THE ROLE OF SAME UNIT COORDINATION IN THE DEVELOPMENT OF ADVANCED MULTIPLICATIVE REASONING SCHEMES: A LONGITUDINAL STUDY

Rui Ding, Boyuan Xue, Xixi Deng, Jiajia Li

Northeast Normal University

Same-Unit Coordination (SUC) has been proposed as a necessary requisite for Mixed-Unit Coordination (MUC) and Place-Value, Base-10 reasoning (PV-B10), yet longitudinal evidence is limited. We followed 225 students in Grades 3-4 across three waves and employed Latent Class Growth Modeling to identify the role of SUC on the advanced multiplicative schemes. Two latent classes (high-decreasing vs. low-increasing) were identified based on their developmental patterns on SUC. Mixed-design repeated-measures ANOVAs showed significant Class × Time interactions for both MUC and PV-B10: the high-decreasing class declined on both schemes, whereas the low-increasing class improved. Results suggest SUC as an adaptor in the development of advanced multiplicative reasoning schemes.

INTRODUCTION

This study employs a longitudinal design to investigate elementary students' developmental trajectories in the Same-Unit Coordination (SUC) scheme and its dynamic role in the advanced multiplicative reasoning (MR) schemes, specifically, Mixed Unit Coordination (MUC) and Place Value, Base 10 (PV-B10).

Prior cross-sectional studies have established two key findings: first, a puzzling grade-reversal phenomenon, where SUC performance appears to decline across grade levels (Ding et al., 2019); second, SUC is a necessary but insufficient condition for advanced MR schemes (Tzur et al., 2018). These findings raise critical questions that cross-sectional data cannot resolve: Is the reversal an artifact of measurement or a real developmental pattern? How does the growth (or decline) in SUC predict and facilitate the subsequent development of MUC and PV-B10 over time?

To address these questions, we analyze longitudinal assessment data to: (1) map the true developmental trends of SUC across grades, moving beyond static snapshots; and (2) examine how within-student changes in SUC relate to the development of advanced MR schemes. And this approach moves beyond association to provide mechanistic insight into the developmental function of a foundational scheme.

CONCEPTUAL FRAMEWORK

We adopt a constructivist perspective in which a scheme functions as a cognitive apparatus for assimilation: a goal-setting recognition template, the mental operations enacted to achieve that goal, and the anticipated result (Piaget, 1985; von Glasersfeld,

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 347-354. PME.

1995). From a units-and-operations stance, schemes are distinguished by the kinds of units a learner operates on and the operations coordinated across unit levels (Steffe, 1992; Ulrich, 2016). Given our research foci, we only focused on three schemes in multiplicative reasoning identified by Tzur et al. (2013), including SUC, MUC and PV-B10.

SUC involves operating additively on equal-size composite units while maintaining awareness that each composite unit is composed of 1s; critically, the intended result is expressed in numbers of composite units rather than totals of 1s. Consider the task: *A has 7 boxes of mooncakes, each box containing 5 mooncakes. B has 8 boxes of mooncakes, each box also containing 5 mooncakes. How many boxes do A and B have altogether? How many boxes does B have more than A?* Students' goal is to figure out the sum or differences of the composite units (boxes).

MUC requires coordinating a mix of unit types (typically a compilation of composite units together with a quantity of 1s that can be reorganized into those composite units), thereby intentionally sequencing operations that reorganize 1s into composite units and then unite them with an existing compilation. Consider the task: *K makes 7 wraps, each containing 4 Chinese dates. K will place 48 more Chinese dates in wraps of 4 dates each. How many wraps would K then have in all?* A child assimilating this task into the MUC scheme recognizes a situation involving (a) one compilation of same-size composite units and (b) a composite unit comprised of 1s that can be organized into the same-size composite units. The global goal is to determine how many composite units (wraps) would comprise the encompassing compilation.

Finally, **PV-B10** is treated as a special case of MUC in which unit coordination is constrained by powers of ten and governed by place-value conventions (e.g., coordinating 1s, 10s, and 100s to compare or combine quantities). Consider the task: *Schools keep apples in boxes, bags, or singletons, with 1 box = 10 bags of apples, and 1 bag = 10 apples. School A has 2 boxes, 4 bags, and 19 apples; School B has 1 box, 16 bag, and 11 apples. Which school has more apples and by how many?* Having constructed MUC as explained above, a child assimilating this task into the PV-B10 scheme recognizes a situation involving two encompassing, numerical units comprised of a mix of levels of units. At the least, such a numerical mix of units involves a coordination of 1s and a compilation of one or more 10s.

