education, NCES. 2024.

https://nces.ed.gov/ccd/xls/SY_2023-24_School_Membership_Companion_2024-252.xlsx

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[9] National Center for Education Statistics. Common Core of Data: ED Public Data Frequently Asked Questions. U.S. Department of Education, NCES. Accessed 2026-10-16.

https://nces.ed.gov/ccd/quickfacts.asp

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[10] Fang Fang Chen. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. Structural Equation Modeling: A Multidisciplinary Journal, 14(3), 464-504. 2007-07-31.

https://doi.org/10.1080/10705510701301834

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[11] Michael G. Hudgens; M. Elizabeth Halloran. Toward Causal Inference With Interference. Journal of the American Statistical Association, 103(482), 832-842. 2008-06.

https://doi.org/10.1198/016214508000000292

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[12] Clare Rutterford; Andrew Copas; Sandra Eldridge. Methods for Sample Size Determination in Cluster Randomized Trials. International Journal of Epidemiology, 44(3), 1051-1067. 2015-06.

https://doi.org/10.1093/ije/dyv113

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[13] Marion K. Campbell; Gilda Piaggio; Diana R. Elbourne; Douglas G. Altman; for the CONSORT Group. CONSORT 2010 Statement: Extension to Cluster Randomised Trials. BMJ, 345, e5661. 2012-09-04.

https://doi.org/10.1136/bmj.e5661

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[14] John Barnard; Donald B. Rubin. Small-Sample Degrees of Freedom with Multiple Imputation. Biometrika, 86(4), 948-955. 1999-12.

https://doi.org/10.1093/biomet/86.4.948

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

[15] OpenAI. ChatGPT [AI-assisted drafting software]. Software/tool disclosure, not a human author or reviewer. Accessed 2026-10-16.

https://chatgpt.com/

Citation only; no source article redistribution or instrument adaptation rights asserted. Original authors and rights-holders retain their rights; source text, figures and copyrighted scale items are not republished.

Authoritative source / evidence | checked 2026-10-16

Original primary-study protocol and executed public-data methods; cited authors retain attribution. Administrative categories do not authorize cultural representation or personal disclosure. No source findings are reassigned to the named human author.

A11 | PRIMARY-STUDY PROTOCOL AND REPRODUCIBLE METHODS RESEARCH

Beyond the Published Portfolio: An Optional-Publication Trial of Learning Transfer, Provenance Mastery and Scholarly Portability

Dr. Freedom Cheteni | Human author

ORCID 0009-0006-3507-0094

Prepared 2026-10-15 | Edition 1.0 | 4,405 article words

Primary-study protocol and reproducible methods; separate human editorial record.

Keywords: publication as learning, student agency, DataCite, provenance mastery, mentor-cluster trial, scholarly portability

Read online / article companion / citation exports

Contribution

A matched-pair trial of optional supported publication against a resource-matched private authentic audience, with independent transfer and provenance outcomes; completed 300-record metadata audit and 20,000-replicate-per-scenario sensitivity analysis.

Prospective primary mentor-cluster trial; completed bounded DataCite REST metadata audit and artificial-data design analysis. No participant learning effects, authentication or full-text reuse eligibility are reported.

Abstract

Publication can give learning a consequential audience while also creating incentives to mistake visibility for understanding. This article operationalizes publication as learning infrastructure through a prospective, matched-pair mentor-cluster trial and two completed methodological analyses. The proposed trial recruits 96 mentor-course clusters, paired within educational settings, with 12 consenting learners per cluster. It compares an optional supported publication pathway with an equally resourced private authentic-audience pathway. Both conditions receive the same inquiry time, feedback opportunities, technical support and academic assessment; neither course credit nor study participation depends on public release. The primary estimand is the equal-cluster intention-to-treat effect of offering the supported pathway on change in blinded transfer performance. Provenance mastery is a gated secondary outcome; publication counts are process descriptions. A frozen DataCite REST audit examines 300 education-related DOI records, comprising bounded alphabetic samples of datasets, software and dissertations from publication years 2024-2025. Creator ORCID entries occur in 119 records, resource-version strings in 136 and version relations in 97; these fields do not authenticate identity or establish student authorship, quality, consent or learning. Nine artificial-data scenarios, each with 20,000 replicates, expose the sensitivity of candidate design power to clustering and retention. No participant effects are reported. The contribution is a falsifiable educational design and an inspectable provenance architecture that preserves a right not to publish, separates editorial decisions from assessment and evaluates learning beyond the portfolio's public appearance.

1. The problem: public outputs are not learning outcomes

A student's work may deserve an audience beyond a classroom. A community partner can challenge an unsupported recommendation, a repository can preserve a reproducible analysis and an editor can require a more careful account of limitations. Yet these possibilities do not establish that publishing causes learning. Strong students may publish more often; well-resourced families may supply additional editing; selective journals may accept work after the most difficult intellectual changes have already occurred. Counting publications can reward these advantages while obscuring the instructional process.

The corresponding conceptual article in this volume distinguishes assessment, submission, acceptance, publication and software contribution. The present article preserves that separation and asks a narrower empirical question: does offering a supported, optional scholarly pathway improve transferable reasoning beyond equally serious inquiry addressed to a private authentic audience? The comparison matters because a weak

control would confuse publication with additional adult attention, technical help or the simple experience of being taken seriously.

Movement Thinking supplies an orientation toward adaptable inquiry and accountable revision [1]. It does not supply evidence that public exposure is educationally necessary. This protocol treats a decision not to publish as a legitimate exercise of agency, rather than a failure of ambition. Its original contribution is a design in which learning, portability, editorial acceptance and public identification can vary independently, making a favorable result harder to manufacture through the platform's own activity counts.

2. Intellectual contribution and competing explanations

Wiley and Hilton distinguish practices enabled by open permissions from the broad range of activities described as open pedagogy [2]. Their conceptual clarification is useful here: a public URL, an openly licensed work and a task that requires learners to revise reusable material are different conditions. This study therefore does not call every uploaded assignment open scholarship. It specifies which activities the intervention adds, who retains decision authority and what evidence could contradict the proposed benefit.

Cho and MacArthur's investigation of revision with peer and expert reviewing provides a research precedent for treating feedback and revision as substantive instructional activities [3]. Its population, tasks and design do not establish an effect of this proposed publication pathway. We retain feedback in both arms precisely so that a difference cannot be attributed merely to one group receiving a reader. The new question concerns a supported route into an enduring scholarly record, including explicit rights, attribution, version and audience decisions.

Three rival explanations guide design choices. An attention explanation predicts comparable improvement when private work receives equally committed readers. A selection explanation predicts that publication correlates with performance because learners choose into it. A compliance explanation predicts better-looking records without better understanding. Equal support addresses the first, randomized offers address the second at the offer level, and independent transfer tasks address the third. None is eliminated by a DOI badge.

3. A provenance contract for learning work

The proposed provenance contract separates six objects: an assessed learning artifact; a student's contribution statement; a versioned submission; an editorial decision; a released publication; and a software companion that helps a reader inspect evidence. An artifact can be assessed without being submitted. A submission can be declined while the learner earns full course credit. An accepted output can remain under embargo or require removal of material that cannot be shared. A companion is a distinct software object, rather than an additional human author.

This architecture draws a limited technical analogy to the W3C provenance model, which distinguishes entities, activities and responsible agents [16]. The protocol does not claim that adopting a vocabulary makes the evidence true. Instead, it uses separate records to make contradictions inspectable. A revision event points to prior and subsequent artifacts; a contribution assertion names its responsible human; and permission to expose identifying information has its own record and withdrawal limits. Private consent documents never become public metadata.

The contract is intentionally exportable. A learner can retrieve a human-readable contribution statement and a machine-readable manifest without adopting a particular journal, learning management system or AI account. Export preserves dates, source identifiers, version relationships and rights assertions without translating a teacher's assessment into editorial acceptance. The trial tests whether learners understand these distinctions, not merely whether the platform can fill the fields for them.

4. Why DataCite is the methods database

DataCite provides a public metadata interface covering multiple research-output types. Its documented query syntax permits field restrictions, publication-year ranges, resource-type filters, pagination and DOI-name sorting [4]. This makes it suitable for an executed feasibility audit of the metadata structures that a course-linked publication system might need to exchange. It is not a database of measured student learning, a registry of authenticated student status or a complete census of educational publications.

DataCite's schema distinguishes mandatory identification fields from recommended and optional descriptive properties [6]. Optional does not mean educationally unimportant: a record may identify a resource adequately for discovery while omitting the information a learner needs to explain its version or reuse conditions. Conversely, richly populated metadata can describe an unsuitable or inaccessible resource. The audit examines supplied fields and relationships; it does not infer academic quality from their presence.

The metadata policy explicitly covers deposited metadata available through DataCite APIs under a CC0 waiver, while excluding the linked resources and retaining applicable individual rights [5]. Accordingly, the public research materials release a privacy-minimized metadata projection and aggregate results. They do not reproduce linked dissertations, datasets, code files, student portfolios or personal profiles. Metadata reuse and permission to republish a research output remain separate decisions.

