Research Articles

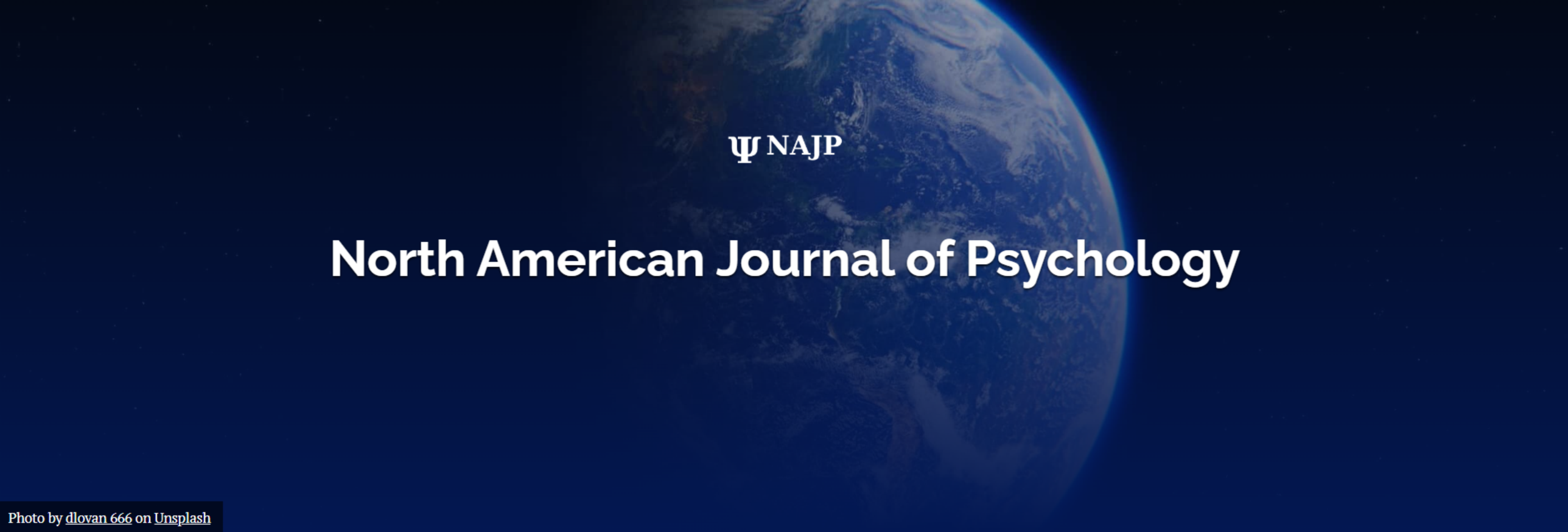
July 29, 2026 PDT

Validating a Two-Factor School Discipline Measure for Adolescents: An Unrestricted Factor Analysis in Indonesian Samples

Yohanes Budiarto, Dr., Bonar Hutapea

School discipline | two-factor solution | bifactor model | punitive | non-punitive

 • https://doi.org/10.65696/001c.165924


ΨNAJP
North American Journal of Psychology
Photo by dlovan 666 on Unsplash

RESEARCH ARTICLES

July 29, 2026 PDT

Validating a Two-Factor School Discipline Measure for Adolescents: An Unrestricted Factor Analysis in Indonesian Samples

Yohanes Budiarto, Bonar Hutapea

The study compares two models of school discipline, finding both a unified perception and two distinct approaches. It recommends a hybrid scoring approach for comprehensive evaluation of discipline practices.

Abstract

RESEARCH ARTICLES

July 19, 2026 PDT

Factors Associated With Stress Among Social Workers in Public Social Service Institutions in the Southeastern Region of Vietnam

Lua Vu, Tuan Pham Van

This study identifies optimism, job satisfaction, perceived social support, and turnover intention as factors affecting stress levels in Vietnamese social workers, explaining 42.7% of stress variance.

Abstract

RESEARCH ARTICLES

July 18, 2026 PDT

Examining the Contributions of Knowledge and Constructivist Beliefs on Filipino Pre-Service Mathematics Teachers' Intentions to Use Real-Life Tasks

Juvenal Dalimeg, Antonio Ponciano Rodil, Kimberly Ann S. Cantilero

This study reveals that pre-service mathematics teachers' constructivist beliefs and self-efficacy are key factors in their intention to use real-life tasks in teaching, more so than their knowledge base.

Abstract

THEORETICAL AND REVIEW ARTICLES

July 08, 2026 PDT

Psychological and Structural Determinants of Silo Mentality: A Multi-level Scoping Review

Ana Fitriani, Seger Handoyo, Berlian Gressy Septarini

The study identifies factors contributing to silo mentality in organizations, including defensive attitudes, limited emotional awareness, hierarchical structures, and competitive incentives. It suggests interventions should target multiple levels simultaneously.

Abstract

RESEARCH ARTICLES

July 03, 2026 PDT

Assessing School-Related Depression in Adolescents: Psychometric Validation of a Brief Screening Scale and Its Complementarity With the DASS-D in Vietnam

Hong Ta, Duc Dao

This study validates the School Depression Screening Scale (SDSS) as a reliable tool for identifying depression in Vietnamese students, complementing the DASS-21 depression subscale.

