Educational Psychology Review (2026) 38:39
https://doi.org/10.1007/s10648-026-10131-w

REPLICATION

I Feel Competent, Therefore I Am: Self-concept and Skill Interact at Different Speeds

Fernando Núñez-Regueiro[1] · Herbert W. Marsh[2,3] · Pascal Bressoux[1] · Anatolia Batruch[4] · Marinette Bouet[5] · Marco Bressan[6] · Genavee Brown[7] · Fabrizio Butera[4] · Anthony Cherbonnier[7] · Céline Darnon[8] · Marie Demolliens[8] · Anne-Laure de Place[1] · Olivier Desrichard[9] · Luc Goron[7] · Brivael Hémon[7] · Pascal Huguet[8] · Eric Jamet[7] · Vincent Mazenod[5] · Nathalie Mella[9] · Estelle Michinov[7] · Nicolas Michinov[7] · Nana Ofosu[9] · Laurine Peter[7] · Céline Poletti[6] · Isabelle Régner[6] · Mathilde Riant[1] · Anaïs Robert[8] · Ocyna Rudmann[4] · Camille Sanrey[1] · Arnaud Stanczak[8] · Farouk Toumani[5] · Emilio Paolo Visintin[4] · Pascal Pansu[1]

Received: 16 September 2025 / Accepted: 10 March 2026

Abstract
Do perceptions about one's competence shape learning, or are they simply reflections of actual skills? This study revisits this longstanding question by replicating and extending preliminary findings by Marsh et al. (2024) on the temporal dynamics linking students' academic self-concept (i.e., their perceived academic competence) and their academic skills. Using longitudinal data from a large-scale field study ($N > 9000$ students, 3 measurement points), we tested how academic self-concept and skills relate to each other over time. Consistent with Marsh et al., results revealed a consistent temporal asymmetry: Academic skills predicted concurrent changes in self-concept within the same semester (contemporaneous effects), whereas self-concept predicted changes in academic skills across semesters (lagged effects). These findings were robust to several stress tests, including measurement error, unmeasured confounding, and competing models of change. Together, the results are consistent with a renewed theory of learning behavior, in which perceived competence and skills influence each other at different speeds. This temporal asymmetry helps integrate short-term and long-term cognitive-motivational processes in theories of learning behavior. It also underscores the importance of aligning intervention strategies and model specifications with the timescales of the underlying psychological processes, with implications for both fundamental and intervention research.

Extended author information available on the last page of the article

Published online: 11 April 2026

Keywords Academic self-concept · Learning behavior · Reciprocal effects model · Timescales · Structural equation modeling

Introduction

What does it truly mean to be competent? The philosopher René Descartes famously declared, "*Cogito*, *ergo sum*"—"I think, therefore I am"—arguing that thought itself is the surest proof of existence. In the learning sciences and developmental psychology, a parallel idea has emerged: "I feel competent, therefore I am (competent)". This view suggests that a learner's perceived competence can be the foundation for building skills, and that fostering this perception should be a priority for parents and educators. An opposing view, however, reverses the direction: It is the growth of skills that fuels the feeling of competence—"I am competent, therefore I feel competent"—and thus improving performance should come first. Preliminary findings by Marsh et al. (2024) suggest that both views are true, but that they might operate on different timescales. Their study among German students found indeed that changes in skills acquisition in mathematics predicted short-term changes in mathematics self-concept (within a semester), whereas changes in self-concept predicted long-term changes in skills acquisition (on the following semester).

In this study, we replicate and extend the findings by Marsh et al. (2024) using a large-scale dataset in the French context ($N>9{,}000$ students) and longitudinal data in two domains, namely mathematics and French (language art). We also develop a theoretical framework that explicitly accounts for fast versus slow dynamics of change. In this framework, perceived competence can shift quickly in response to recent performance ("I am competent, therefore I feel competent"); but it can also, in turn, slowly shape skill development over time ("I feel competent, therefore I become competent"). This temporal perspective provides a basis for unifying theories operating on different timescales and for calibrating research designs more effectively as a function of fast and slow dynamics of change.

How Feelings of Competence Relate to Skills Acquisitions: Past and Present Perspectives

Historically, reciprocal effects between feelings of competence and skills acquisitions have been discussed for decades in the form of a motivational process known as self-enhancement theory, and a cognitive process known as skill development theory (Calsyn & Kenny, 1977; Wesarg-Menzel et al., 2023). Self-enhancement theory posits that perceived competence—also referred to as academic self-concept—can promote skill acquisition by increasing motivation to learn. Competence-related beliefs are known to support greater effort, engagement, and persistence; extensive research documents these motivational pathways within self-concept and related frameworks (Marsh, 1990; Marsh & Craven, 2006; Valentine et al., 2004). Research on self-efficacy similarly shows that capability beliefs predict task engagement and persistence (Bandura, 1982; Pajares & Schunk, 2001; Shum & Fryer, 2026). In this view, self-perceptions are not merely reflections of performance, but active contributors to learning. Echoing Descartes's *cogito ergo sum*, this account holds: *I feel competent*, *therefore I become competent*. The implication is that educators should support learners' self-perceptions—especially when confidence is fragile—by highlighting personal progress, emphasizing success, and offering encouragement that promotes positive self-concept in a domain (Calsyn & Kenny, 1977; Valentine et al., 2004).

By contrast, skill development theory holds that self-perceptions are effects, not causes, of learning. In this view, learners form perceptions about their competence based on their actual performance in a learning domain: Success fosters confidence, while failure erodes it. Echoing the reverse dictum, this account holds: *I am competent*, *therefore I feel competent*. The implication is that educators should focus less on the motivational processes related to how learners feel about themselves, and more on the cognitive processes that effectively support learning, notably by designing instruction that provides structured guidance for novices and minimally guided instruction for more advanced learners (Hattie & Timperley, 2007; Hunt, 1975; Kirschner et al., 2006; Tobias & Duffy, 2009).

Although initially opposed, these two perspectives are now widely regarded as complementary. The Reciprocal Effects Model (REM) formalized this integration, positing bidirectional influences between academic self-concept and academic achievement (Marsh, 1990; Marsh & Craven, 2006). In this framework, perceived competence contributes to skill development (*self-enhancement effect*), while skill development reinforces perceived competence (*skill development effect*). Many contemporary theories of school motivation and learning have since incorporated the REM (Núñez-Regueiro et al., 2022; Valentine et al., 2004), and longitudinal meta-analyses consistently report positive, reciprocal effects between self-concept and skills (Huang, 2011; Marsh & Craven, 2006; Valentine et al., 2004; Vu et al., 2024).

Despite its popularity, the REM remains debated, notably regarding the causal interpretation of these reciprocal effects and its conceptualization of how developmental processes unfold. First, most REM studies rely on analytic models, such as the cross-lagged panel models (CLPMs), that do not distinguish between long-term evolutions in self-concept and skills (baseline levels or "traits") and short-term evolutions (momentary fluctuations or "states") (Hamaker et al., 2015; Usami et al., 2019). This confounding can lead to spurious findings. For example, if trait levels correlate positively (e.g., individuals with lower skills consistently reporting lower self-concept), the CLPM will represent this trait-level correlation in the guise of prospective effects between states, creating the illusion of reciprocal causality even when state-level effects are actually null or reversed (Hamaker et al., 2015; Lucas, 2023; Núñez-Regueiro et al., 2022). To address this issue, researchers have adopted more refined analytic models that separate trait and state components (for details, see Supplemental Material A; Núñez-Regueiro et al., 2025; Usami et al., 2019). For example, the random intercept CLPM (RI-CLPM) (Hamaker et al., 2015) posits that traits are stable dispositions shaped by background factors (e.g., socioeconomic status, gender, early learning environments), which are separated from dynamic states. Studies showed that REM effects, particularly the effects of self-concept on skills, tended to disappear when estimated using these modeling approaches, suggesting they reflect state-trait confounding rather than causal effects (Burns et al., 2020; Ehm et al., 2019; Hübner et al., 2023; Núñez-Regueiro et al., 2022).

Second, previous statistical models used to test the REM assume that reciprocal effects unfold over a fixed time lag, typically from one measurement occasion to the next (e.g., year T to year T+1). However, recent advances in longitudinal modeling showed that if reciprocal effects operate over shorter time intervals (e.g., weeks or months) than the interval between measurements (e.g., semesters or years), they will appear contemporaneous, i.e., occurring within the same occasion (i.e., without time lag) (Muthén & Asparouhov, 2024; Speyer et al., 2024). In this case, forcing a lagged structure onto the data can distort findings. This means that previous tests of the REM (positive or negative) could potentially be misleading. It also opens a theoretical opportunity: By contrasting lagged and contemporaneous effects, empirical models can help clarify the timescales at which self-concept and skills influence each other and embrace more precise temporal interpretations of learning behavior (Bailey et al., 2024; Murayama & von Keyserlingk, 2025). Recent work by Marsh and colleagues points in this direction (Marsh et al., 2024). Using repeated-measures data, they found that skill effects on self-concept were contemporaneous (skills at year T predicted self-concept at year T), whereas self-concept effects on skills were lagged (self-concept at year T predicted skills at year T+1).

In the present study, we build on the findings reported by Marsh et al. (2024) while extending their work in three ways. First, Marsh et al.'s study was mainly exploratory and did not provide a theoretical rationale for why self-concept and skills should interact at different speeds, that is, through contemporaneous versus lagged effects. Here, we advance such rationale by situating the REM within a timescales continuum framework that formalizes varying speeds in developmental processes (see next section). Second, Marsh et al. relied on a single domain (mathematics), a single population of German secondary students, and measures specific to the PALMA framework (i.e., original six items rated on a 5-point Likert scale for self-concept; standardized tests in arithmetic, algebra, geometry and fractions for skills) (Pekrun et al., 2007). By using the ProFAN dataset, we offer a conceptual replication that examines the fast-slow dynamics of the REM in a different educational context (i.e., French vocational students) and with alternative measures of self-concept and skills in both mathematics and French (language arts). A conceptual replication has the advantage of testing whether the findings by Marsh et al. (2024) reflect the underlying theoretical processes and constructs, rather than particular features of the PALMA design (Fabrigar & Wegener, 2016; Núñez-Regueiro & Wang, 2024).

Third, we extend the replication using a comparative approach that evaluates the reference statistical model in Marsh et al. (2024), i.e., the RI-CLPM-C, against competing models (RI-CLPM, RI-PCM, LCSM, RI-LTVC) that make different claims about the timescales and causal strength of reciprocal effects (for details, see SM-A; Núñez-Regueiro et al., 2025). More specifically, models that disaggregate trait and state components but differ in their temporal assumptions (i.e., RI-CLPM and RI-PCM) enable testing whether reciprocal effects operate only across semesters or only within semesters. LCSMs assume that trait-like differences emerge from accumulated state changes and therefore test whether trait-state entanglement provides a better approximation of co-development. Finally, STARTS and RI-LTVC variants assess whether apparent reciprocal effects between self-concept and skills could instead arise from measurement error or latent third-variable confounders. This comparative modeling framework thus allows to determine whether the temporal asymmetry identified through the reference model RI-CLPM-C by the CL-REM remains the most plausible account once these competing explanations are formally tested.