5. Frozen sample and reproducible selection

The completed audit uses three queries sharing the expression subjects.subject:education AND publicationYear:[2024 TO 2025]. The resource-type-id parameter is respectively dataset, software and dissertation. Each request specifies page[size]=100, page[number]=1, sort=name, affiliation=true and publisher=true. Exact percent-encoded URLs, retrieval times, byte counts and SHA-256 digests are preserved in the frozen-query manifest. The analysis includes precisely the 100 returned records in each stratum, totaling 300 unique DOI records.

This is a bounded alphabetic stress test, selected for inspectability and heterogeneous resource structures. It is not random, representative, exhaustive or restricted to school students. Alphabetic DOI ordering can concentrate repository prefixes, and the education subject query can include broad or inconsistent subject assignments. These limitations prevent population prevalence estimates. The returned total-hit counts are recorded as API metadata but do not become denominators for the 300-record audit; no unseen pages are described as inspected.

Before summarization, the tool checks unique normalized DOI keys and the stated publication-year window. The public projection omits names, abstracts and actual ORCID values while retaining audit-bearing counts, field-presence indicators, relationship metadata and resource dates. Its own hash is published. Re-running the live query may produce different bytes because repositories update metadata. The frozen projection reproduces the reported aggregates; raw response hashes document the original source bytes retained with the research record.

6. Operational definitions and safeguards against category errors

A creator ORCID indicator requires a nameIdentifiers entry whose declared scheme is ORCID. A second indicator checks its ISO 7064 checksum and expected syntax. Neither operation observes an OAuth transaction or proves that the identifier belongs to the named creator. DataCite's creator structure supports identifier and affiliation information [7]; ORCID's authentication guidance concerns a separate account-holder authorization workflow [12]. The audit never labels a checksum-valid identifier authenticated, and never interprets creator count as a verified author count.

Resource version means a nonempty attributes.version value. Metadata version is the registry's revision-related attribute and is retained separately, not substituted for the resource version [9]. Version relations are the explicitly declared HasVersion, IsVersionOf, IsNewVersionOf and IsPreviousVersionOf relationships. Relation types and target identifiers are preserved because a generic link cannot tell whether something is a supplement, a revision, a citation or a component [8]. No graph edge is silently converted into evidence of intellectual contribution.

Rights-list presence and rights-URI presence are separate descriptive indicators. Neither is a verified reusable-full-text decision [10]. PublicationYear, an Issued date, the registry-created timestamp, registered timestamp and top-level published timestamp are also recorded separately [11]. In particular, a null published value is not proof that the underlying resource was never published. Registry registration is not an editorial acceptance date. A landing URL is not a demonstrated accessible full-text file, and a null contentUrl does not prove the file cannot be downloaded elsewhere.

7. Completed metadata audit: availability and heterogeneity

All 300 sampled DOI records have a landing-URL field and registry-created and registered timestamps. A declared Issued date appears in 284 records. The top-level published timestamp and contentUrl are absent in every sampled record. These findings identify an interface hazard: a platform that treats the wrong timestamp as publication or equates the API contentUrl with all possible access routes would misrepresent this sample. The audit did not follow the landing pages or verify the linked full text.

Creator ORCID entries occur in 119 records, comprising 239 entries; all 239 pass the implemented syntax and checksum check. The stratum counts are 19 datasets, 67 software records and 33 dissertations with at least one such entry. This is evidence about deposited metadata only. No account was authenticated, no identity match was adjudicated and no

inference was made about the creators' age, enrollment, contribution or consent to a new publication.

A resource-version string occurs in 136 records: 66 datasets, 68 software records and two dissertations. Explicit version relations occur in 97 records: zero datasets, 69 software records and 28 dissertations. The differing patterns demonstrate why a single version-completeness score would conceal different repository practices. Related identifiers appear in 138 records; 42 records point to another DOI within the bounded sample. None of the recorded DOI relations is a self-link under the implemented normalization.

Rights lists occur in 199 records and rights URIs in 173. Language is supplied in 191 and formats in 36. The results suggest that an educational export workflow needs explicit missing-value handling and a second-stage rights check, rather than assuming a discoverable DOI arrives with a complete reuse and portability package. They do not estimate global repository quality, diagnose a publisher's compliance or identify an ingestion-ready collection. The sample contains DOI records, including versions and components, rather than 300 proven distinct research projects.

8. Prospective study population and recruitment

The planning design recruits 96 mentor-course clusters in 48 matched pairs, with 12 baseline-cohort learners per cluster, totaling 1,152 learners if recruitment succeeds. These are design targets, not a recruited sample. Settings may include secondary and introductory postsecondary research courses, but pairs must share educational stage, course domain and institution or a prespecified closely comparable setting. The study does not mix minors and adults within an unmatched comparison or presume one syllabus serves every participant.

A cluster is a stable instructional group supervised by one mentor team. A mentor cannot teach both assigned conditions; shared staff create a larger interference unit requiring redesign before allocation. Eligible courses provide sustained inquiry, an accessible private-audience route and sufficient paid support in either arm. A documented roster fixes eligibility, invitations, consent and baseline assessment before allocation. Sampling cannot select only work already likely to be published or families willing to make their children publicly identifiable.

Independent ethics review, institutional agreements, an accessible consent process, funding and a timestamped registration must precede activation. For minors, the applicable permission and assent arrangements require institutional determination;

nothing in this article supplies approval. Research participation, routine instruction, optional public release and ORCID connection have separate choices. Students who decline the research retain the ordinary course and its assessment opportunities. Recruitment flow and exclusions are reported without identifying nonparticipants.

9. Pairing, allocation and the intervention contrast

Pairs are formed before assignment using course domain, educational stage, baseline cluster mean transfer score and scheduled support. An independent custodian freezes the roster and pairing rules, then assigns one cluster in each pair to the supported pathway using independently generated fair binary draws. The seed and allocation file remain access-controlled until allocation is irrevocable. The methods package specifies the procedure but contains no purported study roster or executed participant allocation.

Both arms receive a 20-week inquiry course, ten 60-minute mentor sessions, two rounds of substantive reader feedback, equivalent technical assistance and the same assessment schedule. Reader recruitment, expected response time and feedback prompts are matched. The comparator addresses a real, pre-agreed audience in a protected setting: for example, a community advisory group that receives a confidential recommendation and asks questions. Private does not mean perfunctory, and public does not mean more academically demanding by definition.

The intervention additionally offers a structured, optional pathway through contribution negotiation, a rights-and-source audit, versioned submission preparation, a release decision and referral to an independent editorial process. Comparable time in the comparator is used for private provenance explanation and audience-responsive revision. Submission is supported, never required. An editor may reject a technically competent submission. The causal contrast is therefore an offer embedded in instruction, rather than guaranteed publication, journal prestige or receipt of a DOI.

10. The primary outcome and assessment design

The primary outcome is change from baseline to week 20 on a newly specified 24-point transfer assessment, with six dimensions scored 0-4: framing an unfamiliar problem, selecting relevant evidence, recognizing uncertainty, addressing a competing account, revising under a changed constraint and justifying a decision for an affected audience. It does not award points for publication, polished design, DOI possession or use of a favored technical vocabulary. This is a proposed instrument requiring pilot development, not a validated scale.

Parallel tasks present fictional but consequential decisions outside a student's own project. A task may ask whether a community should adopt a proposed environmental-monitoring practice, supply conflicting evidence and then alter a budget or privacy constraint. Students explain the initial decision, revise it and identify what would change their mind. Task forms are counterbalanced within clusters. No real person's health, finances or personal story is needed to demonstrate reasoning, and permitted accommodations are specified before allocation.

Two masked human raters score every assessment independently. The primary value is their mean total; a third reviewer audits discrepancies without selectively replacing inconvenient scores. Pilot artifacts, excluded from the confirmatory cohort, support cognitive interviews, accessibility checks and rater calibration. Reports include agreement and disagreement by dimension, not just a favorable reliability summary. Large ceiling effects, unreliable form equivalence or scoring that reveals condition require revision before recruitment and renewed simulation, rather than retrospective optimization.

11. Provenance mastery, portability and other outcomes

The gated secondary outcome is a 16-point provenance-mastery assessment with four 0-4 dimensions: distinguish an original contribution from source material; explain the difference between resource and metadata rights; reconstruct a defensible version chain; and justify an identity or audience-disclosure choice. Learners respond to unfamiliar, privacy-preserving case records. Knowing where to click is insufficient: a correct answer must explain why a DOI is not evidence of review, or why an ORCID string alone does not authenticate its holder.

A separate technical portability audit exports and reimports a synthetic or permission-cleared artifact package. It checks eight items: stable artifact key, human contribution assertion, source-identifier preservation, resource version, typed version relation, rights assertion, access state and distinct assessment/editorial status. Each item is scored present-and-equivalent or not. This technical index is a process endpoint, not a substitute for learning. It can be evaluated without releasing the learner's assessed work or registering a DOI.