Abstract

Validating A Two-Factor School Discipline Measure for Adolescents: An Unrestricted Factor Analysis in Indonesian Samples

Yohanes Budiarto & Bonar Hutapea
Universitas Tarumanagara

This study examined the dimensional structure of the School-DDI (an adaptation of the Dimensions of Discipline Inventory) for adolescents in Indonesian school contexts, comparing two exploratory models: the bifactor model and the traditional two-factor model. Data were obtained from 1,300 secondary students who responded to 23 items (after excluding three items with severe floor effects: V5, V9, V10) on a 5-point scale assessing perceptions of punitive and non-punitive disciplinary practices. Exploratory analyses were conducted using FACTOR (Ferrando & Lorenzo-Seva, 2017; Lorenzo-Seva & Ferrando, 2006) with polychoric correlations, MORGANA extraction, and Robust Promin rotation. Optimal Parallel Analysis supported a two-dimensional structure. The bifactor model showed excellent fit (CFI = .999; RMSEA = .016; BIC = 905.45) and revealed a dominant general discipline factor (ECV = .690; UniCo = .756), alongside two specific factors representing punitive-specific and non-punitive-specific disciplinary styles; Factor Determinacy Indices (FDI) were high (GF = .903, F1 = .891, F2 = .900). The two-factor model also showed excellent fit (CFI = .993; RMSEA = .034; BIC = 1010.03) and yielded a simpler, interpretable structure with a positive inter-factor correlation (r = .235); FDI values were exceptionally high (F1 = .927, F2 = .967), supporting individual-level diagnostic applications. McDonald's omega reliability was ω = .913. External validation with academic procrastination (M = 2.41, SD = 0.96) showed divergent patterns across models: in the bifactor solution, the specific punitive factor (F1) correlated positively with procrastination (r = .170, $p < .05$), the specific non-punitive factor (F2) correlated negatively (r = –.095, $p < .05$), and the general factor (GF) showed no significant association (r = .031, p = .326); in the two-factor solution, the non-punitive factor (F1) showed a non-significant negative correlation (r = –.055, p = .062) and the punitive factor (F2) correlated positively (r = .135, $p < .05$). All fit statistics are reported after the LOSEFER empirical correction (Lorenzo-Seva & Ferrando, 2023). Findings highlight that school discipline reflects both a unified perception of climate and two distinct operational styles. A hybrid scoring approach is recommended to

Author info: Correspondence should be sent to: Dr. Yohanes Budiarto, Faculty of Psychology, Universitas Tarumanagara, Jakarta, Indonesia.
Email: yohanesb@fpsi.untar.ac.id

North American Journal of Psychology, 2026, Vol. 28, No. 3, 1602-1618
 https://doi.org/10.65696/001c.165924

preserve both general and specific variance for comprehensive evaluation of school discipline practices.

Keywords: school discipline, two-factor model, bifactor model, unrestricted factor analysis, academic procrastination, Indonesian adolescents

INTRODUCTION

School discipline is a system of rules, monitoring, and corrective measures designed to foster good behaviour or change inappropriate behavior (Way, 2019). It also involves strategies for managing student behaviour, including self-management and classroom management practices (Smith et al., 2022). Discipline is necessary to maintain an orderly environment and is conducive to both study and students' moral and educational development (Ching & Jaffri, 2015). Disciplinary actions can range from minor reprimands to exclusion from school, with varying implications for students' educational experiences (Welsh, 2022).

Research has shown that school discipline significantly impacts student behaviour and academic achievement, particularly in junior high school (Awiria, 2021; Saputri et al., 2025; Via & Padang, 2021). The use of structured discipline strategies can make students' experiences more orderly, resulting in better behaviour and achievement (Ijaz et al., 2024; Njoroge & Nyabuto, 2014). Research indicates that positive perceptions of school rules correlate with disciplined behavior, while negative perceptions can link to resistance and alienation (Indriyani & Putri, 2024; Sheets, 2002).

Studies in Indonesia show that school type moderates discipline practices. Students in private junior high schools experience stricter discipline than those in state schools (Asyhar & Hastuti, 2023). Punitive approaches have been widely implemented but do not appear as effective as prevention-oriented approaches in reducing aggression (Gaffney et al., 2021). Punitive school discipline is more strongly linked to adverse health outcomes (e.g., persistent depressive symptoms) than to improved behavioural outcomes (Duarte et al., 2023). In the Indonesian context, rewards are more effective in improving students' motivation and positive behaviour, while punishments, though moderately effective, should be applied judiciously to avoid adverse psychological effects (Asmira et al., 2023).

Rationale for Adapting the DDI to School Contexts

To date, no instrument comprehensively measures Indonesian adolescents' perceptions of school disciplinary practices in terms of punitive versus non-punitive dimensions. Furthermore, to our knowledge, no prior Indonesian study has systematically developed and validated such a measure at the item level. Existing school discipline instruments in Asian and Western contexts, such as the Authoritative School Climate Survey (Konold et al., 2014) in the United States, the Delaware School Climate Survey–Student (Bear et al., 2011), and the Behavior and Instructional Management Scale (Martin & Sass, 2010) capture global school climate or classroom management orientations rather than a fine-grained set of discrete disciplinary strategies. Systematic meta-reviews confirm substantial heterogeneity in school climate measurement and the absence

of a validated instrument covering specific punitive and non-punitive school discipline practices across cultural contexts (Chow et al., 2024; Gonzálvez et al., 2023).

The Dimensions of Discipline Inventory (DDI; Straus & Fauchier, 2007) was originally designed to capture the range of parental discipline practices, which are broadly summarised as punitive and non-punitive factors. Although parent-child and teacher-student relationships differ in their authority structures, emotional contexts, and institutional constraints, the DDI's theoretical framework is amenable to school-based adaptation. Both contexts involve adult authority over child behaviour through structured rules and consequences. The underlying style continuum, from power-assertive to inductive and supportive practices, is conceptually equivalent across both settings. The adaptation involved replacing "parents" with "teachers/school staff" and "family rules" with "school or classroom rules," without altering the latent disciplinary dimensions. Nevertheless, because this change in referent and context does not guarantee structural equivalence, the adapted instrument's factor structure must be empirically evaluated rather than assumed.

A critical step in cross-referent adaptation involves linguistic and conceptual equivalence. The present adaptation followed a forward-backward translation procedure, with semantic equivalence testing and cognitive interviews involving school counsellors, teachers, and students. Content validity was established through expert review panels and think-aloud interviews to refine wording and comprehension prior to data collection.