This combination of theoretical grounding, conceptual replication, and empirical extensions provide a more systematic framework for further evaluating the results in Marsh et al. (2024) and for refining our interpretation of learning dynamics at school.

Theorizing the Contemporaneous-Lagged Reciprocal Effects Model (CL-REM): A Timescales Continuum Approach

To clarify how academic self-concept and skills interact at different speeds, we propose a refinement of the REM: the Contemporaneous-Lagged Reciprocal Effects Model (CL-REM). This refined theory draws on the timescales continuum framework (Hamaker, 2023), which formalizes how psychological processes unfold at different speeds; some processes thus evolve rapidly and appear transient (microlevel), others accumulate more slowly (mesolevel), and others exhibit long-term stability (macrolevel). These orders of magnitude are defined relative to the temporal context of a study. For example, weekly variations in a process will appear stable in a daily design, but transient in a year-long study (Hamaker, 2023). Timescales must therefore be conceptualized in relation to the developmental window under investigation. In the REM context, self-concept and skills are conceptualized over multi-year developmental periods, such that this multiyear period represents the macrolevel timescale of stable factors, while mesolevel and microlevel scales capture evolutions expected to operate at slower (semestrial or yearly) or faster (weekly or monthly) speeds, respectively. Accordingly, three levels (micro-, meso-, macrolevel timescales) are distinguished in the CL-REM.

Contemporaneous skill development effect (microlevel timescale) According to the CL-REM framework, the contemporaneous effect of academic skills on academic self-concept reflects fast changes operating at the microlevel timescale, which includes month-to-month or week-to-week variations. For example, an improvement in mathematics performance during an examination (e.g., in week T) may increase mathematics self-concept in the following week (e.g., week T+1) by instilling feelings of pride or teacher praise shortly after the event. This mechanism is consistent with control-value theory (Pekrun et al. 2023; Pekrun 2024a), which posits that proximal feedback indicating success can quickly enhance students' sense of competence by eliciting positive achievement-related emotions. It is also consistent with teacher feedback models (Hattie & Timperley, 2007; Mandouit & Hattie, 2023) which emphasize that teacher interventions occurring directly after, or shortly following, a test or performance—such as affirming comments or targeted guidance—can strengthen students' self-concept by clarifying goals and highlighting progress. Because these processes operate on short timescales, the influence of skills on self-concept is likely to appear contemporaneous in longitudinal models with widely spaced measurements (e.g., yearly or semestrial).

Lagged self-enhancement effect (mesolevel timescale) By contrast, the lagged effect of academic self-concept on skills corresponds to slower changes operating at the mesolevel timescale, typically over semesters or school years. Thus, an enhanced mathematics self-concept may gradually increase motivation or engagement in class, ultimately fostering acquisitions in mathematics. Indeed, engagement is understood as a multidimensional process encompassing behavioural (e.g., sustained effort), emotional (e.g., enjoyment or interest), and cognitive (e.g., strategic attention) components, that typically requires extended periods to shift in a coordinated way (Fredricks et al., 2004; Skinner et al., 2008). It also aligns with the process of student motivation, characterized by the gradual internalization (or, conversely, externalization) of learning goals and values, reinforced through repeated rewarding experiences or sustained teacher autonomy support (Bardach & Murayama, 2025; León et al., 2025; Ryan & Deci, 2020). Because such motivational processes unfold over longer periods, the influence of self-concept on subsequent skill acquisition is more likely to emerge with a temporal lag in longitudinal analyses.

Covariance of stable traits (macrolevel timescale) At the macrolevel timescale, spanning years or decades, these reciprocal dynamics are expected to stabilize into enduring trait-like associations between academic self-concept and skills. This covariance does not imply active causal processes occurring at the trait level, but instead reflects the cumulative impact of prior skill development and self-enhancement effects over time. Such traits may evolve over years, but their changes occur on a different—typically much longer—timescale than microlevel or mesolevel dynamics. In practice, the macrolevel covariance is estimated through random intercepts, which requires at least three measurement occasions to capture the stable component of each construct with sufficient precision.

In sum, the CL-REM distinguishes processes and effects that unfold at different speeds. Importantly, the detectability of these effects depends on how measurement intervals align with their timescales. When intervals are very short (e.g., microlevel of weeks), fast effects may appear lagged (instead of contemporaneous), whereas slow processes may not surface at all. Similarly, long intervals can make slow processes appear contemporaneous (instead of lagged) while obliterating the variance necessary to detect fast processes. It is therefore expected that the structure specified by the CL-REM is best examined in mesolevel designs that employ yearly or multi-month intervals, enabling the detection of both skill development effects (as contemporaneous effects) and self-enhancement effects (as lagged effects). Consistent with the CL-REM, Marsh et al.'s (2024) yearly design proved capable of identifying both contemporaneous skill-development effects and lagged self-enhancement effects. A shorter mesolevel interval of comparable magnitude, such as the 8-month intervals of the ProFAN design, can similarly be expected to capture these temporal dynamics, as we will test in this study.

Present Study

This study examines how the relation between academic self-concept (i.e., perceived competence) and academic skills unfolds across different timescales. Guided by the CL-REM framework (Fig. 1) and preliminary findings from Marsh et al. (2024), we hypothesize that academic skills produce short-term changes in self-concept (i.e., contemporaneous effect, time lag ~ 0–2 months, i.e., within a few weeks or months), whereas the influence of self-concept on skills emerges more slowly through long-term changes (lagged effect, time lag ~ 8 months). These effects are expected to occur at the level of short-term states—momentary fluctuations in self-concept and skills—rather than at the level of stable traits, which are shaped by past learning histories and broader contextual factors. This study therefore focuses on testing the temporal structure implied by the CL-REM, leaving the examination of potential mediating processes underlying these reciprocal effects to future research.

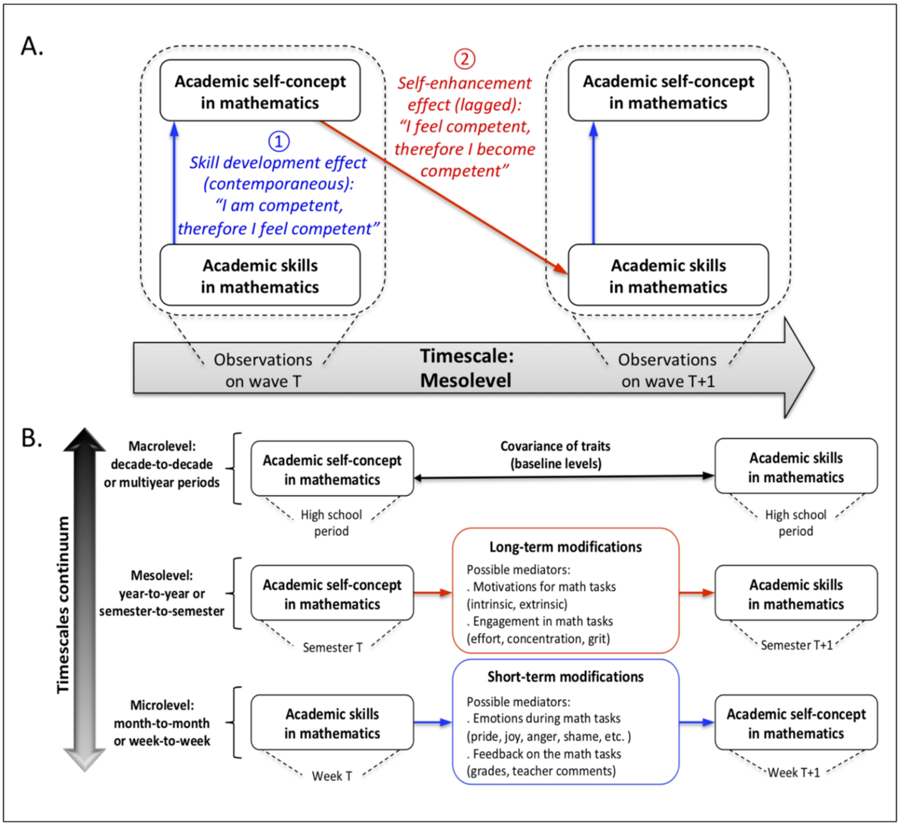

Fig. 1 Hypothesized relations between academic self-concept and skills across timescales. **A**) Contemporaneous skills development effects and lagged self-enhancement effects in a mesolevel design (yearly or semestrial waves); **B**) theoretical explanation in terms of differing timescales of modifications induced by the reciprocal effects, with possible mediators. (adapted from the "timescales continuum" framework; Hamaker, 2023). This study only tests the temporal structure posited by the CL-REM theoretical framework (putative mediators are not investigated)

We tested these predictions using longitudinal data from the ProFAN program (2017–2020), a large-scale study conducted in French high schools, involving over 9,000 students enrolled in two-year vocational programs. This study constitutes a secondary analysis of ProFAN, whose measurement schedule (see Method) was not developed specifically to test the CL-REM but nonetheless provides the temporal structure needed to examine the fast and slow processes. More precisely, the 8-month intervals of the ProFAN design is appropriate to test both the fast, skill-development effects (as lag0 effects) and the slower, self-enhancement effect (as lag 1 effects), because it aligns with yearly or multimonth intervals expected from the CL-REM (Fig. 1). Moreover, academic self-concept and academic skills were assessed in both mathematics and French (language arts) across three waves of data collection from Grade 11 to Grade 12, which is sufficient for separating state-like and trait-like variation via random intercepts. The ProFAN dataset thus provides an independent context for conceptually replicating the fast-slow dynamics of self-concept and skills reported by Marsh et al. (2024), thereby testing their robustness in a new context with alternative measures (Fabrigar & Wegener, 2016).

To operationalize the CL-REM, we used structural equation modeling, specifically a variant of the RI-CLPM that includes both contemporaneous effects of skills on self-concept and lagged effects of self-concept on skills, while also separating stable traits from dynamic states (state-trait disaggregation). This specification, referred to as the RI-CCLPM (RI-CLPM with contemporaneous and cross-lagged effects), directly reflects the temporal asymmetry central to the CL-REM, thereby combining theory and modeling to better capture the dynamics linking self-concept and skills (Marsh & Hau, 2007).

To evaluate the uniqueness and robustness of the RI-CCLPM, we compared it with alternative models that offer competing accounts of how self-concept and skills co-evolve (Supplemental Material A; Núñez-Regueiro et al., 2025). The first group of alternatives retained the state-trait separation but varied the temporal dynamics, with some models allowing only contemporaneous effects (RI-CPM), others only lagged effects (RI-CLPM). The second group of models (i.e., LCSM) assumed that traits emerge from the accumulation of state-level changes, thereby entangling state and trait processes. These comparisons allowed us to test whether trait-state entanglement is supported by the data and whether reciprocal effects persist when the state-trait separation of the RI-CCLPM is removed. We also examined models that attribute the observed effects to artefacts rather than genuine reciprocal influences. These included the STARTS model, which captures residual state-level fluctuations, and the RI-LTVC model, which accounts for unmeasured latent third-variable confounders.

