Abstract

Studies and models on academic self-concept (ASC) have mostly relied on its stability over time, but recent research advancements on individual differences has shown that the majority of psychological constructs tend to change over time. Drawing on literature regarding personality trait change and the RI/EM, we conducted a study aimed at investigating characteristics of the changeability of math and verbal self-concepts across junior high school, as well as examined their relationships with academic performance. The sample consisted of 1674 students, who filled in a self-report questionnaire on math and verbal self-concept at T1 (10yrs) and T2 (13yrs), whereas math and verbal achievement at T1 and T2 were measured by standardized test scores. Results attested (a) that both math and verbal self-concept, on average, decrease significantly over the course of junior high school; (b) that a large variability exists in the way students change; (c) that the way students change in one academic self-concept is not related to changes in the other academic self-concept. In regards to academic achievement, we found reciprocal positive longitudinal effects in matching domains and low-positive or non-significant longitudinal relationships in non-matching domains. In sum, the ability to contrast the overall negative trend of ASCs is associated with amelioration in academic achievement at the end of junior high school. From a practical standpoint, these findings suggest the importance of (a) assessing and intervening on ASCs during junior high school; (b) intervening in math and verbal self-concept separately; (c) taking into account student's own way of changing.

Keywords: academic self-concept, academic achievement, latent change score models, reciprocal internal/external frame-of-reference model (RI/EM)

Academic self-concept change in junior high school students and relationships with academic achievement

Self-concept is a multifaceted and hierarchical construct (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2006; O'Mara, Marsh, Craven, & Debus, 2006; Shavelson, Hubner, & Stanton, 1976) broadly defined as "a person's self-perceptions formed through experience with and interpretations of one's environment" (Marsh & Martin, 2011, p. 60). Self-perceptions related to self-concept include feelings of self-worth, competence, self-confidence, ability and self-acceptance. Hence, one's self-concept is significantly influenced by attributions for one's behavior, by reinforcements and by the evaluations of significant others. Unlike other self-related constructs (e.g., self-esteem and self-efficacy beliefs), self-concept requires a self-evaluation of competences (a) in a specific and restricted domain and (b) largely based on past circumstances and accomplishments (Marsh et al., 2019).

Self-concept is implicated in various psychological processes such as self-regulation of behavior (e.g., Carver & Scheier, 1998), motivation (e.g., Eccles & Wigfield, 2002), and emotions (e.g., Higgins, 1987; Markus, 1977; Markus & Kitayama, 1991). It is an organized structure of knowledge that influences processing information relevant to the individual and affects behavior (e.g., Hattie & Marsh, 1996; Markus, 1977; Markus & Wurf, 1987) in various contexts such as social, emotional, and educational contexts (e.g., Chen, Yeh, Hwang, & Lin, 2013; Harter, 2012; Marsh & Craven, 2006; Pekrun et al., 2019). The relevance of positive self-concepts has been particularly highlighted in educational settings (e.g., Marsh & Craven, 2006; Valentine et al., 2004). Indeed, positive academic self-concepts are not only desirable goals, but also a means of facilitating subsequent learning and various academic outcomes (Marsh & Scalas, 2010).