Research Objectives and Rationale for the Analytical Strategy

The present study had three primary objectives: (1) to evaluate the factor structure of the School-DDI in a large Indonesian adolescent sample using Unrestricted Factor Analysis; (2) to compare a two-factor correlated model with a bifactor model; and (3) to examine the external validity of the resulting factor scores using academic procrastination as a criterion.

Academic procrastination was selected as the external criterion because it represents a behaviorally anchored, theoretically plausible outcome of disciplinary climate. Self-determination theory (Deci & Ryan, 2000; Ryan & Deci, 2000) predicts that structured, clear, and autonomy-supportive environments reduce motivational deficits, with procrastination as a prototypical indicator. Coercive or unpredictable disciplinary environments may increase it. Future research could complement this with direct measures of behavioral infractions or school connectedness.

METHOD

We conducted a cross-sectional, multi-site psychometric study. The analytical strategy centred on Unrestricted Factor Analysis (Steenkamp & Maydeu-Olivares, 2023), implemented in the FACTOR program (Ferrando & Lorenzo-Seva, 2017; Lorenzo-Seva & Ferrando, 2006, 2013).

The choice of UFA over a conventional EFA-CFA two-stage sequence requires explicit justification. In the traditional approach, EFA is used to identify the factor structure, and CFA is then applied, often to the same sample, with fixed zero cross-loadings. This is problematic for three reasons. First, when the population model is genuinely multidimensional with non-negligible cross-loadings, forcing items to load on only one factor introduces specification errors that bias factor correlations and

structural parameter estimates. Simulation studies show that cross-loadings as small as .20 can substantially bias these estimates in CFA (Steenkamp & Maydeu-Olivares, 2023). Second, fitting a restrictive CFA to the same data used for EFA capitalises on chance. Third, for newly adapted instruments, the exact item-to-factor mapping cannot be specified a priori with certainty (Lorenzo-Seva & Ferrando, 2020).

UFA imposes only the minimum constraints required for identification, without fixing non-target loadings to zero. Any restricted (CFA) model is nested within the corresponding UFA model, making UFA fit an upper bound to CFA fit (Steenkamp & Maydeu-Olivares, 2023). All 1,300 participants were analysed in a single UFA stage; this is not a two-step EFA–CFA design. Structural robustness was assessed via the Solomon split-half method (Lorenzo-Seva, 2022), which split the sample into two equivalent halves (Ratio Communality Index = .989). In addition to the two-factor UFA, a pure exploratory bifactor model (Lorenzo-Seva & Ferrando, 2019a) was estimated as an alternative structural hypothesis, with both models using MORGANA extraction, an algorithm that minimizes matrix residuals without distributional assumptions, designed specifically for non-normal ordinal data (Lorenzo-Seva & Ferrando, 2019b).

Participants and Setting

The sample consisted of 1,300 junior high school students recruited across multiple sites in Indonesia. Participants ranged from 11 to 14 years old ($M = 13.25$, $SD = 1.07$). Regarding sex, 43.6% ($n = 567$) were male and 56.4% ($n = 733$) were female. The sample encompassed students from both state and private schools across different regions. Future publications should report the full breakdown by school type and region to allow evaluation of the generalizability of the discipline factor.

Instrumentation

School-DDI The original DDI domains (Straus & Fauchier, 2007) were adapted to reference teachers in the school context. After forward–backward translation and cognitive interviews, twenty-six items were retained on a 5-point scale (1 = Never to 5 = Very Often) covering the past 1–3 months. Three items (V5, V9, V10) were subsequently excluded due to severe floor effects (see Results), leaving 23 items for all factor-analytic procedures.

Academic Procrastination Scale Academic procrastination was measured using a validated self-report scale on a 5-point frequency scale ($\alpha = .87$; $\omega = .88$; $M = 2.41$, $SD = 0.96$).

RESULTS

All fit statistics reported below are based on the LOSEFER empirical correction (Lorenzo-Seva & Ferrando, 2023), an empirical rescaling procedure that adjusts chi-square-based fit indices for inflation due to non-normality and large sample sizes, yielding more trustworthy CFI and RMSEA estimates under ordinal data conditions. All 1,300 participants were included in a single UFA framework.

Pre-Factor Analysis

Three items (V5, V9, V10) were excluded due to severe floor effects. Almost all of the 1,300 participants selected the lowest response option (Never), leaving essentially no variance for differentiating disciplinary

practices. After exclusion, 23 items remained. Item distributions were clearly non-normal (multivariate kurtosis = 743.39, $p < .001$), justifying polychoric correlations and MORGANA with robust corrections. KMO = .933 (95% CI [.912, .935]), and Bartlett's test was highly significant ($\chi^2(253) = 14{,}857.4$, $p < .001$).

The Solomon split-half (RCI = .989) confirmed internal stability. Optimal Parallel Analysis (Timmerman & Lorenzo-Seva, 2011) supported retaining two factors: the first two eigenvalues (40.90% and 19.11% of variance) substantially exceeded the 95th percentile of random data (9.95% and 9.09%), while the third (5.37%) fell below the threshold (8.51%).

Closeness to Unidimensionality

Unidimensionality indices indicated that a strictly one-factor solution would be too reductive. Explained Common Variance (ECV) = .690 (95% CI [.670, .714]) and Unidimensional Congruence (UniCo) = .756 (95% CI [.724, .789]) both fell well below the thresholds recommended for essential unidimensionality (ECV > .85 and UniCo > .95; Ferrando & Lorenzo-Seva, 2018), and The Mean of Item Residual Absolute Loadings (MIREAL = .327) exceeded the recommended threshold of < .30. These values confirm the presence of multidimensionality and motivated the comparison of two-factor and bifactor solutions.

Two-Factor Model Results

Goodness of Fit

After LOSEFER correction the fit indexes are: RMSEA = .034 (95% CI [.032, .034]), CFI = .993 (95% CI [.992, .994]), NNFI = .991, GFI = .994, AGFI = .993, BIC = 1010.03. These robust fit statistics confirm that the two-factor structure replicated the polychoric correlation matrix with exceptional precision even under non-normal data conditions.