Finally, we assessed the robustness of the RI-CCLPM by including second-order effects (from time T to T+2, lag2) to capture possible unmeasured time-varying confounders, and by adding background covariates (e.g., age, gender, parental SES, parental education level) to adjust for stable individual differences that could influence both constructs.

Method

Participants

The present study is a secondary analysis of data from the ProFAN project, a large-scale longitudinal research initiative conducted from 2017 to 2020 by the French Ministry of Education to investigate the effects of cooperative learning, particularly the jigsaw classroom instructional technique, on a broad array of outcomes including academic achievement, socio-emotional functioning, and school adaptation (for details, see SM-B). The project involved two cohorts of students (2017 and 2018) from 109 vocational secondary schools across 10 “Academies” (School Districts), enrolled in one of three vocational pathways (business, care services, or electricity and connected environments). The initial target sample comprised 10,035 students, but the final sample was reduced to 9,357 after excluding students with non-admissible responses (i.e., satisficing, 3.4%) and missing administrative data (i.e., no class identifiers, 6.6%). The final sample included students aged between 15 and 17 years, of whom 43.8% were male. Based on mother occupational status, 51% of students came from disadvantaged backgrounds, 29% from intermediate backgrounds, and 20% from advantaged backgrounds.

Within ProFAN, students followed the ordinary curriculum throughout the school year (approximately 225 class hours per semester). The ProFAN experimentation consisted in the incorporation of short learning sessions (5 to 10 h per semester), which implemented varying degrees of cooperation across conditions. Specifically, students were assigned to one of three conditions: A “jigsaw” condition involving strong cooperation through resource interdependence ($n=3,383$), a “cooperation” condition involving weak cooperation without resource interdependence ($n=3,207$), and a “business-as-usual” condition involving individual work (no cooperation; $n=2,767$). As this experimental design is part the data-generating process, multigroup analyses were conducted to assess whether reciprocal effects between self-concept and skills differed across conditions; no significant differences were found (Fig. 3). This is consistent with previous ProFAN studies, which did not find evidence for main effects of experimental condition on student beliefs such as intrinsic motivation, self-regulation, or social competences (Riant et al., 2024; Rudmann et al., 2024). Accordingly, our analyses focused on the pooled sample of 9,357 students, with multigroup results demonstrating the robustness of the findings across conditions.

Procedure

Data for the ProFAN study were collected via online questionnaires and standardized tests administered at three time points: the autumn of Grade 11 (October-November, T1; before intervention), the late spring of Grade 11 (May-June, T2; during intervention), and the spring of Grade 12 (April-May, T3; after intervention). All data were anonymized prior to analysis. The questionnaires were delivered through a secure online platform developed specifically for the ProFAN project and comprised 251 items assessing a broad range of constructs, including socio-emotional competencies, school adjustment, academic performance, and education-related beliefs. Students completed the questionnaires during school hours under the supervision of a teacher, with each session lasting approximately one hour. Standardized academic tests were similarly administered within the same academic term as the questionnaires, typically a few weeks to 2 months apart. Given this relatively short and unsystematic lag, and in line with prior research modeling short-term reciprocal processes, we treated these measures as contemporaneous, that is, as occurring within a shared timeframe of academic experience. In contrast, the spacing between successive waves reflected intervals of approximately 8 months of schooling (when accounting for the 2-month summer break), corresponding to semester-level variation in students' academic experience.

Measures

Academic self-concept Academic self-concept (ASC) was assessed using the six-item scale of Perceived Competence adapted from Harter's *Self-Perception Profile for Children* (Harter, 1988), consistent with prior research (e.g., Núñez-Regueiro et al., 2022). Two disciplinary-specific ASC measures were administered—one for mathematics and one for French (language arts)—using a 7-point Likert-type response scale ranging from 1 (*not at all true for me*) to 7 (*very true for me*). The same six items were used in both domains, with subject-specific wording. Items included: (1) *I feel as good as other students my age in < Discipline>*, (2) *I often forget what I learn in < Discipline> (reversed)*, (3) *I'm rather slow at doing my work in < Discipline> (reversed)*, (4) *I feel like I do very well in my work in < Discipline>*, (5) *I have a lot of trouble finding the right answers in < Discipline> (reversed)*, and (6) *I do my work very well in < Discipline>*. Confirmatory factor analyses supported a latent variable model in which (1) item uniquenesses were correlated across time (i.e., between-wave correlated residuals for each item), and (2) method effects were modeled by correlating residuals of the negatively worded items (Items 2, 3, and 5) within each wave. This approach reduces measurement error and method bias, improving construct validity. Measurement invariance across the three time points (T1-T3) was assessed through a sequence of nested models. Results supported scalar longitudinal invariance (i.e., equal configuration, factor loadings, and item intercepts across waves), for both mathematics [$chi2\,(127) = 990.1$, $p < .001$, $CFI = 0.986$, $TLI = 0.983$, $RMSEA = 0.031$, $SRMR = 0.065$] and French [$chi2\,(127) = 1175.1$, $p < .001$, $CFI = 0.968$, $TLI = 0.961$, $RMSEA = 0.034$, $SRMR = 0.075$; see SM-B]. Therefore, changes in ASC scores reflected genuine change in self-perceptions rather than changes in item functioning (Brown, 2015; Widaman et al., 2010). Reliability estimates were satisfactory for ASC in mathematics [$\omega = .795\,(T1)$, $.805\,(T2)$, $.810\,(T3)$], and somewhat lower but acceptable for French [$\omega = .681\,(T1)$, $.678\,(T2)$, $.685\,(T3)$]. Complementary analyses (SM-E) showed that reliability in French could be improved by removing the negatively worded items (Table S9); however, this refinement did not alter the substantive results (Figure S6), indicating that the findings are not sensitive to this potential measurement issue.

Academic skills Standardized tests in mathematics and French were developed in collaboration with vocational teachers to align with the national curriculum. Student responses to the tests were digitized, resulting in approximately 37,000 test scripts. These scripts were anonymized and randomly assigned to independent vocational high school teachers—unaffiliated with the ProFAN experiment—for blind grading. This procedure ensured objective and unbiased assessment of students' academic performance across all conditions.

Covariates The data collection also measured students' age, gender, number of siblings, parental SES, past grade retention, and past achievement levels in mathematics and French (4-category ordinal variables, from 1 = poor performance to 4 = excellent). These covariates were included in the analyses as part of the robustness checks for observed confounders.

Analytic Strategy

Structural equation models To investigate the reciprocal effects between academic self-concept and skills, a series of structural equation models were estimated using Mplus version 8.8 (Hallquist & Wiley, 2018; Muthén & Muthén, 2022). All models posited alternative data-generating processes for the co-evolution of self-concept and skills (for details, see Supplemental Material A). These included the traditional cross-lagged panel model (CLPM) (Marsh et al., 1999; Marsh & Craven, 2006), the random intercept CLPM (RI-CLPM) (Hamaker et al., 2015), the RI-CLPM with contemporaneous effects (RI-CCLPM) (B. Muthén & Asparouhov, 2024), a variant with residual occasion-specific error (STARTS) (Kenny & Zautra, 1995), and a variant substituting reciprocal effects by loading on a latent third variable confounder (RI-LTVC) (Kenny & McCoach, 2025; Sorjonen et al., 2025). In robustness checks, effects were also tested against confounders, first by adding autoregressive effects between non-adjacent occasions (unobserved confounders) (Lüdtke & Robitzsch, 2022; Murayama & Gfrörer, 2024; VanderWeele et al., 2020), and then by adding covariate effects on common factors (random intercepts, accumulating factors) (Kievit et al., 2018; Mulder & Hamaker, 2021). Note that the term "contemporaneous" effect denotes a lag0 effect, that is, a directional relation between two constructs measured within the same wave (B. Muthén & Asparouhov, 2024). In the present design, a single wave typically spanned between a few weeks to 2 months (see Procedure). By contrast, the lag1 effects reflected changes across semesters, corresponding to an interval of approximately 8 months (i.e., lag1 effect = 8 months).
Across all models, scalar-invariant latent constructs for self-reported data (i.e., academic self-concept) were used while incorporating correlated uniquenesses and method effects, thereby reducing measurement error and improving interpretation (Supplemental Material B). Moreover, although the three measurement occasions correspond to distinct phases of the ProFAN intervention (pre-intervention, during intervention, and post-intervention), all regression paths were set to be time-invariant, unless this compromised model identification (e.g., freeing autoregressive paths in the full RI-CLPM-C) (B. Muthén & Asparouhov, 2024). Constraining parameters to be invariant across phases serves to improve estimation efficiency and interpretation of common within-person dynamics. To evaluate whether this constraint was

tenable given the phase structure, we compared the invariant model with a model allowing parameters to vary across the T1-T2 and T2-T3 intervals. Because the time-varying specification did not improve model fit ($\Delta CFI > -0.01$) and revealed no systematic interval-specific differences (see Fig S2, SM-E), we retained the invariant model for parsimony and interpretability.

Inference used two-tailed Wald tests under large-sample normal theory with cluster-robust standard errors (B. Muthén & Satorra, 1995) to account for data clustering (students nested within classrooms); tests therefore remain valid under clustering and non-normality. Missing data were handled using full information maximum likelihood (FIML), which provides unbiased parameter estimates under the assumption that data are missing at random (MAR) (Graham, 2012). This yielded analysis samples of N=9,142 students for mathematics and N=9,041 for French, or 97–98% of the full sample (N=9,357). More precisely, missingness ranged from 15% (T1) to 55% (T3) for self-concept, and from 33% (T1) to 49% (T3) for skills, reflecting typical attrition patterns in large-scale longitudinal studies. Attrition patterns were similar across experimental conditions (between 42% and 44%) and higher in the second cohort (57% in 2018 vs. 28% in 2017). Preliminary analyses indicated that missingness was associated with background variables (being male, older, with siblings, grade retention) and with levels of self-concept and skills (Table S7, SM-E). Logistic regression analyses further showed that missingness was significantly predicted by self-concept and skills levels on a given wave (T1, T2 or T3), a pattern consistent with a MAR mechanism conditional on observed variables (Table S8, SM-E). Sensitivity analyses indicated that the final models were robust to alternative missing data treatment, as the main findings were reproduced when individuals with missing data over time (T2, T3) were removed from the analyses (Fig. S3, SM-E). In terms of estimation, non-missingness covariance coverage ranged from 0.17 to 0.80, with high coverage for within-wave paths (≥0.79) and lower values for distant or cross-domain paths (0.17–0.31.17.31), but always above the recommended threshold (0.10) (L. Muthén & Muthén, 2022) and ensuring adequate model estimation.