Within the developmental progression of MR, SUC is considered a foundational scheme for MUC and PV-B10 (Tzur et al., 2013). Building on this perspective, we conducted a longitudinal design to examine how SUC functions in the development of MUC and PV-B10 over time.

METHODS

Data collection and participants

We collected the data from grades 3-4 students in two elementary schools (A and B) in Northeast China. A longitudinal design was adopted with 3 measurement waves: T1

(Fall 2024), T2 (Spring 2025) and T3 (Fall 2025), with approximately one semester between adjacent waves. At each wave, the same cohort of participants completed MR assessment under same consistent time limits, administration procedures and scoring criteria to ensure the comparability of data across time points. Ultimately, a total of 225 participants completed all three assessment waves from **school A** (N = 144) and **school B** (N = 81), including 84 3rd graders (37.3%) and 141 4th graders (62.7%).

Measures and procedure

Measures were drawn from the Chinese version of the MR assessment (including five MR schemes), a 22-item instrument designed for students in Grades 3-6. This Chinese MR assessment was developed through expert translation and review, and its content and structural validity have been examined in prior work. The original English MR assessment was developed in the U.S. through adapting tasks from teaching experiments that grounded the MR developmental model (Tzur et al., 2013; Tzur & Xin, 2022). Rather than administering the full 22-item instrument at each time point, we only focused on the 10 items involving SUC, MUC, and PV-B10 schemes.

We assembled three test forms administered at T1, T2, and T3. The T1 administration used a subset of items from the 22-item pool. At T2, we administered a form that included anchor (common) items shared with T1, whereas at T3 we administered a form that shared a different set of anchor items with T2 (i.e., the anchor items linking T2 and T3 differed from those linking T1 and T2). This planned adjacent-wave overlap enabled test linking across the three administrations and supported longitudinal comparability.

Data analysis

Data analysis was conducted using multiple statistical software packages for different analytic purposes. Firstly, Winsteps 3.27 was used to examine the psychometric properties of the MR assessment administered at T1, including reliability, discrimination, unidimentionality, and item fit index. Test forms across waves were linked using an anchor-item design within the Rasch model via separate calibrations with fixed anchoring. Anchor-item difficulty parameters were first estimated in a reference wave and then fixed to these values when calibrating the other waves, placing item parameters and person measures on a common logit scale. Based on the linked metric, we extracted students' person measures (i.e., ability estimates) for each scheme at T1, T2, and T3 to enable direct longitudinal comparisons. Second, Mplus 8.0 was used to fit a latent class growth model (LCGM) to identify heterogeneous trajectories in students' performance on the SUC scheme over time. Finally, SPSS 27.0 was used to conduct mixed-design repeated-measures ANOVAs to compare mean students' ability for the MUC and PV-B10 schemes across the three time points and across the growth-pattern groups identified by the LCGM.

Psychometric properties of assessment

Firstly, we examined the reliability, separation, and item-fit properties of the MR assessment. According to Fisher (2007), the item reliability of 0.98 was excellent (> 0.94) and the person reliability of 0.76 was fair (> 0.67). The item separation (8.98) was high and the person separation (1.79) was acceptable (Boone et al., 2014). Evidence for unidimensionality was also supported, the eigenvalue of the first contrast in the unexplained variance was 1.6 (< 3.0) (Linacre, 2025). Item difficulty estimates were associated with relatively small standard errors (0.16-0.44). The Infit MNSQ for all the items fitted quite well with the expected model (0.78-1.25), while the Outfit MNSQ for three items (two for mDC and one for PV-B10) had underfit (2.04, 2.48, and 1.66, respectively) (Linacre, 2002). In general, most of the indicators show that all items fit well with the expected model, so the MR assessment can be used to investigate MR development among students in Grade 3-4.