Factor Structure

In the rotated pattern matrix, Factor 1 (non-punitive discipline) was anchored by items reflecting supportive and autonomy-oriented practices: N1 (λ = .658), N4 (λ = .598), N7 (λ = .672), C19_1 (λ = .714), N10 (λ = .597), N11 (λ = .654), N12 (λ = .731). Factor 2 (punitive discipline) was defined by items reflecting coercive or restrictive practices: P8 (λ = .833), P10 (λ = .815), P12 (λ = .811), P9 (λ = .789), P1 (λ = .723), P11 (λ = .742). The inter-factor correlation was positive (r = .235), suggesting that students who perceive more punitive discipline also tend to perceive more non-punitive discipline, consistent with a general discipline climate interpretation.

However, several items showed mixed or unexpected loadings. Items N2 (.247 on F1, .590 on F2), N3 (-.292 on F1, .542 on F2), and N9 (.193 on F1, .630 on F2) loaded more strongly on the punitive factor than on the non-punitive factor, despite being conceptually designed to measure supportive practices.

In the structure matrix, these items showed cross-loadings: N2 (.385 on F1, .648 on F2), N3 (-.164 on F1, .474 on F2), and N9 (.341 on F1, .675 on F2) (see Table 1). This pattern suggests that these items may blend "control with care" or "firmness with fairness". This conceptualization may be particularly salient in the Indonesian cultural context, where clear expectations and structured guidance are valued as forms of care. These items appear to blend supportive intent with control and pressure,

potentially reflecting culturally embedded notions of 'firmness with care' in hierarchical Indonesian school contexts.

Based on these loading patterns, items N2 and N3 are recommended for rewording to explicitly distinguish relational warmth from structural control, while N9 may be retained with a theoretical note as a culturally embedded item. It needs further qualitative examination of the interpretation of items in the Indonesian school context. A full item-level mapping of the original DDI domains to the adapted School-DDI structure is available as downloadable supplementary material via figshare: https://doi.org/10.6084/m9.figshare.32833760

Factor Score Quality

Factor-score quality was exceptional. The Factor Determinacy Index (FDI) values were $FDI_1 = .927$ (95% CI [.924, .937]) and $FDI_2 = .967$ (95% CI [.961, .970]). ORION (Overall Reliability of Informative prior Oblique N-EAP scores) reliability estimates were $ORION_1 = .860$ and $ORION_2 = .934$. Both FDI values exceeded the recommended .90 threshold (Ferrando & Lorenzo-Seva, 2018), indicating that the School-DDI factor scores are precise enough not only for group-level research but also for individual-level diagnostic use by school counsellors. They may identify individual students who perceive disproportionately punitive discipline for targeted intervention. The Sensitivity Ratios ($SR_1 = 2.474$, $SR_2 = 3.773$) and Expected Percentages of True Differences ($EPTD_1 = 90.7\%$, $EPTD_2 = 94.3\%$) confirm that individual differences on these dimensions can be reliably differentiated.

Reliability

McDonald's ordinal omega (ω) = .913, greatest lower bound (GLB) to reliability = .968, and standardised Cronbach's alpha = .906, confirming excellent scale-level reliability.

External Validity

In the two-factor model, the non-punitive factor (F1) showed a small, non-significant negative correlation with academic procrastination (r = -.055, 95% CI [-.110, .003], p = .062), while the punitive factor (F2) showed a statistically significant positive correlation (r = .135, 95% CI [.082, .196], $p < .05$). These effect sizes are small in practical terms (< 2% of variance), and causal conclusions are not warranted from this cross-sectional correlational design.

Bifactor Model Results

Goodness of Fit

The bifactor model demonstrated excellent fit after the LOSEFER correction: RMSEA = .016 (95% CI [.015, .017]), CFI = .999 (95% CI [.998, .999]), NNFI = .998, GFI = .996, AGFI = .995, and BIC = 905.45. The lower BIC relative to the two-factor model (905.45 vs. 1010.03) indicates that the bifactor model provides a more parsimonious account of the data.

Factor Structure

In the bifactor rotated pattern matrix, Factor 1 (F1, punitive) was defined by P2 ($\lambda = .658$), P4 ($\lambda = .651$), P7 ($\lambda = .639$), C3_1 ($\lambda = .586$), P8 ($\lambda = .574$), P12 ($\lambda = .532$), P9 ($\lambda = .511$), P1 ($\lambda = .516$), P11 ($\lambda = .476$). Factor 2 (F2, non-punitive-specific) was defined by N1 ($\lambda = .705$), C19_1 ($\lambda = .683$), N12 ($\lambda = .658$), N7 ($\lambda = .572$), N4 ($\lambda = .466$), N10 ($\lambda = .437$),

N11 (λ = .379). The General Factor (GF) received meaningful loadings across all 23 items, with particularly high values for P10 (λ = .694), N6 (λ = .634), P8 (λ = .627), N9 (λ = .572), and N10 (λ = .584). The inter-factor correlation between the two specific factors was negative (r = -.286), indicating that after removing general discipline variance, punitive and non-punitive tendencies are negatively related, as theoretically expected.

Once the general factor was partialled out, the three cross-loading items showed clarified roles: N2 (F1-specific = .507, GF = .483), N3 (F1-specific = .382, GF = .207), and N9 (F1-specific = .427, GF = .572). Specifically, N9 emerged primarily as a general-discipline item with only modest punitive-specific loading, whereas N2 and N3 carried moderate specific-punitive loadings alongside substantial general-factor contributions (see Table 1). This confirms that their cross-loading in the two-factor solution reflected shared general-discipline variance rather than true punitive content. The mathematical shift in loadings across the two solutions provides concrete evidence that bifactor modelling disambiguates item meaning by separating general-climate from style-specific variance.