One challenging aspect of the modeling concerned the simultaneous inclusion of contemporaneous and lagged effects, which is known to create convergence issues (e.g., non-positive variance-covariance matrix, negative R2s) (B. Muthén & Asparouhov, 2024). We followed recommendations from these authors to estimate both effects using nonduality parametric constraints (i.e., $0 < \beta\prime_2\beta\prime_3 < 1$ in Fig. S1, Supplemental Material A), and to eliminate nonsignificant effects or effects estimated at zero. This approach enabled establishing whether the reciprocal effects between self-concept and skills were lagged, contemporaneous, or both. Another challenge was estimating measurement error within the STARTS, which usually requires 4 or more waves to reach stable solutions (Hamaker et al., 2015; Kenny & Zautra, 1995). It can nevertheless be identified with 3 waves of data (as in this study) by constraining the error to be invariant across waves (Usami et al., 2019). Such constraints were effective with the error in measures of academic self-concept (as these were based on a measurement model yielding more precise estimates of the true score), but resulted in convergence issues—i.e., Heywood cases—with error for measures of skills (single item measures). Therefore, we only estimated residual state variance for academic self-concept.

Models' performances were assessed using conventional goodness-of-fit criteria ($CFI \geq .90$ and $RMSEA \leq .08$ for acceptable fit, $CFI \geq .95$ and $RMSEA \leq .06$ for excellent fit; Hu & Bentler, 1998). Meeting these criteria can be seen as necessary condition for ruling out misspecification in the models, although it does not suffice for selecting a best-fitting solution among candidate models. Indeed, very different models may yield quasi-identical values on these indices—particularly highly complex models with enough flexibility to match a wide range of data structures, regardless of their correspondence to the true data-generating process (Kenny & McCoach, 2025; Marsh et al., 2024; Muthén & Asparouhov, 2024). In contrast, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) assess model fit while penalizing complexity. To compare multiple models, we computed weighted AIC (wAIC) and weighted BIC (wBIC), which estimate the relative likelihood that each model is the best approximation of the data among those considered. These weights sum to one and offer a probabilistic interpretation, with higher values indicating stronger empirical support (Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004). We considered these criteria while keeping in mind that they require appropriate model specification, including statistically significant defining parameters; in cases of conflict, we prioritized appropriateness of model specification over fit indices alone. To aid interpretability of effect sizes (see Fig. 4), average standardized reciprocal effects (β) were converted to correlations using $r = 0.98\beta + 0.05$ (Peterson & Brown, 2005), and then to Cohen's *d* using $d = \frac{2r}{\sqrt{1-r^2}}$ (Borenstein et al., 2021).

Ethics and Data Availability

The ProFAN experimentation was commissioned and authorized by the French Ministry of Education under an official ministerial decree (Arrêté du 21 mars 2017). Formal participation agreements were signed between each participating school and the corresponding Academic Directorate (Rectorat). All data were anonymized prior to analysis, and data collection and processing complied with French data-protection regulations, including the General Data Protection Regulation (GDPR).

The data is private and is the property of the ProFAN consortium. Due to legal and institutional restrictions, the authors do not have permission to share the dataset directly. Researchers interested in accessing the data and materials should contact the ProFAN program administrators (Monteil et al., 2022) or the corresponding author, who can assist in initiating the formal request procedure. Analysis code for replicating the comparative modeling framework is available at the OSF repository: https://osf.io/84bzp/ (see also Núñez-Regueiro et al., 2025).

Results

Descriptive Patterns and Correlations

Figure 2 shows the distribution and evolution of academic self-concept and skills across the three waves. Mean self-concept scores in both mathematics and French

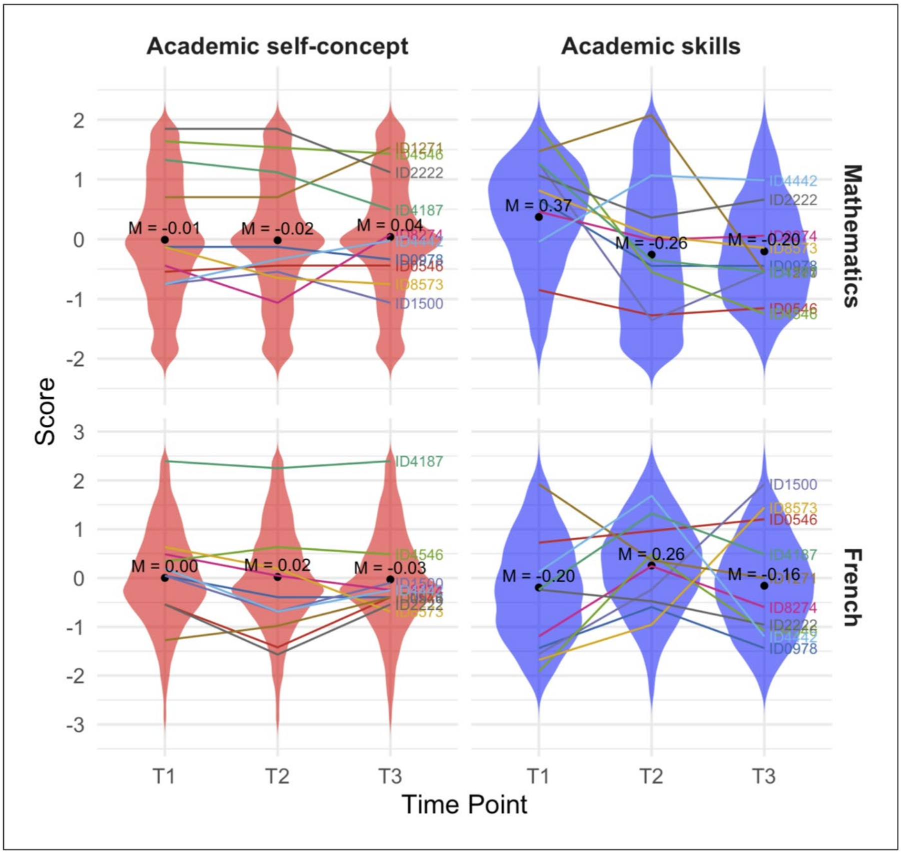

Fig. 2 Distributions of scores of academic self-concept and skills across time points and domains in the ProFAN dataset. Academic self-concept reflects students' perceived competence in a given domain (sample item, "I feel like I do very well in my work in < Discipline>"), measured using validated self-report scales. Academic skills were assessed via standardized tests aligned with the national curriculum in mathematics and French (language arts). To ensure comparability, all scores were grand-standardized across waves (M = 0, SD = 1). Density plots illustrate the distribution of scores at each wave, reflecting central tendencies and variability. Dashed lines represent individual-level trajectories for a random subsample of 10 students, highlighting the heterogeneous, dynamic nature of changes in self-perceptions and skills over time

remained relatively stable over time. In contrast, mean skill scores showed slight shifts (e.g., in mathematics at T1), reflecting variations in the tests to align with different segments of the curriculum over time.

To illustrate individual variability, Fig. 2 also plots the trajectories of a subset of students. Some students showed parallel changes in self-concept and skills within the same domain, whereas others maintained a relatively stable self-concept despite changes in performance. For example, student ID4187 exhibited a parallel decline in mathematics self-concept and skills over time, whereas student ID4546 sustained a high mathematics self-concept despite falling scores. Similarly, student ID0978 showed decreases in both self-concept and skills for French, while student ID1500 improved markedly in French skills with little corresponding change in self-concept. Other students displayed minimal change, such as student ID8573, whose mathematics scores and self-concept remained largely flat with only slight declines toward the final wave.

Correlation analyses (see Table S3, Supplemental Material B) provide additional insight into the temporal structure of the data. Within-wave correlations between academic self-concept and skills were modest (i.e., ranging from $r = .27to.37$ in mathematics and $r = .10to.20$ in French) suggesting that, although related, these constructs capture distinct processes on each time point. In contrast, correlations between adjacent waves of the same construct (e.g., French self-concept on T1 and T2) were substantially higher, typically above $r = .65$, consistent with the influence of stable, trait-like characteristics. Taken together, these patterns suggest that part of the variance in self-concept and skills reflects enduring traits shaped by background or prior learning history, while a substantial portion (not captured by traits) reflects more dynamic, state-like fluctuations that respond to recent academic experiences.

Models Operationalizing to the CL-REM (RI-CCLPM Specifications)

We first estimated models that empirically operationalize the CL-REM framework (Fig. 1), namely RI-CCLPM specifications. These models distinguish stable traits (random intercepts) from dynamic states (wave-specific residuals), while allowing for both contemporaneous effects (lag0, denoted ß') and lagged effects (lag1, denoted ß) between self-concept and skills dynamic states. Following recommended guidelines (B. Muthén & Asparouhov, 2024), we began with saturated RI-CCLPMs including both contemporaneous and lagged cross-effects, and then removed nonsignificant effects to improve parsimony. The resulting models converged properly and showed excellent fit to the data (e.g., $CFI = 0.98$ and $RMSEA = 0.02$ in mathematics, $CFI = 0.96$ and $RMSEA = 0.03$ in French; see Appendix).

Path diagrams for the RI-CCLPM solutions in mathematics and French are shown in Fig. 3 A and B. In both domains, academic self-concept showed moderate autoregressive effects ($\beta_1 = [.391, .398]$ in mathematics and $\beta_1 = [.411, .431]$ in French; all $p < .001$), indicating partial stability in state-level fluctuations across semesters. Autoregressive effects for academic skills were smaller ($\beta_4 = [.075, .154]$ in mathematics and $\beta_4 = [.090, .100]$ in French), indicating lower stability, likely due to differences in test contents across semesters (curriculum changes). Importantly, both models supported fast-slow dynamics in the cross-effects, in alignment with the CL-REM framework (Fig. 1). In mathematics, academic skills had contemporaneous effects on self-concept at each wave ($\beta_{3'} = [.157, .221], p < .001$), alongside smaller but significant lagged effects ($\beta_3 = [.051, .075]\,, p = .011$). Mathematics self-concept, in turn, predicted future skill gains in the form of lagged effects ($\beta_2 = [.111, .159], p < .001$), whereas contemporaneous effects were nonsignificant and fixed to zero for parsimony. A similar pattern was evidenced in French: Academic skills predicted variations in self-concept within the same semester via contemporaneous effects ($\beta_{3'} = [.113, .115], p < .001$), and self-concept predicted subsequent skill variations via lagged effects only ($\beta_2 = [.075, .080]\,, p = .015$). Other effects