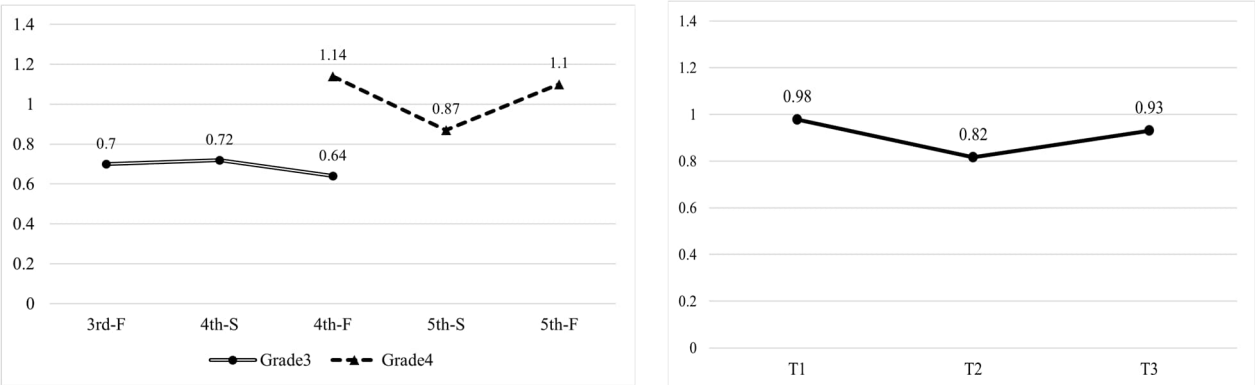
RESULTS

Firstly, we explored the development of students' ability on the SUC scheme. Then, we conducted a LCGM to characterize heterogeneity in students' longitudinal change on SUC and to determine the optimal number of latent classes. Finally, we evaluated whether the resulting SUC-based classes differentiated students' longitudinal patterns on the MUC and PV-B10 schemes.

The development of elementary students on the SUC scheme

Figure 1 depicts students' development of the mean ability on SUC. Especially, students' ability decreased slightly from Grade 4 to Grade 5 (i.e., from 4th-F to 5th-F; see Figure 1 (a)).

Figure 1: Students' development of mean ability on SUC.



(a) Third and fourth graders (b) All students

The overall mean ability for all students on SUC at the three time points was 0.98 (SD = 1.46), 0.82 (SD = 1.55), and 0.93 (SD = 1.36), respectively (see Figure (b)). Specifically, the mean ability decreased from T1 to T2 and rebounded from T2 to T3. However, the mean ability at T3 remained slightly lower than at T1, suggesting a

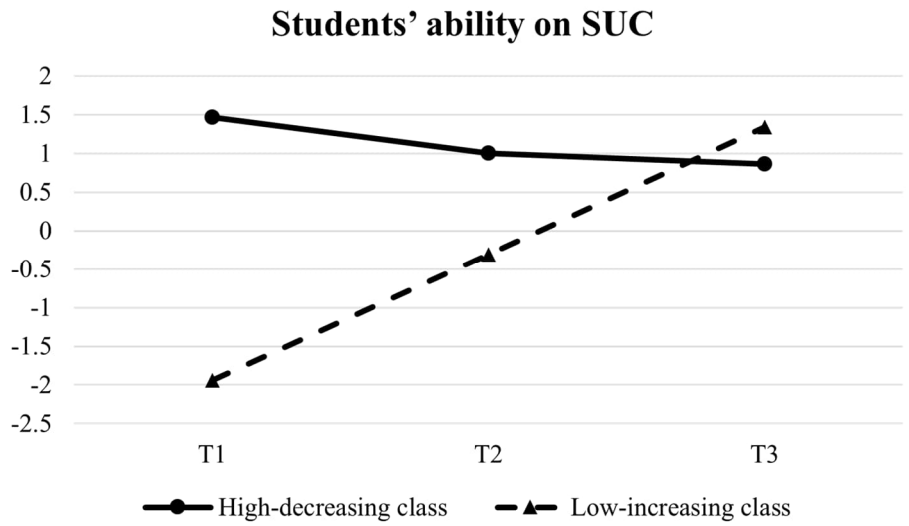
modest overall decreased trend. The coefficients of variation (CV = SD / Mean) were 1.49, 1.89, and 1.46, respectively. Given the substantial between-student variability, we next conducted LCGM to identify latent classes with distinct SUC developmental trajectories over time.

Latent growth classes identified on the SUC scheme

Then, the LCGM was performed to identify the latent trajectory classes. The fit indices collectively supported the 2-class model as the optimal solution. Specifically, the 2-class model showed lower AIC, BIC, and aBIC values than competing models, and both the Lo-Mendell-Rubin test and Bootstrap Likelihood Ratio Test were significant ($p < .001$). Moreover, the classification quality was excellent (Entropy = .97), indicating clear separation between classes (Lubke & Muthén, 2007). Considering statistical fit, parsimony, and theoretical interpretability, we retained the 2-class model as the most appropriate representation of the heterogeneity in students' developmental trajectories. The larger class comprised 85.8% of the sample (N = 193; 68 3rd Graders and 125 4th Graders) and the smaller class comprised 14.2% (N = 32; 16 3rd Graders and 16 4th Graders) of the sample.