Unidimensionality in Bifactor Context

Within the bifactor solution, overall UniCo = .578 (95% CI [.561, .597]) and ECV = .434 (95% CI [.408, .447]), both well below the essential-unidimensionality thresholds (ECV > .85; UniCo > .95; (Ferrando & Lorenzo-Seva, 2018). This confirms that even after modelling a general factor, the two specific factors account for substantial additional variance and cannot be ignored.

Factor Score Quality

FDI values in the bifactor solution were: FDI(GF) = .903 (95% CI [.893, .917]), FDI(F1) = .891 (95% CI [.882, .905]), FDI(F2) = .900 (95% CI [.889, .912]). All three exceeded the .90 threshold (Ferrando & Lorenzo-Seva, 2018), indicating that even the bifactor-specific-factor scores are sufficiently precise for individual-level assessment—for example, a school counsellor could use separate general and specific scores to distinguish students who perceive their school as broadly high-control (high GF) from those who specifically experience disproportionate punishment (high F1-specific).

ORION estimates were ORION(GF) = .815, ORION(F1) = .794, ORION(F2) = .810. The Generalized H index values (H-Latent: F1 = .792, F2 = .826, GF = .817) indicate well-defined latent variables likely to be stable across studies.

External Validity

In the bifactor model, the punitive-specific factor (F1) showed a statistically significant positive association with academic procrastination (r = .170, 95% CI [.117, .221], p < .05), the non-punitive-specific factor (F2) showed a statistically significant negative association (r = -.095, 95% CI [-.146, -.035], p < .05), and the general factor (GF) showed no significant association (r = .031, 95% CI [-.033, .091], p = .326).

These effect sizes are small, and causal claims are not warranted from this cross-sectional correlational design. The associations suggest that a specific disciplinary style, rather than general discipline intensity, is the primary predictor of procrastination tendencies, though replication using longitudinal and experimental designs is needed to establish directionality.

Table 1 Factor Loadings for the Two-Factor and Bifactor Solutions (Rotated Pattern Matrix, Post-LOSEFER Correction)

Item	TF-F1 (NP)	TF-F2 (P)	BF-F1 (P-sp)	BF-F2 (NP-sp)	BF-GF
N1	.658	-.041	.205	.705	.173
N2*	.247	.590	.507	.261	.483
N3*	-.292	.542	.382	-.246	.207
N4	.598	.040	.005	.466	.396
N5	.461	.271	.144	.336	.498
N6	.337	.250	-.120	.094	.634
N7	.672	-.267	-.157	.572	.183
N8	-.158	.611	.346	-.213	.412
N9*	.193	.630	.427	.128	.572
N10	.597	.213	.062	.437	.584
N11	.654	.084	-.098	.379	.579
N12	.731	-.064	.037	.658	.320
C19_1	.714	-.042	.125	.683	.262
P1	.032	.723	.516	.007	.524
C3_1	.111	.634	.586	.177	.426
P2	.064	.629	.658	.196	.329
P4	.064	.609	.651	.193	.330
P7	.035	.665	.639	.138	.367
P8	.053	.833	.574	.010	.627
P9	-.166	.789	.511	-.190	.495
P10	-.017	.815	.413	-.143	.694
P11	-.062	.742	.476	-.099	.515
P12	-.046	.811	.532	-.081	.567

Note. TF = Two-Factor; BF = Bifactor; NP = Non-Punitive; P = Punitive; P-sp = Punitive-specific; NP-sp = Non-Punitive-specific; GF = General Factor. * Items with cross-loading patterns discussed in text. All values derived from the FACTOR program output after LOSEFER correction.

DISCUSSION

The present study compared two psychometric models for the School-DDI among 1,300 Indonesian junior high school students: a two-factor correlated model and a bifactor model. Both models were estimated via UFA in a single analytical stage without a two-step EFA-CFA sequence. Both models yielded psychometrically sound solutions, yet they offer complementary accounts of how discipline is perceived in school settings.

The Two-Factor Model: Practical Utility

The two-factor solution performed exceptionally well (RMSEA = .034, CFI = .993, NNFI = .991, BIC = 1010.03). Factor loadings were strong and substantively interpretable: punitive items (e.g., P8 = .833, P10 = .815) anchored one dimension, while non-punitive items (e.g., N7 = .672, N12 = .731, C19_1 = .714) anchored the other. The positive inter-factor correlation (r = .235) suggests a shared general-climate component.

Particularly noteworthy are the FDI values (FDI(F1) = .927 and FDI(F2) = .967), both exceeding the .90 threshold recommended for individual-level assessment (Ferrando & Lorenzo-Seva, 2018). This means the School-DDI factor scores are precise enough for individual-level diagnostic applications by school practitioners. Two concrete

examples illustrate this practical value: (a) School climate assessment, administrators can compute school-level means on both subscales to profile their institution's disciplinary approach and evaluate changes over time following policy interventions; and (b) Intervention targeting, school counsellors can use individual-level scores to identify students with particularly negative perceptions of punitive discipline, directing them toward restorative or PBIS-based supports and tracking changes in perception following intervention.

The Bifactor Model: Theoretical Depth

The bifactor model provided a superior fit (RMSEA = .016, CFI = .999, BIC = 905.45) and revealed structural features that were invisible in the two-factor solution. The general discipline factor (ECV = .690, UniCo = .756) represented a shared perception of the school's overall disciplinary intensity, regardless of specific method. The UFA-based bifactor approach avoids the zero-loading constraints of CFA, which would have forced cross-loading items into arbitrary assignments and inflated inter-factor correlations (Ferrando & Lorenzo-Seva, 2000; Steenkamp & Maydeu-Olivares, 2023).