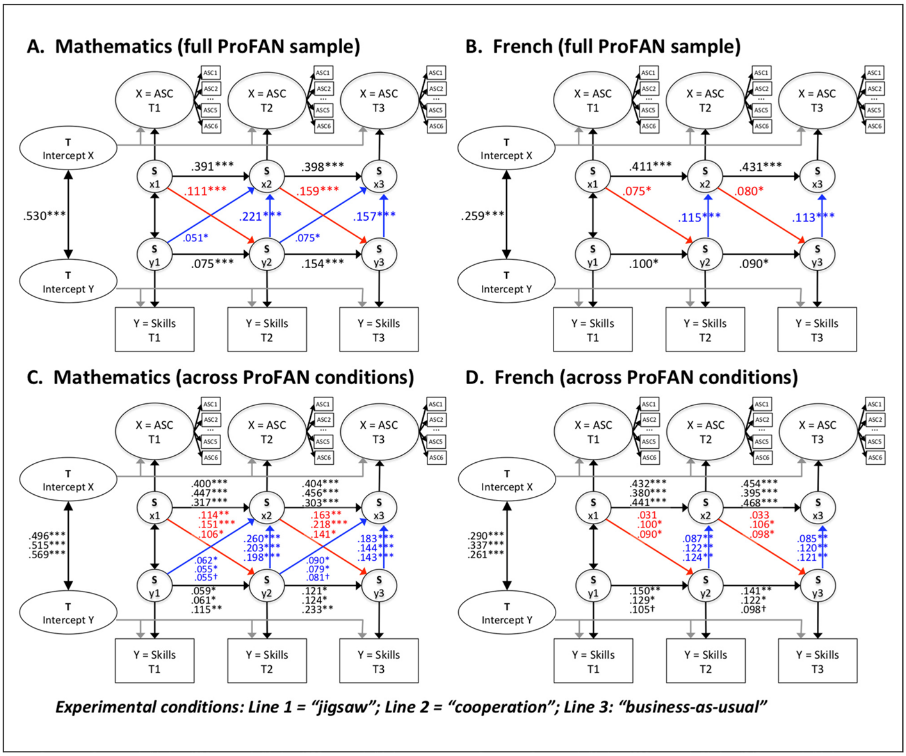

Fig. 3 Final models of reciprocal effects between academic self-concept and skills in the full sample and across conditions. N=9142 (mathematics), 9041 (French). Academic self-concept (ASC) is operationalized as perceived competence; skills as standardized test performance. The models disaggregate between-person traits ('T', macrolevel timescale) from within-person dynamic states ('S') and estimate reciprocal effects between states either as contemporaneous effects within the same semester (microlevel timescale) or as lagged effects across semesters (mesolevel timescale). In both the full sample **A**, **B** and across experimental conditions **C**, **D**, self-concept predicts changes in skills across semesters, whereas skills predict changes in self-concept within semesters (and marginally, for mathematics, across semesters). These patterns provide the best approximation of the data structure, and are robust to alternative specifications of change processes and to confounding effects (i.e., carryover effects, covariate effects). See Supplementary Tables S10 and S11 for full estimates, SEs, Wald statistics, and confidence intervals (*Statistics* sheets). †$p \leq .1$; *$p < .05$; **$p < .01$;***$p < .001$

(lagged effects of skills, contemporaneous effects of self-concept) were nonsignificant and omitted.

Moreover, both RI-CCLPM solutions (Fig. 3A and B) evidenced positive correlations between the trait-like factors (i.e., intercepts of self-concept and skills) representing stable developmental differences between students. The stronger trait correlation in mathematics ($r = .530, p < .001$) than French ($r = .259, p < .001$) may reflect the fact that reciprocal effects between self-concept and skills accumulated over earlier school years (e.g., in primary or lower-secondary education) were more pronounced in mathematics than in French. This interpretation is also supported by the larger reciprocal effects in mathematics (compared to French) observed during the present high school period (Fig. 3).

Multigroup RI-CCLPM analyses indicated that the pattern of contemporaneous and lagged effects was consistent across experimental conditions in ProFAN (Fig. 3C and D). Formal comparisons of models constraining effects to be equal or different across conditions showed no meaningful differences, either in terms of model deviance ($\Delta\chi^2\,(10) = 11.4, p = .325$ in mathematics, $\Delta\chi^2\,(8) = 4.5, p = .807$ in French) or in approximate fit indices ($\Delta CFI = 0.000$ and $\Delta RMSEA = 0.000$ in both domains). These results indicate that a single-group, full-sample RI-CCLPM solution (Fig. 3A and B) provides an adequate representation of the data, yielding more stable and statistically efficient estimates.

Alternative Models of Self-concept and Skills Co-evolutions

To evaluate the robustness and specificity of the RI-CCLPM solutions, we compared them to alternative models making distinct assumptions about change processes. These included variants with only lagged or contemporaneous effects, as well as variants addressing potential sources of misspecification, namely the STARTS (accounting for occasion-specific error) and the RI-LTVC (modeling an unobserved third-variable confounder). Although all models showed acceptable conventional fit, the RI-CCLPM stood out as the best-fitting solution across both academic domains—combining consistent model convergence, superior performance on information-theoretic criteria (wAIC, wBIC), and interpretable model parameters.

These comparisons yielded several informative contrasts (see Supplementary Table S4 SM-D for details, and Supplementary Tables S14 to S27 for full estimates, SEs, Wald statistics, and confidence intervals). For example, the RI-CLPM and RI-CPM variants, which assumed fully lagged or contemporaneous effects, failed to detect reciprocal effects between self-concept and skill in French, indicating that assuming symmetric temporal dynamics may obscure key relations among constructs. In contrast, reciprocal effects remained robust in mathematics under both specifications. LCSM variants yielded similar effects as the RI-CCLPM solutions, but often failed to converge to a proper solution and showed worse fit to the data, indicating that modeling trait emergence from accumulating states did not align well with the data structure.

Similarly, the RI-LTVC and STARTS variants, designed to rule out spuriousness due to latent confounding or residual error, also failed to challenge the RI-CCLPM results. In both domains, their core parameters—latent confounder loadings in the RI-LTVC and residual variances in the STARTS—were consistently nonsignificant. This indicates that the observed reciprocal effects in the RI-CCLPM (Fig. 3) are unlikely to be artifacts of unmeasured third variables or transient fluctuations unaccounted for by the trait-state decomposition.

Together, these results confirm that the reciprocal effects uncovered by the RI-CCLPM specifications are robust to alternative accounts of skills and self-concept co-evolutions. Full comparisons and diagnostics are presented in Supplemental Material C.

Robustness to Time-varying and Time-invariant Confounders

We further tested whether the effects in the RI-CCLPM solutions could be accounted for by time-varying or time-invariant confounding influences (see Supplementary Tables S5 and S6 in SM-D for details, and Supplementary Tables S28 to S31 for full statistics). To address time-varying confounds, we extended the RI-CCLPMs by including second-order autoregressive and cross-lagged paths, capturing potential delayed carryover effects spanning two semesters (lag2 effects, from T to T+2). In mathematics, the original effects remained significant, while the new second-order effects were nonsignificant. This indicates that the observed effects were not simply artifacts of longer-term lagged dependencies. In French, the lagged effect of self-concept on skills disappeared when including the second-order path ($\beta_2 = [.028, .033], \mathrm{p} = .435$), but the added CL2 path was also nonsignificant ($\mathrm{p} = .587$)—even when estimated in isolation—suggesting it does not constitute a genuine carryover effect.

To address time-invariant confounds, we re-estimated the RI-CCLPMs while allowing trait factors to be predicted by stable student background characteristics including age, gender, parental SES, parental education, number of siblings, grade retention, and prior academic achievement. This adjustment controls for between-person heterogeneity, that is, individual differences in long-term academic baselines that might bias within-person estimates. Several covariates significantly predicted trait self-concept and skills (e.g., older age and grade retention predicted lower skills; gender showed opposite effects on self-concept and achievement), but the RI-CCLPM solution remained robust. In mathematics, self-concept still predicted later skill development ($\beta_2 = [.120, .177], p < .001$), and skills predicted immediate changes in self-concept ($\beta_{3\prime} = [.154, .218], p < .001$). In French, self-concept again predicted subsequent skills ($\beta_2 = [.073, .080]\,, p = .016$), and skills predicted concurrent self-concept ($\beta\prime_3 = [.103, .108], p < .001$).

We further probed the influence of stable student characteristics using multigroup variants of the RI-CCLPM. Analyses by gender (Fig. S4, SM-E) showed highly comparable effects for male and female students. In mathematics, the temporal asymmetry was replicated with similar effect sizes across gender. In French, contemporaneous skill effects were observed in both groups, whereas the lagged self-concept effect on skills was weaker among male than female adolescents; however, formal model comparisons indicated that allowing effects to vary across gender did not improve model fit, suggesting that these differences should be interpreted cautiously. Multigroup analyses by initial ability level (Fig. S5, SM-E) likewise indicated that the temporal structure of reciprocal effects was highly consistent across low-, moderate-, and high-achieving students, providing no evidence that the observed fast-slow asymmetry depends on initial skill level.

Together, these analyses confirm that the asymmetrical dynamics captured by the RI-CCLPM solutions (Fig. 3)—contemporaneous effects of skills and lagged effects of self-concept—are robust to a range of both time-varying and time-invariant confounding influences.

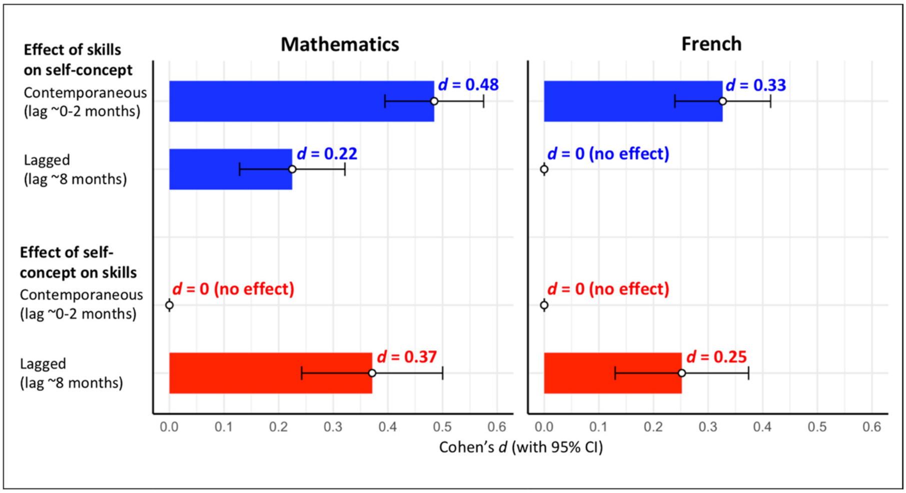

Fig. 4 Average effect size estimates of reciprocal influences between academic self-concept and skills. Contemporaneous effects reflect within-semester associations (e.g., from skills at T1 to self-concept at T1), and correspond to a lag of approximately 0–2 months in this study. Lagged effects reflect between-semester influences (e.g., from self-concept at T1 to skills at T2, or from T2 to T3) and correspond to a lag of approximately 8 months. Bars represent the average contemporaneous or lagged effects from the final models (see Fig. 3, for ß estimates), expressed in effect sizes metrics as Cohen's *d* values. Error bars indicate 95% confidence intervals. The results highlight a consistent temporal asymmetry: academic skills mainly predict short-term or "fast" gains in self-concept (within semesters), while self-concept predicts long-term or"slow" gains in academic skills (between semesters)

Effect Sizes and Temporal Interpretation

Standardized effect sizes for the RI-CCLPM results are displayed in Fig. 4. These reflect dynamic, within-person effects—capturing how changes in one construct (e.g., academic self-concept) predict changes in the other (e.g., academic skills), over and above prior state fluctuations and long-term stability traits. To aid interpretation, we approximated effect sizes as Cohen's *d* (Cohen, 1988), which here represents the expected standardized change in one variable following a one standard deviation increase in the other.