Two latent growth classes based on SUC (abbr. SUC-based classes) are visualized in Figure 2. The majority of students (85.8%) were classified into Class 1, characterized by a relatively high initial level (intercept = 1.77, $p < .001$) followed by a significant decline over time (slope = −0.32, $p < .001$). This latent growth class was thus labeled as the **high-decreasing** class. In contrast, Class 2 comprised 14.2% of the sample and showed a low initial level (intercept = −3.58, $p < .001$) but a significant increase over time (slope = 1.64, $p < .001$), thus was identified as the **low-increasing** class.

Figure 2: Comparison of ability growth trajectories between two SUC-based classes.



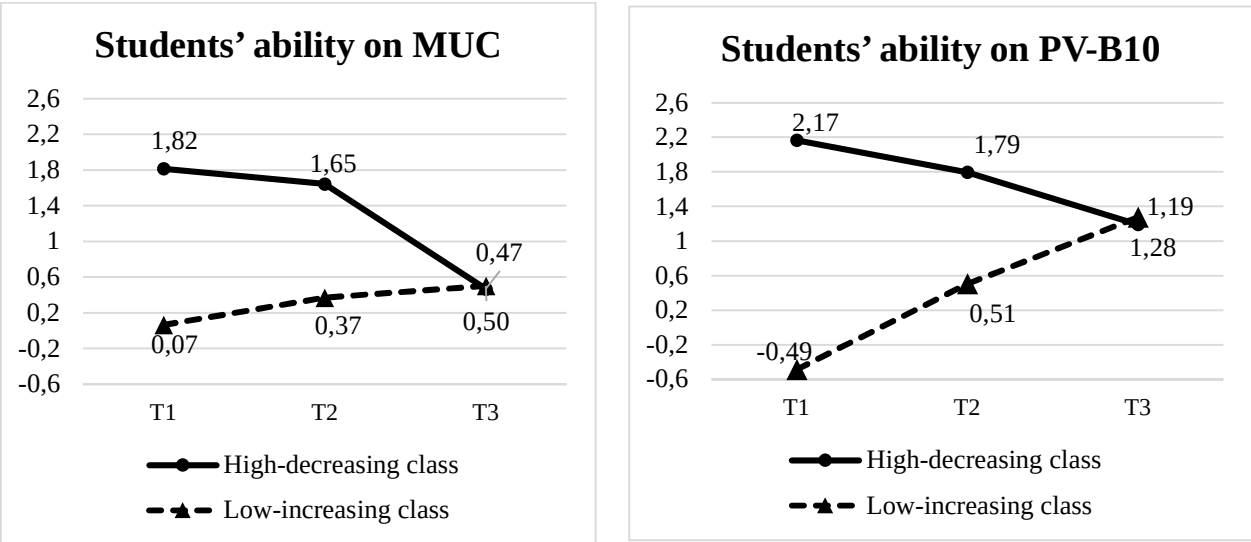
Class-by-time effects on the MUC and PV-B10 schemes

After identifying the SUC-based classes, we examined whether these two classes also exhibited distinct growth patterns on MUC and PV-B10. Therefore, we conducted separate 2 × 3 (Class × Time) mixed-design repeated-measures ANOVAs for MUC and PV-B10. Figure 2 presents the estimation of students' ability on MUC and PV-B10 across T1-T3 for the two SUC-based classes.

The Class × Time interaction in students' ability on MUC was significant ($F_{(2, 446)}$ = 12.09, p < .001, η_p^2 = .051), indicating that the two classes exhibited different developmental patterns on MUC. Descriptively, high-decreasing class showed a marked decline from T1 to T3, whereas low-increasing class showed an increasing trend across the three time points. Bonferroni-adjusted pairwise comparisons showed that, in the high-decreasing class, the change in MUC from T1 to T2 was not significant, whereas decreased significantly from T2 to T3 (p < .001). For the low-increasing class, the person ability showed an increasing trend over time, but this change was not statistically significant.