The bifactor framework clarified the cross-loadings of N2, N3, and N9. These items carry meaningful general-factor loadings (GF: .483, .207, and .572, respectively), indicating they primarily reflect shared discipline-climate variance rather than a true punitive disciplinary style. In Indonesian educational contexts, where high power distance and cultural expectations of obedience are normative (Afifi et al., 2025), students may interpret structurally supportive practices through a lens of authority and control. Future item revisions should: explicitly distinguish relational warmth from structural control in items N2 and N3; consider retaining N9 for theoretical reasons while flagging it as a culturally embedded item; and develop additional items targeting teacher–student communication quality, perceived fairness, and emotional safety to strengthen the non-punitive factor's discriminant validity.

The divergent external validity patterns between models further demonstrate the value of the UFA–bifactor approach. In the two-factor solution, the punitive factor (F2, which absorbed general-climate variance) showed a positive association with procrastination (r = .135), whereas the non-punitive factor showed no significant effect (r = -.055). In the bifactor solution, once general-discipline variance was partialled out, the punitive-specific factor showed r = .170, the non-punitive-specific factor r = -.095, and the general factor was unrelated to procrastination (r = .031). This illustrates how aggregating specific-style variance into the punitive factor in the two-factor model attenuated the non-punitive factor's predictive signal. These correlational associations are small in magnitude and should be interpreted with caution: disciplinary style may be related to procrastination tendencies but does not appear to be a dominant predictor, and the causal directionality cannot be inferred from the cross-sectional design.

Reconciling Model Comparison: Statistical Evidence

The BIC difference between the two competing structures (104.58 for the bifactor) should not be taken as a trivial observation of relative magnitude, but rather should be investigated. In line with common rules of thumb for Bayesian model comparison, a BIC difference > 10 is

interpreted as very strong evidence for the lower-BIC model (Raftery, 1995), a pattern supported in independent applications to structural equation modeling and general Bayesian model selection (Kass & Raftery, 1995). From this perspective, the current ΔBIC value is significant enough to warrant a strong case for a bifactor structure, which has not been the case.

Rather than setting aside BIC evidence, the question needs to be understood in the context of a broader warning from the bifactor literature. Given that general factor loadings may be relatively low, and, above all, that fit indices do not measure the relative strength of the general factor and the specific factors, fit-index comparisons tend to favor bifactor specifications. This pattern is not unique to the current data set; one would expect to see it in many data sets if the complexities of the data were unmodeled and understood to cause a bias in favour of the bifactor model over other multidimensional structures, as shown by standard fit statistics.

Researchers who have employed bifactor models have tended to focus on fit superiority and have neglected other complementary measures, such as omega reliability, factor determinacy, explained common variance, and the percentage of uncontaminated correlations, which are necessary for a full psychometric interpretation of the data (Pretorius, 2021). If this is considered in the context of the present results, the bifactor structure is better supported than the BIC value alone suggests. The general factor's ECV (.690) and UniCo (.756) are lower than the thresholds that are traditionally considered as essential unidimensionality (ECV > .85; UniCo > .95; Ferrando & Lorenzo-Seva, 2018), which are the same thresholds the present manuscript calls elsewhere. This general factor is strong enough to be modeled separately, but not strong enough to warrant specific factors. This condition exemplifies a conflict between statistical fit indices and substantive interpretability.

Although the two-factor model demonstrates superior statistical fit, the substantial magnitude of the general factor complicates the model selection. Specifically, these statistical criteria alone cannot determine whether the construct is better operationalized as a single unidimensional trait with specific sub-components or as two distinct yet correlated dimensions. The hybrid scoring recommendation does not dodge a conclusion. It simply adjusts for a known issue: when a general factor is present but not dominant, specific-factor scores lose their statistical value. By accounting for this, the approach creates a much cleaner, more accurate bifactor model.

Cultural Context and Generalizability

Indonesia's high power-distance culture may amplify the perceived intensity of both punitive and non-punitive practices and produce a stronger general discipline factor than would be expected in Western samples. The pattern of cross-loading items (N2, N3, N9) blending control with care is consistent with non-Western conceptualizations of benevolent authority. Whether the bifactor or two-factor structure generalizes to other Asian educational systems (e.g., Japan, South Korea, China) or to Western contexts remains an empirical question for cross-cultural replication studies that use measurement invariance testing.

Limitations of the Study

Several limitations should be noted. First, the cross-sectional design precludes causal inference; longitudinal designs are needed to establish whether disciplinary climate precedes or follows changes in academic motivation. Second, the sample breakdown by school type (public vs. private), socioeconomic status, and region was not fully reported, limiting the evaluation of the discipline factor's generalizability. Third, academic procrastination was the sole external criterion; future studies should include additional outcomes (e.g., school connectedness, direct behavioral infractions, academic achievement) for a more comprehensive test of external validity. Fourth, although forward-backward translation was conducted, more detailed reporting of evidence for semantic equivalence would strengthen confidence in the cross-cultural applicability.

Future item revision cycles should: (a) reword N2 and N3 to more clearly distinguish relational warmth from structural control; (b) consider retaining N9 with theoretical documentation as a culturally embedded item; (c) develop additional items targeting teacher–student communication quality, perceived fairness, and emotional safety to strengthen the non-punitive factor's coherence; and (d) test the adapted scale in diverse school types and regions to evaluate cross-group invariance using multigroup confirmatory approaches.

Recommendations for Future Research

This study investigated the internal structure of the School-DDI using Unrestricted Factor Analysis in the FACTOR program, comparing bifactor and two-factor models with all 1,300 participants in a single analytical stage. This avoids the specification errors introduced by a conventional EFA-CFA sequence. UFA was chosen because it allows the full loading matrix to be estimated empirically, detects cross-loadings that would be forced to zero in CFA, and provides fit statistics that serve as an upper bound for any restricted CFA model (Steenkamp & Maydeu-Olivares, 2023). All fit statistics were evaluated after the LOSEFER empirical correction to ensure trustworthy inference under non-normal ordinal data conditions.