In mathematics, the contemporaneous effect of academic skills on self-concept was moderate-to-large (d = 0.48), while the lagged effect of self-concept on skills was). In French, a similar asymmetry emerged, with a contemporaneous effect of skills of moderate size ($d = 0.33$) and a lagged effect of self-concept of comparable size ($d = 0.25$). In contrast, the alternative cross-effects—i.e., lagged effect of skills and contemporaneous effect of self-concept—were null or negligible in size (d = 0.22), suggesting that only specific temporal effects were reliably present.

This pattern of modeling results are consistent with theoretical predictions from the CL-REM framework, suggesting that academic skills modify self-concept relatively quickly (in less than a semester), while self-concept influences skills development more gradually (across a full semester).

Discussion

The present study provides a large-scale replication of preliminary evidence by Marsh et al. (2024) that the reciprocal effects between academic self-concept and skills unfold at different speeds. Replicating their findings, we observed a temporal asymmetry: Academic self-concept predicted slower, long-term changes in skill development, whereas skills predicted faster, short-term changes in self-perceptions. This pattern was consistent across mathematics (as in Marsh et al.) and French (new domain) and remained stable across a wide range of robustness and sensitivity analyses. Beyond replication, this study develops a theoretical framework that explicitly distinguishes fast and slow processes in learning behavior, thereby explaining why skills effects emerge more rapidly, whereas self-concept effects emerge more gradually over time. More broadly, returning to the foundational question (i.e., do perceptions of competence shape skill development, or are they merely reflections of actual skills?), our findings are consistent with the view that both processes co-exist, but that they operate on different timescales. For intervention and longitudinal researchers, these findings also highlight how temporal misspecification can obscure theoretically meaningful dynamics even in well-designed studies.

Main Findings: Self-concept and Skills Interact at Different Speeds

Across both mathematics and French, skills had contemporaneous effects on self-concept over a shorter lag (lag = 0–2 months, approximately), while self-concept predicted gains in skills over a longer lag (8 months approximately). In practical terms, a one standard deviation increase in skills was related to a gain of one-third to one-half of a standard deviation in self-concept, observed over the same measurement window spanning a few weeks to few months. Conversely, a one standard deviation increase in self-concept was related to a gain of one-quarter to one-third of a standard deviation in skills the following semester, observed approximately 8 months later. These findings are consistent with the theoretical assumptions of the CL-REM framework (Fig. 1), which posits that reciprocal influences occur at distinct speeds—some fast (skill effects on self-concept), others slow (self-concept effects on skills).

The observed asymmetries remained robust across a range of alternative specifications designed to test competing explanations (Núñez-Regueiro et al., 2025). First, we ruled out the possibility that stable individual differences, such as those shaped by earlier experiences and background conditions (e.g., gender, socioeconomic status, prior school trajectories), might account for the observed reciprocal effects. We did so by separating baseline levels of self-concept and skills that had stabilized by high school (traits), from more dynamic processes responsive to current academic experiences (states; RI-CLPM and variants). Second, we addressed the potential influence of measurement error by using latent factors and by testing the robustness of the final models while accounting for occasion-specific fluctuations in the constructs (STARTS). Third, we tested whether the associations could be explained by common, unmeasured factors influencing both constructs simultaneously, such as a general mood or liking for school (RI-LTVC). We also compared models that conceptualize traits as emerging from the accumulation of prior state-level changes (LCSM), contrary to the state-trait separation posited by the CL-REM. Finally, we adjusted for time-varying and time-invariant confounders, including prior dynamic states, sociodemographic factors (age, gender, parental SES, parental education), and initial skills level. Across all checks, the same core pattern emerged: Skills were related to within-wave (or short-term, “fast”) changes in self-concept, while self-concept was related to between-wave (or long-term, “slow”) changes in skills. While no model can prove causality definitively, the convergence across specifications strengthens the plausibility of the CL-REM as the best account of how self-concept and skills co-evolve in the present sample and timeframe.

Limitations and Research Perspectives

Several limitations should be noted. Although the breadth of our model comparisons—testing multiple competing specifications, controlling for a wide range of potential confounders, and ruling out alternative temporal structures—provides greater confidence in temporal ordering and robustness than most observational studies, the data remain correlational and causal claims should still be made with caution. A notable limitation is the fact that the study constitutes a secondary analysis of data from the ProFAN experimentation, which involved distinct experimental conditions as well as qualitatively differences phases, with T1 occurring prior to the intervention, T2 during the intervention, and T3 after it. Relatedly, because the duration of assessments within each wave varied across classrooms as a function of phase implementation constraints, contemporaneous effects reflect within-wave associations unfolding over a few weeks to a few months, rather than instantaneous change. Although model comparisons indicated that allowing parameters to vary across experimental conditions (Fig. 3) or across phases (Fig. S2, SM-E) did not alter the detection of effects under investigation, the observed temporal asymmetries between skill development and self-enhancement effects (i.e., lag0 vs. lag1 effects) should be interpreted with caution, as they may partly reflect condition-specific or phase-specific variations alongside the fast-slow principle posited by the CL-REM. Future studies based on purer observations (nonexperimental designs) will be needed to further assess the robustness of the observed effect asymmetries.

Relatedly, the sample consisted of vocational high school students in France, which may limit generalizability to other educational contexts, given possible differences in academic history, self-perception, or motivation. Importantly, however, Marsh et al. (2024) reported the same pattern of fast-slow effects between self-concept and skills among German high school students, suggesting that these dynamics are not restricted to the French context. Future research should nevertheless test the CL-REM across more diverse populations, such as younger students populations or other academic domains. Another limitation is that the measures for French self-concept showed modest reliability. This was partially mitigated by the use of latent factors, and sensitivity analyses indicated that the lower reliability did not affect the magnitude of the reciprocal effects (Fig. S6, SM-E). The smaller effects observed in the French domain (vs. mathematics) therefore likely reflect substantive domain differences rather than measurement artifacts. Alternatively, these differences could reflect subgroup variations, including gender differences in reciprocal effects (Fig.

S4, SM-E); clarifying how such factors (domain, gender) interact with the temporal dynamics identified here represents an important direction for future research and theory. Finally, while our measures were validated and curriculum-aligned, other influences—such as school climate, peer comparisons, and parental support—may also shape self-perceptions and learning trajectories; incorporating these factors could enrich future applications of this framework.

Implications for Theory, Research, Intervention

Despite these limitations, the findings are consistent with a temporally structured view of psychological development in learning. Academic self-concept and skills both contain stable trait components and dynamic state components. Our results suggest that reciprocal effects occur at the state level, with traits remaining stable across Grades 11-12 in this sample. Yet, as articulated in the CL-REM framework, traits are not fixed and likely emerge, in part, from the long-term accumulation of earlier reciprocal effects between self-concept and skills, and in part from background characteristics, as evidenced by the analysis of covariate effects on traits (age, gender, parental SES, etc.). This perspective clarifies the distinct roles of trait and state processes in shaping learning behavior.

Our findings carry broader methodological and conceptual implications. Much of the psychological literature relies on models that impose uniform lag structures and conflate stable traits with dynamic states. Yet, developmental processes may unfold over multiple, partially overlapping timescales. When statistical models fail to accommodate this complexity, they risk misattributing effects or overlooking meaningful temporal patterns. In our analyses, for example, models that included only contemporaneous or only lagged effects failed to detect significant reciprocal relations between self-concept and skills in French. This absence of detectable effects was not due to a lack of association between the constructs (as the best-fitting model showed), but rather to a mismatch between the models' assumptions and the actual timing of these effects. This is because misspecifying the timescales of an effect (e.g., as occurring on T + 1 instead of T) can lead the model to misallocate variance, thereby distorting the estimation of the remaining parameters. As a result, true effects may be obscured or underestimated, even when they exist (Muthén & Asparouhov, 2024; Speyer et al., 2024). The takeaway is that mismatches between speeds of change and model temporal structure can obscure core psychological dynamics. As argued elsewhere (Marsh et al., 2024; Murayama & von Keyserlingk, 2025), psychological constructs are often theoretically vague regarding temporalities of change, and empirical studies too rarely justify their chosen time lags. Greater theoretical precision regarding timescale—and closer alignment between model structure and theoretical assumptions—may be essential to resolve longstanding inconsistencies in the literature and advancing explanatory models.

Refining the temporal structure of psychological models also offers a promising path toward theoretical integration. As proposed in the CL-REM, different reciprocal effects may be driven by distinct mediation mechanisms operating on different timescales. Although the present study focused on direct effects between self-concept and skills, future research could incorporate these intermediate mechanisms and complete the empirical exploration of the broader CL-REM framework (Fig. 1). For instance, microlevel skill-development effects (i.e., when recent academic performance shapes self-concept) likely reflect fast processes such as achievement emotions (Pekrun et al. 2023; Pekrun 2024a) and teacher feedback (Hattie & Timperley, 2007; Mandouit & Hattie, 2023). Success in an exam or praise from a teacher can quickly boost perceived competence by reinforcing a sense of control and value, and clarifying learning goals and purpose. In contrast, mesolevel self-enhancement effects (i.e., when self-concept shapes later skill development) may reflect slower processes such as motivational regulation (Bardach & Murayama, 2025; León et al., 2025; Ryan & Deci, 2020) and student engagement (Fredricks et al., 2004; Skinner et al., 2008), which require sustained experiences to modify involvement with the learning tasks.

Recognizing these differentiated timescales, the CL-REM helps integrate distinct theories (control-value theory, teacher feedback, student engagement, motivational regulations) along a common temporal continuum. Rather than offering competing views of the same outcomes, each theory can specify its focal timescale and contribute to an integrated framework of learning dynamics. This perspective echoes recent calls to move beyond theory fragmentation in psychology, by fostering structured complementarities across models (here, the timescale continuum) rather than perpetuating conceptual silos (Pekrun 2024b).