Secondary outcomes include the six-week delayed transfer task, learner control over audience decisions, mentor and learner time, accessibility failures and unwanted exposure. Publication, submission, rejection, embargo and withdrawal counts are descriptive trajectories reported by assigned arm. The study does not rank students by those counts. No analysis conditions the primary effect on eventual acceptance, and a student's decision to keep work private does not remove that student from follow-up.

12. Estimand, estimator and multiplicity

The primary causal target is the equal-cluster average effect of offering the pathway on the baseline cohort's mean transfer-score change at week 20. Each cluster receives equal weight, irrespective of enrollment or publication uptake, and each matched pair contributes equally. The design therefore answers a policy question about a typical recruited instructional group. A student-weighted contrast may be reported secondarily, with its different target stated. Neither random assignment nor a registry audit establishes national representativeness.

For complete outcomes, calculate each cluster's baseline-cohort mean change, subtract the comparator mean from the intervention mean within each pair, and average the 48 pair differences. The standard error is their sample standard deviation divided by the square root of the pair count; the prespecified approximate interval uses the t distribution with 47 degrees of freedom. Pair-matched design reasoning is informed by Imai, King and Nall [13]. The supplied diagnostic tool implements this calculation and a sign-flip test of the sharp null for complete pair differences.

One two-sided primary test uses alpha .05. The provenance-mastery test is confirmatory only if the primary test rejects, following a fixed-sequence gate; otherwise it is descriptive with an interval. Other outcomes and subgroup contrasts are explicitly exploratory. The sharp-null sign-flip result and average-effect interval answer different questions. Neither supports a claim that every learner benefits, nor identifies an effect of publication itself among those who choose it.

13. Missing outcomes, clustering and sensitivity

Outcome collection continues after non-submission, editorial rejection, mentor changes or intervention discontinuation when consent permits. Missingness is reported by arm, pair and reason. The main analysis targets the fixed baseline cohort, not a favorable available-case subset. The analysis specification proposes 50 multilevel posterior-predictive imputations of bounded endline scores under a missing-at-random working model, including baseline score, assignment, educational stage, task form, prespecified access indicators, pair intercepts and mentor-cluster variation. Model priors, diagnostics and small-sample pooling rules are supplied separately and require an implementation check before registration [18].

Missing-at-random is not guaranteed by including many predictors. Prespecified pattern-mixture sensitivities shift imputed treatment-arm outcomes downward by one and two rubric points relative to the working model, repeat symmetric comparator shifts and

report bounded 0-24 extreme-case ranges. A wholly missing cluster is not discarded with its partner to manufacture complete pairs. Such a loss triggers explicit sensitivity analysis and changes the strength of the conclusion. The restricted Python diagnostic deliberately does not implement imputation or accept raw participant data.

Mentor spillovers, shared audiences and cross-cluster friendships are recorded as plausible interference, not assumed absent because the allocation file is tidy. The primary assigned-policy effect remains interpretable only under its stated exposure assumptions. A per-protocol analysis is exploratory. An instrumental-variable claim about actual publication is not planned: the offer can change motivation, rights literacy and feedback engagement before publication, violating a simple exclusion story.

14. Executed artificial-data design analysis

The completed design analysis follows an explicit aim, data-generating mechanism, estimand, method and performance-measure structure [15]. It uses 48 pairs, two clusters per pair and 12 baseline learners per cluster. Gaussian cluster and individual mean errors generate artificial change-score means, with a constant standardized effect of .30. The within-pair change-score intracluster correlation is varied across .05, .12 and .20; independent learner retention is 1.00, .85 or .70. Shared matched-pair variation cancels in the paired contrast.

Each of nine scenarios uses 20,000 replicates with random seed 20261016. At retention .85, simulated power is .96435, .88125 and .77520 for correlations .05, .12 and .20 respectively. Corresponding Monte Carlo standard errors are .00131, .00229 and .00295. Across all scenarios, power ranges from .74930 to .98020, null rejection from .04740 to .05265 and nominal 95% interval coverage from .94735 to .95260. These results were computed, not proposed or copied from another article.

The simulation does not validate a rubric, supply an empirical correlation or predict an actual institutional result. Retention is unrelated to outcomes, at least one learner per cluster is retained by construction, no cluster is lost and interference is absent. The effect size is measured in the artificial within-pair individual change-score standard deviation, not an established conversion to rubric points. Pilot evidence and actual recruitment constraints must replace these planning assumptions before a confirmatory sample is finalized.

15. Equity, burden and the right not to publish

A public portfolio can distribute opportunity unequally. Learners may have safety concerns, disability-related access needs, culturally specific reasons to control an audience or previous experiences of unwanted exposure. The protocol makes privacy-preserving work a complete academic route. Consent to research does not authorize publicity; course enrollment does not authorize permanent indexing; connecting an ORCID account does not authorize a journal to claim authorship. A withdrawn public-release choice is honored within documented technical and legal limits, without promising erasure from every external copy.

Both groups receive budgeted support, accessible task formats and equivalent assessment credit. Publication fees, where relevant to the intervention, cannot become a student-level treatment barrier or a reward for a favorable finding. A funding and fee plan must be resolved before recruitment. Independent safeguarding review can pause a release, and suspected unwanted exposure receives a response regardless of its effect on the study's completion count. Resource needs and barriers are measured directly instead of inferred from demographic stereotypes.

Weekly time diaries sample designated baseline, midpoint and endline windows. The candidate sustainability boundary is more than 45 additional minutes of uncompensated mentor work per learner per month without a learning gain judged educationally meaningful before allocation. This is a planning threshold for stakeholder ratification, not a validated universal rule. Burden, privacy events and inequitable uptake remain reportable harms even when the primary mean difference is positive.

16. Falsification, interpretation and governance

The theory is weakened if publication offers improve metadata completion and submission volume without improving independent transfer performance. It is further challenged if provenance mastery remains superficial, if learners cannot explain their own source choices, or if improvements depend on additional unpaid adult labor. A private authentic-audience pathway performing equally well would support an audience-and-feedback account over a claim that public release is necessary. Such a finding would be a useful design result, not a failed institutional publicity campaign.

Editorial acceptance is outside the instructional treatment's control. Editors receive submissions through the venue's ordinary criteria, may require revision or decline them, and cannot be asked to accept work to improve treatment compliance. If the intended author or platform leader has an editorial role, an independent handling arrangement

must govern a submission's decisions. A course grade, study treatment or publication fee cannot purchase a favorable editorial outcome. Trial reporting should follow applicable cluster-reporting principles, including flow at both cluster and learner levels [14].

The platform can preserve evidence, retrieve a cited passage or flag an incomplete manifest; it cannot infer consent, assign human authorship or certify peer review from a database match. A paper companion's responses must distinguish the reported 300-record metadata audit and artificial-data analysis from the unexecuted primary trial. This boundary gives superintendents a practical procurement criterion: a system should make unsupported inferences harder to display, rather than merely making records easier to generate.

17. Limitations

The metadata sample privileges alphabetically early DOI prefixes and three output types. Subject labels are supplied by repositories, and a Dissertation type is not an independently adjudicated student thesis. Records may be versions, components or inconsistent descriptions of the same broader project. The audit checks neither full-text accessibility nor source-specific reuse permission and authenticates no creator. Its counts describe a frozen interface sample, not publication quality, peer-review coverage or educational effectiveness.

The trial is a proposal. No participants have been enrolled or allocated in the research package, no outcomes collected and no ethical approval or registration identifier supplied. New instruments may fail during piloting. Volunteer institutions, mentor enthusiasm and the feasibility of private authentic audiences limit generalization. Two educational stages require careful matching and transparent reporting; sufficient statistical power for an overall effect does not establish power for subgroup interactions or rare privacy harms.

Finally, a randomized offer does not identify an isolated effect of being published. The pathway contains rights instruction, version reasoning, contribution negotiation and anticipation of an external decision. Equal time and readers reduce some rival explanations but do not decompose this package. A positive result would justify further replication and component studies, while a null or harmful result should inform the conditions under which a private learning record is preferable.

18. Implications for superintendents and university presses

Publication infrastructure should serve the learner's intellectual development and control over evidence. For a superintendent, that means specifying the audience, support, accessible alternatives and cost before choosing a platform. For a university press, it means preserving independent editorial decisions and portable provenance without turning a course's outputs into a guaranteed publication pipeline. Persistent identifiers can improve reference and continuity; they cannot replace a defensible assessment or the learner's own explanation.

This article supplies a testable next step from the volume's conceptual account. Its completed metadata and design analyses establish what was actually inspected and computed, while its prospective study makes a stronger educational claim available for future testing. The relevant institutional achievement would be an honest, sustainable improvement in transferable reasoning and provenance judgment. The size or visual polish of the public portfolio is not the criterion of success.