The hybrid nature of the identified structure brings a number of questions to the fore that need to be explored in future, methodologically different ways. Causal inferences about the relationship between disciplinary climate and student outcomes in the present study are limited by the cross-sectional design used. Further research is required to determine the causal relationships between perceived changes in disciplinary style and academic motivation and self-regulation, and whether student characteristics affect perceived changes in disciplinary style. Similarly, this type of longitudinal study would allow for the investigation of developmental pathways throughout the secondary school years, during which students' ability to engage in self-regulated learning is developing.

Second, several non-punitive dimensions loaded on the punitive dimension in the two-factor solution suggest that the conceptual separation between supportive control and coercive control may be culturally specific. Qualitative research, such as cognitive interviewing, focus group discussions with students, and classroom observations, would help to shed light on the interpretation of specific disciplinary practices by Indonesian

adolescents, as well as the observed cross-loadings as a result of either genuine conceptual blending or inadequate item wording. This qualitative work needs to guide a systematic revision of items with a focus on new items that differentiate between relational warmth and structural control and, of course, the retention of items that reflect culturally embedded ideas of benevolent authority.

Third, this study examined only one external criterion. Future studies should test the School-DDI's relationship with a broader range of student outcomes, including school connectedness, direct behavioral violations, academic outcomes, and psychological outcomes. Multiple criteria would help resolve whether the general discipline climate factor predicts overall adjustment, and whether specific punitive and non-punitive factors predict differentiated adjustment. Furthermore, interventions conducted with the School-DDI as a pre-post measure would yield useful information about the scale's sensitivity to change and its value in assessing school-based prevention interventions.

Fourth, the sample included students from public and private schools across several areas, but the data did not provide specific breakdowns by important subgroups to draw conclusions about structural equivalence among these subgroups. Future research should explore bifactor and two-factor structures in a multigroup invariance test to determine whether there are significant differences among students across school types, socioeconomic backgrounds, and geographic areas. These analyses are crucial prior to recommending the scale for comparative research or for policy evaluation.

Fifth, it is essential to replicate across cultures. This Indonesian sample seemed to show a greater general discipline factor, which may be due to cultural factors of high power distance and respect for authority. Comparative analysis with other educational settings in Western and Eastern Asia, as well as in other Southeast Asian cultures, would show whether it is a universal or culturally specific pattern in the perception of school discipline. A cross-cultural analysis of measurement invariance would help to determine if the School-DDI is a cross-nationally comparable instrument. Last but not least, an exploratory or "free-ranging" approach was used, which is suitable for newly adapted instruments but which does not necessarily ensure that the identified structure will also be found in independent samples. The factor loadings and item mappings reported here are the target model for a confirmatory factor analysis on an independent sample of students for which the structural validity would be more rigorously tested. Further, test-retest reliability must be determined to provide the scale with the possibility of longitudinal monitoring and evaluation of intervention.

REFERENCES

Afifi, Z., Setiaji, B., & Waston. (2025). Power Distance in Indonesian Schools: A Study of Public General Schools and Pesantren. *Journal of Cultural Analysis and Social Change*, *10*(2). https://doi.org/10.64753/jcasc.v10i2.2063

Asmira, Herlina, S., Mustika, S., & Sulaeman. (2023). The Effectiveness of Educational Punishment in Improving Student Discipline: A Case Study in South Sulawesi. *Didaktika*, *17*(2), 137–145.

Asyhar, A. N., & Hastuti, D. (2023). Differences in Discipline Character Between State and Private Junior High School Students: Analysis of Family and School Discipline Practices, and Self-concept. *Journal of Child, Family, and Consumer Studies*, *2*(3), 179–189. https://doi.org/10.29244/jcfcs.2.3.179-189

Awiria, A. (2021). Hubungan Antara Kedisiplinan Belajar Siswa dengan Prestasi Belajar Siswa Mata Pelajaran Ilmu Pengetahuan Sosial Kelas V Sekolah Dasar. *TERAMPIL: Jurnal Pendidikan Dan Pembelajaran Dasar*, *8*(1). https://doi.org/10.24042/terampil.v8i1.8153

Bear, G. G., Gaskins, C., Blank, J., & Chen, F. F. (2011). Delaware School Climate Survey-Student: Its factor structure, concurrent validity, and reliability. *Journal of School Psychology*, *49*(2). https://doi.org/10.1016/j.jsp.2011.01.001

Ching, L., & Jaffri, H. (2015). Developing Discipline among Students through Social-Emotional Learning: A New Model to Prevent and Reduce Behavior Problems. *Journal of Education and Vocational Research*, *6*, 80–90. https://api.semanticscholar.org/CorpusId:151785511

Chow, J. C., Sayers, R., Fu, Y., Granger, K. L., McCullough, S., Kingsbery, C., & Morse, A. (2024). A Systematic Meta-Review of Measures of Classroom Management in School Settings. *Assessment for Effective Intervention*, *49*(2), 60–74. https://doi.org/10.1177/15345084231208671

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4). https://doi.org/10.1207/S15327965PLI1104_01

Duarte, C. d. P., Moses, C., Brown, M., Kajeepeta, S., Prins, S. J., Scott, J., & Mujahid, M. S. (2023). Punitive school discipline as a mechanism of structural marginalization with implications for health inequity: A systematic review of quantitative studies in the health and social sciences literature. In *Annals of the New York Academy of Sciences* (Vol. 1519, Number 1). https://doi.org/10.1111/nyas.14922

Ferrando, P. J., & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: some aspects of the problem and some suggestions. *Psicologica*, *21*, 301–323. https://api.semanticscholar.org/CorpusID:14830623

Ferrando, P. J., & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema*, *2*(29). https://doi.org/10.7334/psicothema2016.304

Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis. *Educational and Psychological Measurement*, *78*(5). https://doi.org/10.1177/0013164417719308

Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2021). Effectiveness of school-based programs to reduce bullying perpetration and victimization: An updated systematic review and meta-analysis. In *Campbell Systematic Reviews* (Vol. 17, Number 2). https://doi.org/10.1002/cl2.1143