As for students' ability on PV-B10, the results also showed a significant Class × Time interaction ($F_{(2,446)}$ = 28.10, p < .001, η_p^2 = .112), indicating different developmental patterns between the two SUC-based classes. Estimated marginal means showed that the high-decreasing class declined across time (both significant from T1 to T2, and from T2 to T3, p < .05), while the low-increasing class increased (significantly from T1 to T2, p < .001, whereas not significant from T2 to T3).

Figure 3: Comparison of ability growth trajectories on the MUC (left) and PV-B10 (right) schemes between two SUC-based classes.



Overall, the Class × Time interactions indicated differential developmental trajectories across classes on the same scheme.

DISCUSSION

Our findings provide crucial insight into students' longitudinal development on SUC and its role in the advanced MR schemes.

Firstly, a concerning developmental trend emerged: a majority of students exhibited declines not only in SUC but also in advanced MR schemes. This finding partially aligns with the grade-reversal phenomenon noted by Ding et al. (2019), but it contradicts the assumed grade-level progression commonly inferred from cross-sectional studies. We posit that this decline may point to a significant instructional gap, where everyday classroom curricula and teaching may overlook or inadequately reinforce the conceptual coordination of units and operations, particularly the construction of composite units. This highlights a potential disconnect between hypothesized learning trajectories and actual student development.

Moreover, SUC appears to function as a pivotal "developmental adaptor", conditioning the acquisition of advanced MR schemes, like MUC and PV-B10. This role is clearly demonstrated by the divergent longitudinal paths of the two SUC-based classes. Despite a substantial initial performance in MUC and PV-B10, students in the low-increasing SUC class demonstrated accelerated growth in these advanced schemes, nearly closing the gap with their peers in the high-decreasing SUC class by the upper grades. This dynamic "catch-up" pattern, likely driven by the consolidation of SUC, offers a mechanistic explanation for why Grade 4 (ages 9-10) is often identified as a critical juncture in mathematical development (Lin, 2018): it may represent a window during which students with initially lower but improving foundational unit coordination unlock rapid advancement in more complex multiplicative reasoning.

References

Bond, T. G., Yan, Z., & Moritz, H. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences* (4th ed.). Routledge.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.

Ding, R., Tzur, R., Wei, B., Jin, X., Jin, X., Davis, A. (2019). Conceptual Analysis of an Error Pattern in Chinese Elementary Students. In M. Graven, H. Venkat, A. Essien & P. Vale (Eds.), *Proc 43rd Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp.93-100). PME.

Fisher, W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions*, *21*(1), 1095.

Geiser, C. (2020). *Longitudinal structural equation modeling with Mplus: A latent state-trait perspective*. Guilford Press.

Lin, C. D. (2018). *Developmental psychology*. Beijing: People's Education Press. [In Chinese].

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 878.

Linacre, J. M. (2025). *A user's guide to WINSTEPS: Rasch-model computer programs*. Winsteps.

Lubke, G., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(1), 26-47.

Moss, J., & McNab, S. L. (2011). An approach to geometric and numeric patterning that fosters second grade students' reasoning and generalizing about functions and covariation. In J. Cai, & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives* (pp. 277-302). Berlin: Springer-Verlag.

Norton, A., Boyce, S., Phillips, N., Anwyll, T., Ulrich, C., & Wilkins, J. (2015). A written instrument for assessing students' units coordination structures. *Journal on Mathematics Education*, *10*(2), 111-136.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). The University of Chicago Press.

Steffe, L. P. (1992). Schemes of action and operation involving composite units. *Learning and Individual Differences*, *4*(3), 259-309.

Tzur, R., Wei, B., Smith, A., Norton, A., Davis, A., & Johnson, H. L. (2018). Same unit coordination: A conceptual screener for mixed unit coordination and base-10, place value reasoning. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proc. 42nd Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 323-330). PME.

Tzur, R., Johnson, H. L., McClintock, E., Kenney, R. H., Xin, Y. P., Si, L., . . . Jin, X. (2013). Distinguishing schemes and tasks in children's development of multiplicative reasoning. *PNA-Pensamiento Numérico Avanzado*, *7*(3), 85-101.

Tzur, R., & Xin, Y. P. (2022). Nurturing multiplicative reasoning with whole numbers. In Y. P. Xin, R. Tzur, & H. Thouless (Eds.), *Enabling mathematics learning of struggling students* (1st ed., pp. 291-314). Springer - Nature.

Ulrich, C. (2016). Stages in constructing and coordinating units additively and multiplicatively (part 2). *For the Learning of Mathematics*, *36*(1), 34-39.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Falmer.