Method and data statement

Completed work comprises a three-query, 300-record DataCite metadata feasibility audit and nine 20,000-replicate artificial-data simulations. Public materials contain a minimized frozen projection, query and source hashes, executable offline audit and simulation tools, analysis results, proposed instruments and a study specification. They contain no participant records or linked-resource full texts. The raw response hashes and retention boundaries are documented. The primary mentor-cluster study, imputation model, validation pilot and ethical approvals remain prospective; no completed student learning effect is claimed.

AI, authorship and editorial disclosure

Dr. Freedom Cheteni is the intended human author and remains responsible for the argument, data interpretation and any eventual submission. ChatGPT assisted research, coding, analysis and drafting [17]; no precise underlying model version is asserted. DataCite, ORCID, cited authors and institutions are sources, not consenting coauthors or endorsers. This original protocol has no inherited human-review attestation, acceptance decision, publication date or assigned DOI. A source-grounded companion retrieves the article's statements and does not perform human peer review or create participant findings.

Research materials and reproducibility

Inspect the prospective study specification, original draft instruments, executed public-file audit, design simulations and documented tools. No local participant trial outcomes are reported.

Complete research package · ZIP

Methods and reproduction guide

Study specification

Proposed instruments and scoring rubrics

Statistical analysis specification

Frozen metadata projection

Executed metadata audit

Executed design sensitivity analysis

Executable offline research tools

Source and rights audit

Frozen queries and response hashes

Read the corresponding conceptual foundation

Evidence and limits

Claim	Evidence	Limit
A public DOI portfolio cannot by itself demonstrate learning or human authorship.	Separate provenance/transfer outcomes and source-specific metadata semantics. [6, 7, 12, 16]	The metadata audit tests supplied fields, not outcomes, identities or consent.
Portability fields vary across the bounded metadata sample.	Executed three-stratum 300-record audit: 119 creator-ORCID-bearing records, 136 resource versions, 97 version relations, 199 rights lists. [4, 5, 8, 9, 10, 11]	Alphabetic first-page strata are not representative or exhaustive; no full-text rights/access check.
Power depends materially on within-pair residual clustering and retention.	Nine executed 20,000-replicate artificial-data scenarios; power .74930-.98020 for δ = .30. [13, 15]	No empirical ICC, validated rubric SD, informative missingness or cluster loss is modeled.
The causal target is the effect of offering the pathway, not publication itself.	Matched-pair allocation before uptake, equal resources, independent editorial decisions and baseline-cohort follow-up. [2, 3, 13, 14, 18]	The study is prospective; implementation, instrument piloting, ethics determination and registration remain necessary.

References and rights

[1] Freedom Cheteni. Embracing the AI Revolution in Education: A Movement Thinking Approach. K12 Digest. 2023-11-09.

https://www.k12digest.com/embracing-the-ai-revolution-in-education-a-movement-thinking-approach/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[2] David Wiley; John Levi Hilton III. Defining OER-Enabled Pedagogy. The International Review of Research in Open and Distributed Learning, 19(4). 2018-09-26.

https://www.irrodl.org/index.php/irrodl/article/view/3601

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[3] Kwangsu Cho; Charles MacArthur. Student revision with peer and expert reviewing. Learning and Instruction, 20(4), 328-338. 2010-08.

https://www.sciencedirect.com/science/article/pii/S0959475209000747

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[4] DataCite. Queries and filtering. DataCite Support documentation. 2026 (accessed 2026-10-16).

https://support.datacite.org/docs/api-queries

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[5] DataCite. DataCite Data File Use Policy. DataCite Support. 2026 (accessed 2026-10-16).

https://support.datacite.org/docs/datacite-data-file-use-policy

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[6] DataCite Metadata Working Group. DataCite Metadata Schema 4.7: Overview. DataCite Metadata Schema documentation. 4.7 (accessed 2026-10-16).

https://datacite-metadata-schema.readthedocs.io/en/4.7/properties/overview/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[7] DataCite Metadata Working Group. DataCite Metadata Schema 4.7: 2. Creator. DataCite Metadata Schema documentation. 4.7 (accessed 2026-10-16).

https://datacite-metadata-schema.readthedocs.io/en/4.7/properties/creator/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[8] DataCite Metadata Working Group. DataCite Metadata Schema 4.7: 12. RelatedIdentifier. DataCite Metadata Schema documentation. 4.7 (accessed 2026-10-16).

https://datacite-metadata-schema.readthedocs.io/en/4.7/properties/relatedidentifier/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[9] DataCite Metadata Working Group. DataCite Metadata Schema 4.7: 15. Version. DataCite Metadata Schema documentation. 4.7 (accessed 2026-10-16).

https://datacite-metadata-schema.readthedocs.io/en/4.7/properties/version/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[10] DataCite Metadata Working Group. DataCite Metadata Schema 4.7: 16. Rights. DataCite Metadata Schema documentation. 4.7 (accessed 2026-10-16).

https://datacite-metadata-schema.readthedocs.io/en/4.7/properties/rights/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[11] DataCite Metadata Working Group. DataCite Metadata Schema 4.7: 8. Date. DataCite Metadata Schema documentation. 4.7 (accessed 2026-10-16).

https://datacite-metadata-schema.readthedocs.io/en/4.7/properties/date/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[12] ORCID. How do I collect an authenticated ID?. ORCID documentation. 2022-11-14.

https://info.orcid.org/ufaqs/how-do-i-collect-an-authenticated-id/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[13] Kosuke Imai; Gary King; Clayton Nall. The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation. Statistical Science, 24(1), 29-53. 2009-02.

https://doi.org/10.1214/08-STS274

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[14] Marion K. Campbell; Gilda Piaggio; Diana R. Elbourne; Douglas G. Altman; for the CONSORT Group. CONSORT 2010 statement: extension to cluster randomised trials. BMJ, 345:e5661. 2012-09-04.

https://www.bmj.com/content/345/bmj.e5661

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[15] Tim P. Morris; Ian R. White; Michael J. Crowther. Using simulation studies to evaluate statistical methods. Statistics in Medicine, 38(11), 2074-2102. 2019-01-16.

https://doi.org/10.1002/sim.8086

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[16] Paul Groth; Luc Moreau (editors). PROV-Overview: An Overview of the PROV Family of Documents. W3C Working Group Note. 2013-04-30.

https://www.w3.org/TR/prov-overview/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[17] OpenAI. ChatGPT [AI-assisted research, coding and drafting software]. Software disclosure; not a human author or peer reviewer. Accessed 2026-10-16.

https://chatgpt.com/

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

[18] John Barnard; Donald B. Rubin. Small-sample degrees of freedom with multiple imputation. Biometrika, 86(4), 948-955. 1999-12-01.

https://doi.org/10.1093/biomet/86.4.948

Citation only; no third-party full text or instrument republished. Original rights remain with original authors/rights holders. DataCite metadata reuse is separately documented; no linked-resource rights are inferred.

Authoritative source / evidence | checked 2026-10-16

This original primary-study protocol and completed metadata/design analyses retain credit to source authors and institutions. No linked-resource text is republished and no source findings are reassigned. Metadata identifiers do not establish identity, consent, peer review or learning.

Appendix: complete 100-source register

Supporting literature, not 100 new articles

Every source below retains its complete original attribution, venue, publication date, DOI, reusable version and exact article license. The register supports the volume and a continuing research programme. The source papers are not reassigned to Dr. Cheteni. The proposed questions are commissioning briefs, not completed research results.

Purposive screening of the existing 2,619-record open-research catalog plus a separately inspected APA catalog; all 131 open-research API pages read. This is not an exhaustive search of all school-leadership scholarship and these are not claimed to be newly discovered records.

Fresh Europe PMC fullTextXML; exact article DOI, article-level license and body checked. Author groups parsed separately from editorial groups.

Fresh Europe PMC core records; no retraction, withdrawal or expression-of-concern flags found for selected records as of 2026-10-15. Related corrections remain visible. This is not an independent certification of research integrity.

Search and filter the complete source register online

Machine-readable rights, notice metadata and source integrity records

001. School Leadership and Cyberbullying - A Multilevel Analysis

Sara B Låftman; Viveca Östberg; Bitte Modin. International Journal of Environmental Research and Public Health. 2017-10-15.

DOI 10.3390/ijerph14101226

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2017 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students’ digital and public-facing work?

002. A Multilevel Model of Teachers’ Job Performance: Understanding the Effects of Trait Emotional Intelligence, Job Satisfaction, and Organizational Trust

Mingwei Li; Pablo Alejandro Pérez-Díaz; Yaqing Mao; K. V. Petrides. Frontiers in Psychology. 2018-11-30.

DOI 10.3389/fpsyg.2018.02420

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2018 Li, Pérez-Díaz, Mao and Petrides.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

003. Effect of leadership style and organizational climate on employees' food safety and hygiene behaviors in the institutional food service of schools

Wen-Hwa Ko; Hsiu-yu Kang. Food Science & Nutrition. 2019-05-07.

DOI 10.1002/fsn3.1056

Publisher: Wiley. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2019 The Authors. Food Science & Nutrition published by Wiley Periodicals, Inc.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

004. School learning climate in the lens of parental involvement and school leadership: lessons for inclusiveness among public schools

Jason Alinsunurin. Smart Learning Environments. 2020-09-21.