Gonzálvez, C., Bacon, V., & Kearney, C. A. (2023). Systematic and evaluative review of school climate instruments for students, teachers, and parents. *Psychology in the Schools*, *60*(6), 1781–1836. https://doi.org/10.1002/pits.22838

Ijaz, S., Nobles, J., Mamluk, L., Dawson, S., Curran, B., Pryor, R., Redwood, S., & Savović, J. (2024). Disciplinary behaviour management strategies

in schools and their impact on student psychosocial outcomes: A systematic review. *NIHR Open Research*, *4*. https://api.semanticscholar.org/CorpusId:268710153

Indriyani, A., & Putri, D. R. (2024). Hubungan Persepsi Siswa Tentang Pelaksanaan Tata Tertib Sekolah Dengan Sikap Disiplin Siswa SMA Muhammadiyah 1 Surakarta. *Khatulistiwa: Jurnal Pendidikan Dan Sosial Humaniora*, *4*(1), 29–38. https://doi.org/10.55606/khatulistiwa.v4i1.2703

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Konold, T., Cornell, D., Huang, F., Meyer, P., Lacey, A., Nekvasil, E., Heilbrun, A., & Shukla, K. (2014). Multilevel multi-informant structure of the authoritative school climate survey. *School Psychology Quarterly*, *29*(3). https://doi.org/10.1037/spq0000062

Lorenzo-Seva, U. (2022). SOLOMON: a method for splitting a sample into equivalent subsamples in factor analysis. *Behavior Research Methods*, *54*(6). https://doi.org/10.3758/s13428-021-01750-y

Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, *38*(1). https://doi.org/10.3758/BF03192753

Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR 9.2: A Comprehensive Program for Fitting Exploratory and Semiconfirmatory Factor Analysis and IRT Models. *Applied Psychological Measurement*, *37*(6). https://doi.org/10.1177/0146621613487794

Lorenzo-Seva, U., & Ferrando, P. J. (2019a). A General Approach for Fitting Pure Exploratory Bifactor Models. In *Multivariate Behavioral Research* (Vol. 54, Number 1). https://doi.org/10.1080/00273171.2018.1484339

Lorenzo-Seva, U., & Ferrando, P. J. (2019b). Robust Promin: A method for diagonally weighted factor rotation. *Liberabit: Revista Peruana de Psicología*, *25*(1). https://doi.org/10.24265/liberabit.2019.v25n1.08

Lorenzo-Seva, U., & Ferrando, P. J. (2020). Unrestricted factor analysis of multidimensional test items based on an objectively refined target matrix. *Behavior Research Methods*, *52*(1). https://doi.org/10.3758/s13428-019-01209-1

Lorenzo-Seva, U., & Ferrando, P. J. (2023). LOSEFER: A robust correction for the chi-square goodness-of-fit statistic in factor analysis. *Structural Equation Modeling*. https://doi.org/10.1080/10705511.2023.2171462

Martin, N. K., & Sass, D. A. (2010). Construct validation of the behavior and instructional management scale. *Teaching and Teacher Education*, *26*(5). https://doi.org/10.1016/j.tate.2009.12.001

Njoroge, P. M., & Nyabuto, A. N. (2014). Discipline as a Factor in Academic Performance in Kenya. *Journal of Educational and Social Research*, *4*, 289. https://api.semanticscholar.org/CorpusId:143326313

Pretorius, T. B. (2021). Over reliance on model fit indices in confirmatory factor analyses may lead to incorrect inferences about bifactor models: A cautionary note. *African Journal of Psychological Assessment*, *3*. https://doi.org/10.4102/ajopa.v3i0.35

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*. https://doi.org/10.2307/271063

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1). https://doi.org/10.1037/0003-066X.55.1.68

Saputri, R. E., Setiowati, I., Kamila, N. A., Hikma, N., & Aulia, R. (2025). Hubungan Antara Kedisiplinan Dan Prestasi Belajar Siswa Sekolah Dasar. *Jurnal Pengajaran Sekolah Dasar*, *4*(1), 225–236. https://doi.org/10.56855/jpsd.v4i1.1678

Sheets, R. H. (2002). "You're Just a Kid That's There"-Chicano Perception of Disciplinary Events. *Journal of Latinos and Education*, *1*(2), 105–122. https://doi.org/10.1207/S1532771XJLE0102_3

Smith, T. E., Thompson, A. M., & Maynard, B. R. (2022). Self-management interventions for reducing challenging behaviors among school-age students: A systematic review. *Campbell Systematic Reviews*, *18*(1). https://doi.org/10.1002/cl2.1223

Steenkamp, J. B. E. M., & Maydeu-Olivares, A. (2023). Unrestricted factor analysis: A powerful alternative to confirmatory factor analysis. *Journal of the Academy of Marketing Science*, *51*(1). https://doi.org/10.1007/s11747-022-00888-1

Straus, M. A., & Fauchier, A. (2007). *Manual for the Dimensions of Discipline Inventory (DDI)*. Family Research Laboratory, University of New Hampshire. https://doi.org/https://doi.org/10.1037/t37478-000

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*(2). https://doi.org/10.1037/a0023353

Via, I., & Padang, A. T. (2021). Pentingnya Tata Tertib Untuk Membentuk Disiplin Belajar Siswa SMP [The Importance of Rules to Form Student Learning Discipline in Junior High School]. *Kumpulan Artikel Ilmiah Rumpun Ekonomi Dan Ilmu Sosial*, *1*(1), 78–94. https://doi.org/10.19166/kairos.v1i1.3164

Way, S. (2019). School Discipline. In *The Blackwell Encyclopedia of Sociology* (pp. 1–5). Wiley. https://doi.org/10.1002/9781405165518.wbeoss022.pub2

Welsh, R. O. (2022). Schooling Levels and School Discipline: Examining the Variation in Disciplinary Infractions and Consequences Across Elementary, Middle, and High Schools. *Journal of Education for Students Placed at Risk*, *27*(3). https://doi.org/10.1080/10824669.2022.2041998