The identification of distinct timescales also carries practical implications for intervention design (for details, see SM-F). For example, the results on slow self-enhancement effect suggest that interventions targeting self-concept may be genuinely beneficial to learning and skill development, but only if sustained and assessed over long periods of time (semester or year-long window). Psychological supports that help students reframe difficulty as a challenge, underscore progress, or receive encouragement may foster self-perceptions that serve as motivational scaffolds for learning (Walton & Wilson, 2018; Yeager & Walton, 2011). Conversely, the results on fast skill-development effect align with instructional designs that sequence tasks by difficulty, reduce distractions, or provide explicit feedback to enhance skills acquisition (Sweller et al., 2019; Wisniewski et al., 2020); such skill acquisition, in turn, may generate immediate benefits in perceived competence and, in the long run, increased acquisitions, thereby triggering positive reinforcement cycles. Importantly, according to the CL-REM framework, these reinforcement cycles should unfold asynchronously: Effects of self-concept on skills are slow, effects of skills on self-concept are fast. This has direct relevance for thinking clearly about the timing and dosage of interventions within a CL-REM approach.

Conclusion

By replicating Marsh et al. (2024), this study provides additional evidence for a temporally structured pattern of predictive effects between academic self-concept and skills acquisition. The classic question “Do students acquire skills because they feel competent or feel competent because they are skilled?” is often interpreted as implying mutually exclusive explanations for learning behavior. Yet our results are consistent with the view that these processes are not contradictory but, rather, that they unfold on different timescales. Recognizing this temporal structure may be key to shaping not only how we study learning behavior, but also how we support it.

Appendix

Fit Indices of Alternative Models of Self-concept and Skills Co-evolutions

Model	χ^2	DF	CFI	TLI	RMSEA	SRMR	wAIC (AIC)	wBIC (BIC)
Mathematics								
RI-CCLPM (CL-REM)	**1105.2**	**176**	**0.983**	**0.980**	**0.024**	**0.056**	**0.473 (412153.5)**	**0.961 (412694.7)**
CLPM	1330.2	178	0.979	0.975	0.027	0.062	0 (412413.7)	0 (412940.6)
RI-CLPM	1105.1	175	0.983	0.980	0.024	0.056	0.261 (412154.7)	0.015 (412703.0)
RI-CPM	1104.5	175	0.983	0.980	0.024	0.056	0.261 (412154.7)	0.015 (412703.0)
STARTS	1117.9	176	0.983	0.980	0.024	0.057	0.005 (412162.8)	0.009 (412704.0)
RI-LTVC	1214.9	177	0.981	0.978	0.025	0.059	0 (412263.7)	0 (412797.7)
LCSM	2867.7	180	0.951	0.943	0.040	0.131	0 (414273.2)	0 (414785.9)
LCSM-C	3821.1	181	0.934	0.923	0.047	0.132	0 (415425.7)	0 (415931.3)
French								
RI-CCLPM (CL-REM)	**1295.3**	**177**	**0.965**	**0.958**	**0.026**	**0.063**	**0.106 (410174.8)**	**0.690 (410708.0)**
CLPM	1339.2	178	0.963	0.957	0.027	0.065	0 (410229.4)	0 (410755.5)
RI-CLPM	1292.0	175	0.965	0.958	0.027	0.063	0.040 (410176.8)	0 (410724.2)
RI-CPM	1292.0	175	0.965	0.958	0.027	0.063	0.040 (410176.8)	0 (410724.2)
STARTS	1291.4	175	0.965	0.958	0.027	0.063	0.766 (410170.9)	0.004 (410718.3)
RI-LTVC	1297.8	177	0.965	0.958	0.026	0.063	0.047 (410176.5)	0.305 (410709.7)
LCSM[A]	1799.3	179	0.949	0.940	0.032	0.067	0 (410746.9)	0 (411265.9)
LCSM-C[A]	2648.6	180	0.922	0.909	0.039	0.079	0 (411750.0)	0 (412261.9)

N=9142 (mathematics), 9041 (French). *CLPM* Cross-lagged panel model, *RI-CLPM* random intercept CLPM, *RI-CPM* random intercept contemporaneous panel model, *RI-CCLPM* RI-CLPM with contemporaneous effects of skills, *STARTS* stable trait, autoregressive trait and state model with contemporaneous effects (or RI-CCLPM with measurement error), *RI-LTVC* random intercept latent third-variable confounder, *LCSM* latent change score model, *LCSM-C* latent change score model with contemporaneous effects. Models in bold font (*RI-CCLPM*) correspond to the best-fitting solutions identified based on model converge, fit indices, and parameter significance. Parameter estimates for the RI-CCLPM are reported in Fig. 3. Full parameter values (estimates, standard error, Wald test, p-value, 95% confidence intervals) are available on the Supplementary Tables S10 to S27 (Statistics sheets). Model scripts are available online (see Núñez-Regueiro et al., 2025)

[A] These models did not reach proper solutions (non-positive definite variance-covariance matrices), even after applying constraints to obtain identified solutions

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10648-026-10131-w.

Authors' contributions F.N.R: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing-Original Draft, Writing- Review and Editing, Visualization. H.M.: Conceptualization, Validation, Writing-Original Draft, Writing- Review and Editing. P.B.: Conceptualization, Methodology, Investigation, Writing- Review and Editing, Project administration, Funding acquisition. A.B. to E.P.V.: Conceptualization, Methodology, Investigation, Project administration, Funding acquisition. P.P.: Conceptualization, Methodology, Investigation, Writing- Review and Editing, Project administration, Funding acquisition.

Funding This research was supported financially by the French Ministry of National Education, Youth and Sports (MENJS); the Ministry of Higher Education, Research and Innovation (MESRI); "Mission Monteil pour le numérique éducatif", and the "Programme d'investissements d'avenir, expérimentation ProFAN" (PIA).

Declarations

Conflict of interest The authors report no conflict of interest.

References

Bailey, D. H., Jung, A. J., Beltz, A. M., Eronen, M. I., Gische, C., Hamaker, E. L., Kording, K. P., Lebel, C., Lindquist, M. A., Moeller, J., Razi, A., Rohrer, J. M., Zhang, B., & Murayama, K. (2024). Causal inference on human behaviour. *Nature Human Behaviour, 8*(8), 1448–1459. https://doi.org/10.1038/s41562-024-01939-z

Bandura, A. (1982). The assessment and predictive generality of self-percepts of efficacy. *Journal of Behavior Therapy and Experimental Psychiatry*, *13*(3), 195–199.

Bardach, L., & Murayama, K. (2025). The role of rewards in motivation—Beyond dichotomies. *Learning and Instruction, 96*, Article 102056. https://doi.org/10.1016/j.learninstruc.2024.102056

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. Wiley.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford.

Burnham, K. P., & Anderson, D. R. (Eds.). (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer. https://doi.org/10.1007/b97636

Burns, R. A., Crisp, D. A., & Burns, R. B. (2020). Re-examining the reciprocal effects model of self-concept, self-efficacy, and academic achievement in a comparison of the Cross-Lagged Panel and Random-Intercept Cross-Lagged Panel frameworks. *British Journal of Educational Psychology, 90*(1), 77–91. https://doi.org/10.1111/bjep.12265

Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology, 69*(2), 136–145. https://doi.org/10.1037/0022-0663.69.2.136

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.

Ehm, J.-H., Hasselhorn, M., & Schmiedek, F. (2019). Analyzing the developmental relation of academic self-concept and achievement in elementary school children: Alternative models point to different results. *Developmental Psychology, 55*(11), 2336–2351. https://doi.org/10.1037/dev0000796

Fabrigar, L. R., & Wegener, D. T. (2016). Conceptualizing and evaluating the replication of research results. *Journal of Experimental Social Psychology, 66*, 68–80. https://doi.org/10.1016/j.jesp.2015.07.009

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Graham, J. W. (2012). *Missing data: Analysis and design*. Springer Science & Business Media.

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in M plus. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hamaker, E. L. (2023). The within-between dispute in cross-lagged panel research and how to move forward. *Psychological Methods*. https://doi.org/10.1037/met0000600

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102–116. https://doi.org/10.1037/a0038889

Harter, S. (1988). *Manual for the self-perception profile for adolescents*. University of Denver.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*(5), 505–528. https://doi.org/10.1016/j.jsp.2011.07.001

Hübner, N., Wagner, W., Zitzmann, S., & Nagengast, B. (2023). How strong is the evidence for a causal reciprocal effect? Contrasting traditional and new methods to investigate the reciprocal effects model of self-concept and achievement. *Educational Psychology Review, 35*(1), Article 6. https://doi.org/10.1007/s10648-023-09724-6

Hunt, D. E. (1975). Person-environment interaction: A challenge found wanting before it was tried. *Review of Educational Research, 45*(2), 209–230. https://doi.org/10.2307/1170054

Kenny, D. A., & McCoach, D. B. (2025). Ruling out latent time-varying confounders in two-variable multi-wave studies. *Multivariate Behavioral Research, 0*(0), 1–17. https://doi.org/10.1080/00273171.2025.2503829

Kenny, D. A., & Zautra, A. (1995). The trait-state-error model for multiwave data. *Journal of Consulting and Clinical Psychology, 63*(1), 52.

Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M. M., Moutoussis, M., Goodyer, I. M., Bullmore, E., Jones, P. B., Fonagy, P., Lindenberger, U., & Dolan, R. J. (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience, 33*, 99–117. https://doi.org/10.1016/j.dcn.2017.11.007

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

León, J., Núñez-Regueiro, F., & Santana-Monagas, E. (2025). Reciprocal relations between teacher engaging messages and student motivation. *Revista de Psicodidáctica (English ed.),* , Article 500165. https://doi.org/10.1016/j.psicoe.2025.500165

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

Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling: A Multidisciplinary Journal, 29*(6), 888–907. https://doi.org/10.1080/10705511.2022.2065278

Mandouit, L., & Hattie, J. (2023). Revisiting "*The Power of Feedback*" from the perspective of the learner. *Learning and Instruction, 84*, Article 101718. https://doi.org/10.1016/j.learninstruc.2022.101718

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., & Hau, K.-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology, 32*(1), 151–170. https://doi.org/10.1016/j.cedpsych.2006.10.008

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist, 34*(3), 155–167. https://doi.org/10.1207/s15326985ep3403_2

Marsh, H. W., Guo, J., Pekrun, R., Lüdtke, O., & Núñez-Regueiro, F. (2024). Cracking chicken-egg conundrums: Juxtaposing contemporaneous and lagged reciprocal effects models of academic self-concept and achievement's directional ordering. *Educational Psychology Review, 36*(2), Article 53. https://doi.org/10.1007/s10648-024-09887-w

Monteil, J., Séré, A., Demolliens, M., Huguet, P., Batruch, A., Bouet, M., Bressan, M., Bressoux, P., Brown, G., Butera, F., Cepeda, C., Cherbonnier, A., Darnon, C., de Place, A. L., Desrichard, O., Ducros, T., Goron, L., Hemon, B., Jamet, E., & Vives, E. (2022). *Partager l'expertise: L'interdépendance positive, un levier pour de nouvelles compétences ?* [Report]. Ministère de l' Education nationale. https://hal.science/hal-04021297

Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(4), 638–648. https://doi.org/10.1080/10705511.2020.1784738

Murayama, K., & Gfrörer, T. (2024). Thinking clearly about time-invariant confounders in cross-lagged panel models: A guide for choosing a statistical model from a causal inference perspective. *Psychological Methods*. https://doi.org/10.1037/met0000647

Murayama, K., & von Keyserlingk, L. (2025). A critical analysis of the current motivation theories in educational psychology: Why the same theories continue to dominate. *Educational Psychologist*(1). https://doi.org/10.1080/00461520.2025.2473894

Muthén, B., & Asparouhov, T. (2024). Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The case of contemporaneous and reciprocal effects. *Psychological Methods*. https://doi.org/10.1037/met0000661

Muthén, L., & Muthén, B. (2022). *Mplus Version 8.8 user's guide*. Muthén & Muthén.