DOI 10.1186/s40561-020-00139-2

Publisher: Springer Science and Business Media LLC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2020

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

005. Teacher-Rated School Leadership and Adolescent Gambling: A Study of Upper Secondary Schools in Stockholm, Sweden

Gabriella Olsson; Bitte Modin; Sara Brolin Låftman. International Journal of Environmental Research and Public Health. 2021-09-14.

DOI 10.3390/ijerph18189660

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2021 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

006. School principals' social support and teachers' basic need satisfaction: The mediating role of job demands and job resources

Jasper Maas; Simone Schoch; Urte Scholz; Pamela Rackow; Julia Schüler; Mirko Wegner; Roger Keller. Social Psychology of Education. 2022-10-31.

DOI 10.1007/s11218-022-09730-6

Publisher: Springer Science and Business Media LLC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2022

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

007. Linking Distributed Leadership with Differentiated Instruction in Inclusive Schools: The Mediating Roles of Teacher Leadership and Professional Competence

Tiantian Wang; Guoxiu Tian. Behavioral Sciences. 2023-11-30.

DOI 10.3390/bs13120990

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

008. The Chain Mediating Role of Teachers' Job Stress in the Influence of Distributed School Leadership on Job Satisfaction: Evidence from China, the United States, England, and Australia

Jian Li; Eryong Xue; Yuxuan Liu. Behavioral Sciences. 2024-03-27.

DOI 10.3390/bs14040279

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

009. Effort-reward imbalance and burnout among German school leaders. Representative study findings and implications for an underserved occupational group

Nele Groß; Kevin Dadaczynski; Marcus Pietsch. The European Journal of Public Health. 2025-02-01.

DOI 10.1093/eurpub/ckae182

Publisher: Oxford University Press. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2024. Published by Oxford University Press on behalf of the European Public Health Association.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

010. Effects of distributed leadership on teachers' job satisfaction: The mediating role of teacher collaboration

Li Ma; Ying Gu. PLOS One. 2026-04-30.

DOI 10.1371/journal.pone.0347862

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Ma, Gu

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

011. Linking distributed leadership to teachers' innovation: Chain mediating roles of commitment and collaboration in Chinese schools

Yu Zhao; Xiaoyang Li; Hongqin Kang. PLOS One. 2025-09-24.

DOI 10.1371/journal.pone.0333118

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 Zhao et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

012. The Influence of Distributed Leadership on Chinese Teachers' Job Satisfaction: The Chain Mediation of Teacher Collaboration and Teacher Self-Efficacy

Xiaodong Fan; Zuwang Chu. International Journal of Environmental Research and Public Health. 2025-03-26.

DOI 10.3390/ijerph22040507

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

013. The long-term effects of perceived instructional leadership on teachers' psychological well-being during COVID-19

Xiu-Mei Chen; Xiao Ling Liao; I-Hua Chen; Jeffrey H. Gamble; Xing-Yong Jiang; Xu-Dong Li; Cun-Xu Bo. PLOS ONE. 2024-08-19.

DOI 10.1371/journal.pone.0305494

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 Chen et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

014. The relationship between distributed leadership and teachers' teaching enthusiasm: the chain mediation of teacher autonomy and organizational commitment

Yang Zhang; Yipu Hao. Frontiers in Psychology. 2026-08-19.

DOI 10.3389/fpsyg.2026.1896309

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Zhang and Hao.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

015. The relationship between distributed leadership and the enthusiasm and professional collaboration levels of teachers in Türkiye: the moderator effect of teacher optimism

Yurdagül Doğuş. Frontiers in Psychology. 2026-03-04.

DOI 10.3389/fpsyg.2026.1729734

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Doğuş.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

016. The well-being and work-related stress of senior school leaders in Wales and Northern Ireland during COVID-19 “educational leadership crisis”: A cross-sectional descriptive study

Emily Marchant; Joanna Dowd; Lucy Bray; Gill Rowlands; Nia Miles; Tom Crick; Michaela James; Kevin Dadaczynski; Orkan Okan. PLOS ONE. 2024-04-10.

DOI 10.1371/journal.pone.0291278

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 Marchant et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

017. A study to examine correlation between instructional leadership and teachers' motivation: evidence from TRNC

Hayat Kalkan; Ahmet Munir Acuner. Frontiers in Psychology. 2026-05-20.

DOI 10.3389/fpsyg.2026.1759126

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Kalkan and Acuner.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

018. Association between distributed leadership and teachers' organizational commitment: Evidence from public and private schools in the United Arab Emirates

Ali Ibrahim; Shaikha Alshehhi; Nikolaos Tsingilis. PLOS One. 2026-06-23.

DOI 10.1371/journal.pone.0351540

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Ibrahim et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

019. Determinants of instructional leadership: A survey of public secondary schools

Suman Pande; Laxman Acharya; Prakash C. Bhattarai. PLOS One. 2026-08-14.

DOI 10.1371/journal.pone.0355674

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Pande et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

020. Distributed Leadership and Teachers' Social and Emotional Learning Practices in Chinese Schools: A Relational Coordination Perspective Using TALIS 2024 Data

Mengting Qian; Xianjun Yu; Jingyi Yang; Chunshun Yan. Behavioral Sciences. 2026-04-01.

DOI 10.3390/bs16040523

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

021. Effective School Leadership for Supporting Students' Mental Health: Findings from a Narrative Literature Review

Brian P Daly; Annie Resnikoff; Shannon Litke. Behavioral Sciences. 2025-01-01.

DOI 10.3390/bs15010036

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

022. How Collective Efficacy Mediates the Association between Principal Instructional Leadership and Teacher Self-Efficacy: Findings from a Meta-Analytic Structural Equation Modeling (MASEM) Study

Turgut Karakose; Abdurrahman Kardas; Sedat Kanadlı; Tijen Tülübaş; Bilal Yildirim. Behavioral Sciences. 2024-01-25.

DOI 10.3390/bs14020085

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

023. How Distributed Leadership Affects Social and Emotional Competence in Adolescents: The Chain Mediating Role of Student-Centered Instructional Practices and Teacher Self-Efficacy

Zhenyu Li; Wei Liu; Qiong Li. Behavioral Sciences. 2024-02-13.

DOI 10.3390/bs14020133

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

024. Is there a significant relationship between the empowering leadership behaviors of school principals and the psychological resilience of teachers? Understanding the moderating effects of gender and length of time spent with the school principal

Mehmet Sabir Çevik; Emine Doğan. BMC Psychology. 2025-01-08.

DOI 10.1186/s40359-024-02330-3

Publisher: BMC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2025

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

025. Leadership of school principals for school health implementation among primary schools in Mataram, Indonesia: a qualitative study

Hirono Sasaki; Dian Puspita Sari; Cut Warnaini; Fahrin Ramadan Andiwijaya; Rie Takeuchi; Hamsu Kadriyan; Fumiko Shibuya; Jun Kobayashi. Tropical Medicine and Health. 2024-01-02.

DOI 10.1186/s41182-023-00568-y

Publisher: BMC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2023

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

026. Perspectives of School Principals on Mental Health Promotion and Prevention Among School-Aged Children - A 2025 Cross-Sectional Survey in Lodz Administrative Region in Poland

Aleksandra Lewandowska; Mateusz Jankowski; Mariusz Gujski; Agata Andrzejczyk; Justyna Teliga-Czajkowska; Andrzej Silczuk. Healthcare. 2025-06-23.

DOI 10.3390/healthcare13131498

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students’ digital and public-facing work?

027. School Principals’ Perspectives and Leadership Styles for Digital Transformation: A Q-Methodology Study

Peili Yuan; Xinshen Chen; Huan Song. Behavioral Sciences. 2026-01-24.

DOI 10.3390/bs16020165

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which human decisions and learning goals should govern this technology rather than its novelty?

028. School leadership and indigenous student outcomes: A correlational study of Student Needs Management Practices based on Maslow’s Hierarchy of Needs

Yusoff Yahaya; Bity Salwana Alias; Mohd Norazmi Nordin. PLOS One. 2026-07-30.

DOI 10.1371/journal.pone.0350479

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Yahaya et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

029. School principals’ span of control and its cross-sectional associations with signs of exhaustion, perceived work ability, and demanding and supportive circumstances in managerial work

Roger Persson; Ulf Leo; Anna Oudin; Carita Håkansson; Kai Österberg. BMC Public Health. 2026-04-21.

DOI 10.1186/s12889-026-27265-9

Publisher: BMC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2026

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

030. Teachers’ perceived time pressure, emotional exhaustion and the role of social support from the school principal

Jasper Maas; Simone Schoch; Urte Scholz; Pamela Rackow; Julia Schüler; Mirko Wegner; Roger Keller. Social Psychology of Education. 2021-03-11.