Academic self-concepts (e.g., math self-concept) are currently among the most studied positive self-beliefs in Educational Psychology, given that they have been shown to be positively and strongly related to several academic outcomes (Huang, 2011; Marsh et al., 2019; see Trautwein & Möller, 2016, for a review), with consistency of findings attested across individual, group, and county levels (Marsh et al., 2020). The positive effect of academic self-concept on relevant educational outcomes has been highlighted in the literature (e.g., Marsh & Scalas, 2010; Quílez-Robres et al., 2021; Valentine et al., 2004). For example, within the Expectancy-Value Theory framework (e.g., Eccles, 2009; Wigfield & Eccles, 2000), which posits that a person's expectancies of success in a given task in combination with that person's valuing of that task (i.e., task value) are key predictors of academic achievement, effort, engagement, and career choices (e.g., Eccles, 2009; Guo et al., 2015; Watt et al., 2012). Expectancy is often operationalized as academic self-concept in educational research (Eccles & Wigfield, 1995, 2002) and when both expectancy and task value are used as predictors of academic achievement, expectancy results as the strongest predictor of achievement in various school domains (e.g., Eccles & Wigfield, 2002; Scalas & Fadda, 2019; Trautwein et al., 2012). Also, academic self-concept showed positive effects on academic achievement over and above the effect of other relevant variables in the educational context, such as interest (Marsh, et al., 2005) and mastery and performance goals (Plante et al., 2013; Scalas & Fadda, 2019). A number of studies addressed the inter-relationship between academic self-concept, values, and achievement, at both cross-sectional and longitudinal level. Using mixture modeling, Archambault, Eccles, and Vida (2010) found 7 different joint trajectories for *Reading/English Ability Self-Concept* and *Literacy Subjective Task Value* across 1 to 12 grade. Of interest, much of those trajectories implied a negative trend, with non-linear trajectories that increased only when levels of self-concept and values were below the average mean (see Archambault et al., 2010, p. 810). Jacobs et al. (2002), using a three-cohort six-wave dataset spanning from 1 to 12 grade, studied trajectories of change in competence beliefs and values in different domains (math, language arts, and sport) and concluded that "the most striking finding across all domains was that self-perceptions of competence and subjective task values declined as children got older" (Jacobs et al., 2002, p. 509). Gaspard et al. (2018), using a cross-sectional dataset of 5-12 grade students, investigated the effect of achievement to expectancy/value in 5 domains (German, English, Biology, Physics, and Math) and found a negative effect between "far" domains (e.g., a negative effect of math achievements on languages expectancy/value) and a positive effect on “near” domains (e.g., a positive effect exerted by math achievement on physics expectancy/value). That said, the widespread interest of practitioners and researchers in academic self-concept across the school years is understandable.

However, there are still open issues, in particular regarding (a) the way in which students change in academic self-concept levels across time, and (b) the relationship between academic self-concept change and academic achievement. This is mostly due to the paucity of longitudinal studies specifically devoted to the study of academic self-concept stability and change, as highlighted by recent contributions (e.g., Marsh et al., 2019, 2020). Indeed, while there are articles showing that academic competence beliefs have a negative trend across time (e.g., Archambault et al., 2010; Jacobs et al., 2002), the study of the “variance” around this change and its effect on academic achievement (net of stability) is yet to be explored. The study of construct stability and change provides useful information for interventions, such as direction, heterogeneity, and correlates of change (Bleidorn et al., 2019). This is particularly true in Educational research and practice, in which the aim “to understand how and why students change over time” (Grimm, Mazza, & Mazzocco, 2016, p. 342) is among the most important tasks. Thus, knowledge regarding self-concept change and knowledge regarding the relationship between academic self-concept change and achievement might help teachers in practical interventions with their pupils, particularly during early adolescence, when developmental challenges bring to a reorganization of self-concepts and often to a decline in various self-domains as a consequence of identity uncertainty due to biological, social, and cognitive changes (e.g., Harter, 2012). Indeed, corroborating that academic self-concept declines over the years and that this decline has an impact on academic achievement may call for more attention in students’ non-cognitive spheres when planning interventions.

Thus, in this scenario, we have to consider the "change" as a variable in itself (measuring the degree of change across time points), having (a) a *mean* and a (b) *variance*, and - consequently - (c) the possibility to *affect other constructs*. Hence, our study aims at shedding light on the above-mentioned unresolved questions, such as whether math and verbal self-concept are stable constructs or tend to change (the *mean* of the change), whether there is substantial heterogeneity in students' academic self-concept over time (the *variance* of the change), whether changes in verbal and math self-concept are related, and what is the relationship between academic self-concept change and academic achievement (the possibility to *affect other constructs*). Such an examination would encourage practitioners and teachers to consider "personal changes" in ASC rather than examining only ASC differences with respect to others. For example, analyses on personal change in ASC may reveal a declining trend during a particular school phase (e.g., junior high school) and thus suggesting the implementation of preventive interventions capable of countering this decline. Furthermore, a focus on individual change may indicate who may benefit more of such interventions (e.g., students who are high in ASC when compared to others but who have experienced a personal decline during the last year or months). Finally, this focus may also reveal which could be the consequences of a mean-level change (e.g., an effect to performance).