Muthén, B., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25*, 267–316. https://doi.org/10.2307/271070

Núñez-Regueiro, F., & Wang, M. T. (2024). Adolescent well-being and school engagement as a function of teacher and peer relatedness: The more (relatedness) is not always the merrier. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000910. Advance online publication.

Núñez-Regueiro, F., Juhel, J., Bressoux, P., & Nurra, C. (2022). Identifying reciprocities in school motivation research: A review of issues and solutions associated with cross-lagged effects models. *Journal of Educational Psychology, 114*(5), 945–965. https://doi.org/10.1037/edu0000700

Núñez-Regueiro, F., Marsh, H. W., Pekrun, R., Lüdtke, O., & Guo, J. (2025). The reciprocal effects model is robust to alternative modeling specifications: A response to Sorjonen et al., 2025. *Educational Psychology Review, 37*(3), Article 79. https://doi.org/10.1007/s10648-025-10059-7

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. J. Riding, & S. G. Rayner (Eds.), *Self perception* (Vol. 2, pp. 239–266). Ablex Publishing.

Pekrun, R. (2024a). Control-value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review*, *36*(3), 1–36. https://doi.org/10.1007/s10648-024-09909-7

Pekrun, R. (2024b). Overcoming fragmentation in motivation science: Why, when, and how should we integrate theories? *Educational Psychology Review*, *36*(1), 27. https://doi.org/10.1007/s10648-024-09846-5

Pekrun, R., Hofe, V., Blum, R., Frenzel, W., Götz, A. C., T., & Wartha, S. (2007). Development of mathematical competencies in adolescence: The PALMA longitudinal study. In M. Prenzel (Ed.), *Studies on the educational quality of schools: The final report on the DFG priority programme* (pp. 17–37). Waxmann. https://kops.uni-konstanz.de/bitstreams/12dd3457-45e4-4f91-a43f-493515a52e95/download

Pekrun, R., Marsh, H. W., Suessenbach, F., Frenzel, A. C., & Goetz, T. (2023). School grades and students' emotions: Longitudinal models of within-person reciprocal effects. *Learning and Instruction, 83*, Article 101626. https://doi.org/10.1016/j.learninstruc.2022.101626

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *The Journal of Applied Psychology, 90*(1), 175–181. https://doi.org/10.1037/0021-9010.90.1.175
Riant, M., de Place, A.-L., Bressoux, P., Batruch, A., Bouet, M., Bressan, M., Brown, G., Butera, F., Cepeda, C., Cherbonnier, A., Darnon, C., Demolliens, M., Desrichard, O., Ducros, T., Goron, L., Hémon, B., Huguet, P., Jamet, E., Martinez, R., & Pansu, P. (2024). Does the jigsaw method improve motivation and self-regulation in vocational high schools? *Contemporary Educational Psychology, 77*, Article 102278. https://doi.org/10.1016/j.cedpsych.2024.102278
Rudmann, O., Batruch, A., Visintin, E. P., Sommet, N., Bressoux, P., Darnon, C., Bouet, M., Bressan, M., Brown, G., Cepeda, C., Cherbonnier, A., Demolliens, M., De Place, A.-L., Desrichard, O., Ducros, T., Goron, L., Hemon, B., Huguet, P., Jamet, E., & Butera, F. (2024). Cooperative learning reduces the gender gap in perceived social competences: A large-scale nationwide longitudinal experiment. *Journal of Educational Psychology, 116*(6), 903–920. https://doi.org/10.1037/edu0000870
Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860
Shum, A., & Fryer, L. K. (2026). Examining short-term self-efficacy change connections with performance, perceived difficulty, and interest. *Learning and Instruction, 101*, Article 102223. https://doi.org/10.1016/j.learninstruc.2025.102223
Skinner, E. A., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781. https://doi.org/10.1037/a0012840
Sorjonen, K., Melin, B., & Nilsonne, G. (2025). Inconclusive evidence for a prospective effect of academic self-concept on achievement: A simulated reanalysis and comment on Marsh et al. (2024). *Educational Psychology Review, 37*(2), Article 30. https://doi.org/10.1007/s10648-025-10008-4
Speyer, L. G., Zhu, X., Yang, Y., Ribeaud, D., & Eisner, M. (2024). On the importance of considering concurrent effects in random-intercept cross-lagged panel modelling: Example analysis of bullying and internalising problems. *Multivariate Behavioral Research, 0*(0), 1–17. https://doi.org/10.1080/00273171.2024.2428222
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5
Tobias, S., & Duffy, T. M. (2009). *Constructivist instruction: Success or failure?* Routledge.
Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods, 24*(5), 637–657. https://doi.org/10.1037/met0000210
Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist, 39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3
VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: A new template for empirical studies. *Statistical Science, 35*(3), 437–466. https://doi.org/10.1214/19-STS728
Vu, T., Scharmer, A. L., van Triest, E., van Atteveldt, N., & Meeter, M. (2024). The reciprocity between various motivation constructs and academic achievement: A systematic review and multilevel meta-analysis of longitudinal studies. *Educational Psychology, 44*(2), 136–170. https://doi.org/10.1080/01443410.2024.2307960
Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11*(1), 192–196. https://doi.org/10.3758/BF03206482
Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review, 125*(5), 617–655. https://doi.org/10.1037/rev0000115
Wesarg-Menzel, C., Ebbes, R., Hensums, M., Wagemaker, E., Zaharieva, M. S., Staaks, J. P. C., van den Akker, A. L., Visser, I., Hoeve, M., Brummelman, E., Dekkers, T. J., Schuitema, J. A., Larsen, H., Colonnesi, C., Jansen, B. R. J., Overbeek, G., Huizenga, H. M., & Wiers, R. W. (2023). Development and socialization of self-regulation from infancy to adolescence: A meta-review differentiating between self-regulatory abilities, goals, and motivation. *Developmental Review, 69*, Article 101090. https://doi.org/10.1016/j.dr.2023.101090
Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10–18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2019.03087
Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301. https://doi.org/10.3102/0034654311405999

Authors and Affiliations

Fernando Núñez-Regueiro[1] · Herbert W. Marsh[2,3] · Pascal Bressoux[1] · Anatolia Batruch[4] · Marinette Bouet[5] · Marco Bressan[6] · Genavee Brown[7] · Fabrizio Butera[4] · Anthony Cherbonnier[7] · Céline Darnon[8] · Marie Demolliens[8] · Anne-Laure de Place[1] · Olivier Desrichard[9] · Luc Goron[7] · Brivael Hémon[7] · Pascal Huguet[8] · Eric Jamet[7] · Vincent Mazenod[5] · Nathalie Mella[9] · Estelle Michinov[7] · Nicolas Michinov[7] · Nana Ofosu[9] · Laurine Peter[7] · Céline Poletti[6] · Isabelle Régner[6] · Mathilde Riant[1] · Anaïs Robert[8] · Ocyna Rudmann[4] · Camille Sanrey[1] · Arnaud Stanczak[8] · Farouk Toumani[5] · Emilio Paolo Visintin[4] · Pascal Pansu[1]

✉ Fernando Núñez-Regueiro
fernando.nr.france@gmail.com

Herbert W. Marsh
herb.marsh@acu.edu.au

Pascal Bressoux
pascal.bressoux@univ-grenoble-alpes.fr

Anatolia Batruch
anatolia.batruch@unil.ch

Marinette Bouet
marinette.bouet@uca.fr

Marco Bressan
marco.bressan@univ-amu.fr

Genavee Brown
genaveebrown@gmail.com; genavee.brown@northumbria.ac.uk

Fabrizio Butera
fabrizio.butera@unil.ch

Anthony Cherbonnier
anthony.cherbonnier@univ-lille.fr

Céline Darnon
celine.darnon@uca.fr

Marie Demolliens
marie.demolliens@uca.fr

Anne-Laure de Place
anne-laure.de.place@univ-paris8.fr

Olivier Desrichard
olivier.desrichard@unige.ch

Luc Goron
luc.goron@univ-rennes2.fr

Brivael Hémon
brivael.hemon@univ-tlse2.fr

Pascal Huguet
pascal.huguet@uca.fr

Eric Jamet
eric.jamet@univ-rennes2.fr

Vincent Mazenod
vincent.mazenod@isima.fr

Nathalie Mella
nathalie.mella-barraco@unige.ch

Estelle Michinov
estelle.michinov@univ-rennes2.fr

Nicolas Michinov
nicolas.michinov@univ-rennes2.fr

Nana Ofosu
nana.ofosu@unige.ch

Laurine Peter
laurine.peter@gmail.com

Céline Poletti
celine.poletti@unil.ch

Isabelle Régner
isabelle.regner@univ-amu.fr

Mathilde Riant
mathilde.riant@univ-grenoble-alpes.fr

Anaïs Robert
anais.robert@uca.fr

Ocyna Rudmann
ocyna.rudmann@unil.ch

Camille Sanrey
camille.sanrey@unistra.fr

Arnaud Stanczak
stanczak@unistra.fr

Farouk Toumani
farouk.toumani@isima.fr

Emilio Paolo Visintin
emiliopaolo.visintin@unife.it

Pascal Pansu
pascal.pansu@univ-grenoble-alpes.fr

[1] Université Grenoble Alpes, LaRAC, 38000, Grenoble, France

[2] Institute for Positive Psychology and Education (IPPE), Australian Catholic University, Sydney, Australia

[3] Department of Education, University of Oxford, Oxford, United Kingdom

[4] Laboratoire de Psychologie Sociale de l'Université de Lausanne (UNILaPS), University of Lausanne, Lausanne, Switzerland

[5] Université Clermont Auvergne, CNRS, LIMOS, F-63000, Clermont-Ferrand, France

[6] Aix Marseille Univ, CNRS, CRPN, Marseille, France

[7] Univ Rennes, LP3C (Laboratoire de Psychologie : Cognition, Comportement, Communication), F-35000, Rennes, France

[8] Université Clermont Auvergne, CNRS, LAPSCO, F-63000, Clermont-Ferrand, France

[9] Groupe de Recherche en Psychologie de la Santé (GREPS), University of Geneva, Geneva, Switzerland