DOI 10.1007/s11218-020-09605-8

Publisher: Springer Science and Business Media LLC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2021

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

031. The Impact of School Leadership on Inclusive Education Literacy: Examining the Sequential Mediation of Job Stress and Teacher Agency

Yulu Feng; Dan Zhou; Yihong Wei. Behavioral Sciences. 2025-11-17.

DOI 10.3390/bs15111572

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

032. The Influence of Teachers' Self-Efficacy and School Leaders' Transformational Leadership Practices on Teachers' Innovative Behaviour

Mohammed Afandi Zainal; Mohd Effendi Ewan Mohd Matore. International Journal of Environmental Research and Public Health. 2021-06-14.

DOI 10.3390/ijerph18126423

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2021 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

033. The Mediating Role of the Principal-Teacher Relationship in Innovative School Leadership and Teacher Professional Learning According to Turkish Teachers' Perceptions

Mehmet Özdoğru; Yurdagül Doğuş; Muhammet İbrahim Akyürek. Behavioral Sciences. 2025-04-01.

DOI 10.3390/bs15040450

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

034. The Relationship between Principals' Instructional Leadership and Teacher Retention in the Undeveloped Regions of Central and Western China: The Chain-Mediating Role of Role Stress and Affective Commitment

Qifang Zhan; Xiaohan Wang; Huan Song. Behavioral Sciences. 2023-09-04.

DOI 10.3390/bs13090738

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

035. Toward agenda 2030 in education: policies and practices for effective school leadership

Norma Ghamrawi. Educational Research for Policy and Practice. 2023-04-08.

DOI 10.1007/s10671-023-09341-8

Publisher: Springer Science and Business Media LLC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2023

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

036. Why is school leadership key to transforming education? Structural and cultural assumptions for quality education in diverse contexts

Monica Mincu. Prospects. 2022-10-31.

DOI 10.1007/s11125-022-09625-6

Publisher: Springer Science and Business Media LLC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2022

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

037. Explaining the Links Between School Administrator Leadership, Job Satisfaction, and Participatory School Climate: A Machine Learning-Enhanced Multilevel Analysis of TALIS 2024 School Administrator Data

Dönüş Şengür. Behavioral Sciences. 2026-06-26.

DOI 10.3390/bs16071062

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the author.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students’ digital and public-facing work?

038. Leadership, autonomy, and organizational trust as predictors of teacher wellbeing and job satisfaction: a cross-cultural study based on PISA 2022 data

Servet Üztemur; Ali Gökalp; Abdurrahman İlğan; Erkan Dinç. Frontiers in Psychology. 2026-01-02.

DOI 10.3389/fpsyg.2025.1703458

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Üztemur, Gökalp, İlğan and Dinç.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

039. Principal visionary leadership and teacher instructional innovation: mediating roles of teacher workplace well-being and collaboration

Chuang Yang; Zeqing Xu. Frontiers in Psychology. 2026-01-21.

DOI 10.3389/fpsyg.2026.1701099

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Yang and Xu.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

040. Promoting Teachers' Organizational Commitment: The Effects of Authentic Leadership, Teachers' Well-Being and Social-Emotional Competence

Zeqing Xu; Nicholas Sun-Keung Pang. Behavioral Sciences. 2024-09-24.

DOI 10.3390/bs14100862

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

041. Promoting teaching innovation of Chinese public-school teachers by team temporal leadership: The mediation of job autonomy and the moderation of work stress

Kai Li; Guiqin Zhu. PLoS ONE. 2022-07-08.

DOI 10.1371/journal.pone.0271195

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2022 Li, Zhu

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

042. Relationships Between Positive Leadership Styles, Psychological Resilience, and Burnout: An Empirical Study Among Turkish Teachers

Gaye Onan; Lütfi Sürücü; Mustafa Bekmezci; Alper Bahadır Dalmış; Gözde Sunman. Behavioral Sciences. 2025-05-22.

DOI 10.3390/bs15060713

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

043. Teachers' Emotional Intelligence and Organizational Commitment: A Moderated Mediation Model of Teachers' Psychological Well-Being and Principal Transformational Leadership

Mingwei Li; Feifei Liu; Chuanli Yang. Behavioral Sciences. 2024-04-20.

DOI 10.3390/bs14040345

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

044. Teachers' Well-Being and Innovative Work Behavior: A Moderated Mediation Model of Perceived Insider Status and Principal Authentic Leadership

Chao Lu; Zeqing Xu; Qinrui Tian. Behavioral Sciences. 2025-10-19.

DOI 10.3390/bs15101419

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

045. The Influence of Servant Leadership on the Professional Well-Being of Kindergarten Teachers: A Moderated Mediation Model

Dongqing Yu; Wenxin Yue; Shuang Hao; Dengyin Li; Qiurong Wu. Behavioral Sciences. 2025-10-17.

DOI 10.3390/bs15101412

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

046. The impact of authentic leadership on the work engagement of primary and secondary school teachers: The serial mediation role of school climate and teacher efficacy

Yanhong Shao; Wenxuan Jiang; Ningjun Wang; Chao Zhang; Lili Zhang. PLOS One. 2025-05-23.

DOI 10.1371/journal.pone.0320839

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 Shao et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students’ digital and public-facing work?

047. The relationship between inclusive leadership in sustainable education and teachers' subjective wellbeing in Türkiye: the mediating role of affective commitment

Mehmet Nezir Çevik. Frontiers in Psychology. 2025-12-12.

DOI 10.3389/fpsyg.2025.1689020

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2025 Çevik.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

048. A dual mediation model of the association between principal inclusive leadership and teacher innovative behavior

Yuangen Bao. Scientific Reports. 2024-05-29.

DOI 10.1038/s41598-024-63332-0

Publisher: Nature Publishing Group. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2024

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

049. Can teacher ostracism be prevented? Exploring how empowering leadership can mitigate teacher ostracism through work engagement

Alper Uslukaya. The British Journal of Educational Psychology. 2025-02-03.

DOI 10.1111/bjep.12745

Publisher: Wiley. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 The Author(s). British Journal of Educational Psychology published by John Wiley & Sons Ltd on behalf of British Psychological Society.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

050. Chain-mediated effects of servant leadership on the physical health of Chinese elementary school teachers

Hongchao Wu; Ruiqi Zhou; Xiuhong Li; Lihui Yan; Congying Ma. Frontiers in Public Health. 2025-12-03.

DOI 10.3389/fpubh.2025.1691696

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2025 Wu, Zhou, Li, Yan and Ma.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

051. Cross-Level Influence of Group-Focused Transformational Leadership on Organizational Citizenship Behavior among Chinese Secondary School Teachers

Zhuotao Fang; Shun-Chi Yu. Behavioral Sciences. 2023-10-16.

DOI 10.3390/bs13100848

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

052. Design and evaluation of leadership development program using co-creation approaches for an athlete leadership group among elite high school rugby teams

Toshiaki Hirose; Yasutaka Ojio; David Lavallee; Naohiko Kohtake. Frontiers in Sports and Active Living. 2026-03-06.

DOI 10.3389/fspor.2026.1770549

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Hirose, Ojio, Lavallee and Kohtake.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

053. Destructive Leadership and Turnover Intention among Chinese Rural Kindergarten Teachers: The Mediation of Ego Depletion and the Moderation of Kindergarten Affiliation

Can He; Jie Xiong; Yue Zhang; Haiyuan Dou; Jiahui Du. International Journal of Environmental Research and Public Health. 2023-01-22.

DOI 10.3390/ijerph20032030

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

054. Development and validation of a deep learning-based assessment tool for teacher leadership: A case study from Xinjiang, China

Jianwei Dong; Xinya Chen; Chen Chen; Cheng Chen. PLOS One. 2025-09-02.

DOI 10.1371/journal.pone.0331560

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 Dong et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which human decisions and learning goals should govern this technology rather than its novelty?

055. Digital leadership as an environmental determinant of teachers' cognitive and affective mechanisms in artificial intelligence adoption: an integrated UTAUT2-GETAMEL model

Fatma Hümeyra Yücel. Frontiers in Psychology. 2026-06-03.

DOI 10.3389/fpsyg.2026.1822713

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Yücel.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which human decisions and learning goals should govern this technology rather than its novelty?

056. Effects of community youth teams facilitating participatory adolescent groups, youth leadership activities and livelihood promotion to improve school attendance, dietary diversity and mental health among adolescent girls in rural eastern India (JIAH trial): A cluster-randomised controlled trial

Komal Bhatia; Suchitra Rath; Hemanta Pradhan; Subhashree Samal; Andrew Copas; Sumitra Gagrai; Shibanand Rath; Raj Kumar Gope; Nirmala Nair; Prasanta Tripathy; Kelly Rose-Clarke; Audrey Prost. SSM - Population Health. 2022-12-27.

DOI 10.1016/j.ssmph.2022.101330

Publisher: Elsevier. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2022 The Authors. Published by Elsevier Ltd.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students’ digital and public-facing work?

057. Engagement in Transformational Leadership by Teachers Influences the Levels of Self-Esteem, Motor Self-Efficacy, Enjoyment, and Intention to Be Active in Physical Education Students

Carolina Sánchez-García; Rafael E Reigal; Jacobo Hernández-Martos; Antonio Hernández-Mendo; Veronica Morales-Sánchez. Sports. 2024-07-15.

DOI 10.3390/sports12070191

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

058. Enhancing Inclusive Teaching in China: Examining the Effects of Principal Transformational Leadership, Teachers’ Inclusive Role Identity, and Efficacy

Dongsheng Wang; Liang Huang; Xianhan Huang; Meng Deng; Wanying Zhang. Behavioral Sciences. 2024-02-23.

DOI 10.3390/bs14030175

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

059. Exploring Prajnamitra Maitreya Buddhists School Pekanbaru: Do leadership, work environment, and organisational culture influence the teachers' competence and work performance?

Fazal Mohamed Mohamed Sultan; Gunasegaran Karuppannan; Nyoto; Herni Lestari. PLOS ONE. 2023-05-16.

DOI 10.1371/journal.pone.0282941

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 Mohamed Sultan et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

060. How Emotional Contagion among Teachers Affects the Relationship between Transformational Leadership and Team Cohesion

Giulia Paganin; Lorenzo Avanzi; Dina Guglielmi; Carlos-María Alcover; Greta Mazzetti. Behavioral Sciences. 2023-08-16.

DOI 10.3390/bs13080685

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

061. How Transformational Leadership Shapes Teachers' Innovative Behavior: A Serial Mediation Model of Creative Growth Mindset and Work Engagement

Yilin Shen; Yi Li; Kaili Tang; Jia Wu; Qianfeng Li; Shaobei Xiao. Behavioral Sciences. 2026-08-03.

DOI 10.3390/bs16081321

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

062. How do various leadership practices affect professional learning communities? The mediating role of principals’ perceived trust by teachers

Xiaobo Gu; Zhihui Liu; Zhenyuan Hang. PLOS One. 2025-05-06.

DOI 10.1371/journal.pone.0323055

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 Gu et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

063. How paternalistic leadership influences teachers' resistance to STEM reform: the mediating role of teachers' STEM literacy and the moderating role of work engagement

Yao Zhang; Dongchen Zhao. Frontiers in Psychology. 2026-02-23.

DOI 10.3389/fpsyg.2026.1726148

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Zhang and Zhao.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

064. How perceived overqualification affects work engagement among new-generation teachers: the mediating role of harmonious work passion and moderating role of humble leadership

Li Jiang; Xiaoli Xie. Frontiers in Psychology. 2026-09-03.

DOI 10.3389/fpsyg.2026.1911357

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Jiang and Xie.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

065. How transformational leadership is associated with primary and secondary school teachers' social-emotional competence: a multiple mediation analysis

Qilan Li; Jinwei Tu; Biao Zeng. Frontiers in Psychology. 2026-06-03.

DOI 10.3389/fpsyg.2026.1777257

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Li, Tu and Zeng.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

066. Leadership Competence and Psychosocial Safety Climate Implementation in an Evolving School Work Environment

Stefano Cataloni; Darryl Forsyth; David Brougham; Kaye Thorn. International Journal of Environmental Research and Public Health. 2026-04-29.

DOI 10.3390/ijerph23050573

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

067. Leadership and Conflict Management Style Are Associated with the Effectiveness of School Conflict Management in the Region of Epirus, NW Greece

Elisavet Chandolia; Sophia Anastasiou. European Journal of Investigation in Health, Psychology and Education. 2020-02-14.

DOI 10.3390/ejihpe10010034

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2020 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

068. Leading the charge in the education sector: development and validation of the School Implementation Leadership Scale (SILS)

Aaron R Lyon; Catherine M Corbin; Eric C Brown; Mark G Ehrhart; Jill Locke; Chayna Davis; Elissa Picozzi; Gregory A Aarons; Clayton R Cook. Implementation Science : IS. 2022-07-19.

DOI 10.1186/s13012-022-01222-7

Publisher: BMC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2022

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

Related notices are recorded. Consult the original source and linked notice metadata before reuse.

069. Longitudinal associations among team orientation, leadership, and teamwork behaviors in community-based school management

Takuya Yoshida; Hiroyuki Yoshizawa; Ryosuke Asano. PLOS One. 2026-08-27.

DOI 10.1371/journal.pone.0357007

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Yoshida et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

070. Organizational and social work environment factors, occupational balance and no or negligible stress symptoms among Swedish principals - a cross-sectional study

Carita Håkansson; Ulf Leo; Anna Oudin; Inger Arvidsson; Kerstin Nilsson; Kai Österberg; Roger Persson. BMC Public Health. 2021-04-26.

DOI 10.1186/s12889-021-10809-6

Publisher: BMC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2021

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

071. Paternalistic leadership and work engagement: the roles of knowledge sharing and supervisor support among calligraphy teachers in Hebei primary schools

Yu Yang; Lydia Yoke Yean Foong; Xiangna Zhang; Baihui Wang; Fan Yang. Frontiers in Psychology. 2026-07-20.

DOI 10.3389/fpsyg.2026.1821998

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Yang, Foong, Zhang, Wang and Yang.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

072. Performing womanhood in schools: a feminist phenomenological investigation of female teachers' psycho-social health experiences and implications for educational leadership in Turkey

Engin İŞ. Frontiers in Psychology. 2025-12-03.

DOI 10.3389/fpsyg.2025.1692323

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2025 İŞ.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Whose authority, access or unpaid labor changes when this finding is translated into a future-school design?

073. Relationship between Principals' Leadership Styles and Teachers' Behavior

Kazi Enamul Hoque; Zarin Tasnim Raya. Behavioral Sciences. 2023-01-28.

DOI 10.3390/bs13020111

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

074. Revisiting Louis Fry's spiritual leadership model in confessional school teachers using structural equation modeling (SEM)

Lorena Martinez-Soto; Flor Ontiveros Ramírez; Iván Dario Toro Jaramillo; Nelly Rosario Moreno-Leyva. PLOS ONE. 2024-09-17.

DOI 10.1371/journal.pone.0299671

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 Martinez-Soto et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

075. Servant Leadership and Teachers' Emotional Exhaustion - The Mediation Role of Hindrance Stress and Depersonalization

Hongchao Wu; Jun Zhao; Shaoping Qiu; Xiuhong Li. Behavioral Sciences. 2024-11-25.

DOI 10.3390/bs14121129

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

076. Stance of numerous leadership styles and their effect on teaching to sustain academic performance at the high school level

Samra Maqbool; Hafiz Muhammad Ihsan Zafeer; Sufyan Maqbool; Pingfei Zeng; Zineb Draissi; Saima Javed. Heliyon. 2024-08-17.

DOI 10.1016/j.heliyon.2024.e36438

Publisher: Elsevier. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 The Authors

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

077. Teachers' Perceptions of Shared Leadership and Their Relationship with Organizational Attractiveness and Identification: A Structural Equation Modeling Approach

Nesip Demirbilek. Journal of Intelligence. 2025-11-05.

DOI 10.3390/jintelligence13110141

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the author.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: When responsibilities are distributed, do educators gain usable authority or additional unresourced work?

078. The Development of a New Model of Educational Leadership: Leadership for Teacher Flourishing

Katy Granville-Chapman; Matthew T Lee; James Ritchie-Dunham. Humanistic Management Journal. 2024-06-14.

DOI 10.1007/s41463-024-00181-z

Publisher: Springer Science and Business Media LLC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© Crown 2024

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

079. The Multilevel Effects of Principals' Servant Leadership on Kindergarten Teachers' Job Crafting: The Mediating Role of Organizational Identification

Xiaoqing Lin; Runkai Jiao; Feifei Li. Behavioral Sciences. 2026-02-27.

DOI 10.3390/bs16030329

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

080. The Narrow Corridor of Heartfelt Leadership: Social and Emotional Leadership Practices in Bureaucratic School Cultures

Sevgi Yıldız. Behavioral Sciences. 2025-09-26.

DOI 10.3390/bs15101316

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 by the author.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

081. The influence of resonant leadership on teachers' innovative behavior: The mediating mechanism of information processing models

Yuangen Bao. PLOS One. 2025-04-11.

DOI 10.1371/journal.pone.0321763

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2025 Yuangen Bao

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

082. The relationship between principals' pedagogical leadership practices and teachers' job performance: the role of sociodemographic characteristics

Kelemu Zelalem Berhanu. BMC Psychology. 2025-01-30.

DOI 10.1186/s40359-025-02415-7

Publisher: BMC. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© The Author(s) 2025

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What support would allow teachers to exercise informed judgment when the curriculum and tools change?

083. Transformative educational leadership as organizational practice: a practice theory perspective on reconfiguring teacher agency and professional culture in schools

Nur Efendi; Muh Ibnu Sholeh; Nurul Indana; Fahrur Rosikh; Choiruddin; Nur 'Azah; Sholihan; Imro'atus Solikah; Riyanto. Frontiers in Sociology. 2026-08-07.

DOI 10.3389/fsoc.2026.1847035

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Efendi, Sholeh, Indana, Rosikh, Choiruddin, 'Azah, Sholihan, Solikah and Riyanto.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How should a superintendent translate this source into a locally accountable experiment rather than a universal prescription?

084. Understanding teachers' curriculum leadership: a field-mediator mechanism

Yi Wan. Frontiers in Psychology. 2026-04-28.

DOI 10.3389/fpsyg.2026.1663486

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2026 Wan.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

085. Unraveling Reading Achievement Through Educational Leadership, School Actions to Sustain Learning, Digital Self-Efficacy, and ICT-Related Factors: A Multilevel Mediation Analysis of PISA 2022 Türkiye Data

Nermin Er Aydemir; Ahmet Şahin. Journal of Intelligence. 2026-07-03.

DOI 10.3390/jintelligence14070137

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which human decisions and learning goals should govern this technology rather than its novelty?

086. A Systematic Review of Research Patterns and Trends in Early Childhood Education Teacher Well-Being from 1993 to 2023: A Trajectory Landscape

Xinlin Li; Lutong Zhang; John Chi-Kin Lee; Junjun Chen. Behavioral Sciences. 2024-12-10.

DOI 10.3390/bs14121179

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2024 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: What evidence would show that teacher agency improves the quality of learning work rather than only the quantity of projects?

087. An intervention to improve teacher well-being support and training to support students in UK high schools (the WISE study): A cluster randomised controlled trial

Judi Kidger; Nicholas Turner; William Hollingworth; Rhiannon Evans; Sarah Bell; Rowan Brockman; Lauren Copeland; Harriet Fisher; Sarah Harding; Jillian Powell; Ricardo Araya; Rona Campbell; Tamsin Ford; David Gunnell; Simon Murphy; Richard Morris. PLoS Medicine. 2021-11-11.

DOI 10.1371/journal.pmed.1003847

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2021 Kidger et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which concrete working conditions would make relational leadership usable during technological and organizational change?

088. Child-to-Parent Violence: Attitude towards Authority, Social Reputation and School Climate

Gonzalo Del Moral; Cristian Suárez-Relinque; Juan E Callejas; Gonzalo Musitu. International Journal of Environmental Research and Public Health. 2019-07-05.

DOI 10.3390/ijerph16132384

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2019 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students’ digital and public-facing work?

089. Does Accountability Aggravate the Risk of Teacher Burnout? Evidence from the Chinese Education System

Guanyu Li; Kwok Kuen Tsang. Behavioral Sciences. 2023-09-15.

DOI 10.3390/bs13090772

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2023 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Can a future-of-work intervention reduce demands without shifting unpaid repair work to another group?

090. Effectiveness of the TEI Program for Bullying and Cyberbullying Reduction and School Climate Improvement

Rosario Ferrer-Cascales; Natalia Albaladejo-Blázquez; Miriam Sánchez-SanSegundo; Irene Portilla-Tamarit; Oriol Lordan; Nicolás Ruiz-Robledillo. International Journal of Environmental Research and Public Health. 2019-02-16.

DOI 10.3390/ijerph16040580

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2019 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

091. Effectiveness of universal school-based mindfulness training compared with normal school provision on teacher mental health and school climate: results of the MYRIAD cluster randomised controlled trial

Willem Kuyken; Susan Ball; Catherine Crane; Poushali Ganguli; Benjamin Jones; Jesus Montero-Marin; Elizabeth Nuthall; Anam Raja; Laura Taylor; Kate Tudor; Russell M Viner; Matthew Allwood; Louise Aukland; Darren Dunning; Tríona Casey; Nicola Dalrymple; Katherine De Wilde; Eleanor-Rose Farley; Jennifer Harper; Verena Hinze; Nils Kappelmann; Maria Kempnich; Liz Lord; Emma Medlicott; Lucy Palmer; Ariane Petit; Alice Philips; Isobel Pryor-Nitsch; Lucy Radley; Anna Sonley; Jem Shackleford; Alice Tickell; MYRIAD Team; Sarah-Jayne Blakemore; Obioha C Ukoumunne; Mark T Greenberg; Tamsin Ford; Tim Dalgleish; Sarah Byford; J Mark G Williams. Evidence-Based Mental Health. 2022-07-12.

DOI 10.1136/ebmental-2022-300424

Publisher: BMJ Publishing Group. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY. Published by BMJ.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

092. Human-Machine Collaboration Depth and Teacher Burnout: A Dual-Path Moderated Mediation Model of Empowerment and Depletion

Xiaoyu Guo; Man Li; Xin Zhao. Behavioral Sciences. 2026-08-03.

DOI 10.3390/bs16081324

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: Which human decisions and learning goals should govern this technology rather than its novelty?

093. Investigation of the Impact of School Climate and Teachers' Self-Efficacy on Job Satisfaction: A Cross-Cultural Approach

Ioannis G Katsantonis. European Journal of Investigation in Health, Psychology and Education. 2019-10-14.

DOI 10.3390/ejihpe10010011

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2019 by the author.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

094. Mediating Processes in the Relations of Parental Monitoring and School Climate With Cyberbullying: The Role of Moral Disengagement

Maria Giuseppina Bartolo; Anna Lisa Palermiti; Rocco Servidio; Pasquale Musso; Angela Costabile. Europe's Journal of Psychology. 2019-09-27.

DOI 10.5964/ejop.v15i3.1742

Publisher: PsychOpen. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

095. Parental Warmth, Gratitude, and Prosocial Behavior among Chinese Adolescents: The Moderating Effect of School Climate

Haocheng Luo; Qingqi Liu; Chengfu Yu; Yangang Nie. International Journal of Environmental Research and Public Health. 2021-06-30.

DOI 10.3390/ijerph18137033

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2021 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

Related notices are recorded. Consult the original source and linked notice metadata before reuse.

096. Perceptions of School Climate and Internet Gaming Addiction among Chinese Adolescents: The Mediating Effect of Deviant Peer Affiliation

Hongyu Zou; Yuting Deng; Huahua Wang; Chengfu Yu; Wei Zhang. International Journal of Environmental Research and Public Health. 2022-03-18.

DOI 10.3390/ijerph19063604

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2022 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

097. Prevalence of mental health, mental well-being, and school connectedness in UK secondary school students in 2021/2022: Results from the South West - School Health Research Network pilot study

Patricia N. Albers; Emily Widnall; Lorna Hatch; Georgina Hopkins; Catherine Sharp; Judi Kidger; Frank de Vocht; Eileen Kaner; Esther M. F. van Sluijs; Hannah Fairbrother; Russell Jago; Rona Campbell. PLOS One. 2026-08-28.

DOI 10.1371/journal.pone.0355874

Publisher: PLOS. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2026 Albers et al

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

098. School Bullying in Urban China: Prevalence and Correlation with School Climate

Ziqiang Han; Guirong Zhang; Haibo Zhang. International Journal of Environmental Research and Public Health. 2017-09-25.

DOI 10.3390/ijerph14101116

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2017 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

099. School Climate and Responsibility as Predictors of Antisocial and Prosocial Behaviors and Violence: A Study towards Self-Determination Theory

David Manzano-Sánchez; Alberto Gómez-Mármol; Alfonso Valero-Valenzuela; José Francisco Jiménez-Parra. Behavioral Sciences. 2021-03-17.

DOI 10.3390/bs11030036

Publisher: Multidisciplinary Digital Publishing Institute (MDPI). Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

© 2021 by the authors.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

100. School Climate, Loneliness, and Problematic Online Game Use Among Chinese Adolescents: The Moderating Effect of Intentional Self-Regulation

Chengfu Yu; Wentao Li; Qiao Liang; Xuelan Liu; Wei Zhang; Hong Lu; Kai Dou; Xiaodong Xie; Xiong Gan. Frontiers in Public Health. 2019-04-30.

DOI 10.3389/fpubh.2019.00090

Publisher: Frontiers Media SA. Version: Journal article in Europe PMC. License: CC BY 4.0. Verified 15 October 2026.

Copyright © 2019 Yu, Li, Liang, Liu, Zhang, Lu, Dou, Xie and Gan.

Exact license terms

Exact reusable full text / permission evidence

Proposed inquiry: How can school leadership preserve belonging and safety while expanding students' digital and public-facing work?

Article format and publication record

A reusable scholarly reading format

Each article identifies its contribution, defined constructs, source designs, methods, competing explanations and limitations. Linked references and evidence-and-limits tables make the warrant inspectable. The online edition adds section navigation, accessible citation links, BibTeX/RIS exports, print/PDF controls and an extractive companion that retrieves article passages without generating a conclusion.

Editorial record and publication

The public review record preserves Dr. Vickie Aultmon’s review of the five conceptual manuscripts and the earlier A01 conceptual text. The substantial version 3.0 additions to A01 and new primary-study protocols have separate content and review scopes. Author approval, independent editorial decisions, rights and DOI registration are maintained as separate records. Acceptance and public journal publication dates describe actual events.

Reusable online article format

Crossref correspondence workspace