Thus, drawing on both recent advances in the study of individual difference changes (e.g., Anusic & Schimmack, 2016; Bleidorn et al., 2019; Grimm & Ram, 2018; McArdle, 2009; see Wagner et al., 2020, for an accessible and comprehensive conceptual and theoretical discussion) and utilizing the reciprocal internal/external frame-of-reference model (RI/EM; Sewasew & Schroeders, 2019), we conducted a longitudinal study in which we investigated (a) the degree and the direction in which verbal and math self-concept varied over time, (b) their relationship at both the initial and change level and (c) their longitudinal relationship with matching and non-matching academic achievement domains. To do so, we adopted a Latent Change Score framework, which is a latent variable approach (Bollen, 1989; Kline, 2016) widely used to investigate the stability and change of a construct (as well as their predictors, outcomes and correlates) after taking into account the measurement instrument's degree of unreliability (Ferrer, Boker, & Grimm, 2019; Klopack & Wickrama, 2020; McArdle, 2009; McArdle & Nesselroade, 2014). Hence, we adopted a substantive-methodological synergy (Borsboom, 2006) in order to advance research on academic self-concepts (in particular, on their change and their relationship with academic achievement), since to our knowledge, no study has explored these issues through a proper technique, and given that we only found four studies that tested the RI/EM. Shedding light on academic self-concept change and its relationship with academic achievement may provide information on the value of intervening on academic self-concept during specific school years. Specifically, findings may suggest when to intervene on academic self-concept, in which way, and what might be the consequences of change. Following, we first outline the state-of-the-art research regarding the relationship between academic self-concepts and academic achievement, with particular attention to the RI/EM; we then describe why the study of longitudinal change in academic self-concepts may advance our understanding of this field; finally, we present our contribution in more detail.

Academic Self-Concept and Academic Achievement

From a theoretical perspective, self-concept is a hierarchical construct (O'Mara et al., 2006; Shavelson et al., 1976; Trautwein & Möller, 2016). At the apex of this hierarchy lies *global self-concept*, which can be divided into *non-academic self-concept* (e.g., emotional, social and physical) and *academic self-concept* (Arens et al., 2021). The latter can be further disentangled into self-concept in specific academic domains, such as *math self-concept* and *verbal self-concept* (Marsh & Hau, 2004). While math and verbal achievements are assumed to be correlated, research has surprisingly shown that math and verbal self-concepts tend to be poorly correlated (Marsh, 1986; Marsh & Hau, 2004; Marsh et al., 2019). As a consequence, the relationship between academic self-concepts and academic achievement depends on whether domains are matching or non-matching (e.g., Möller et al., 2020; Wan et al., 2021). The literature offers two main theories that seek to understand the underpinnings of the relationship between academic self-concepts and academic achievement: The internal/external frame-of-reference (I/E) model and the reciprocal effect model (REM). In recent years, the I/E model and REM models have been merged into a unique comprehensive framework, the reciprocal internal/external frame-of-reference model (RI/EM). In what follows, we provide an overview of these models.

The Internal/External Frame-of-Reference (I/E) Model

The internal/external frame-of-reference (I/E) model (Marsh, 1986) postulates that one's self-concept in a specific school subject is shaped by both an *external* and an *internal* reference: The former concerns the comparison between one's own performance in a specific school subject and the corresponding performance of other students; the latter regards the comparison between one's own performance in a specific school subject and one's own corresponding performance in other school subjects (Marsh & Hau, 2004). According to this theory, both processes affect academic self-concept. Thus, the joint operation of external and internal processes (which depends on the relative weight given to external and internal comparisons) would lead to correlations between math and verbal self-concepts that are substantially lower than the typical correlations found between math and verbal achievements (see Marsh & Hau, 2004). Furthermore, according to this theory, academic achievement is a positive predictor of academic self-concept when the latter regards the same (matching) domain and a negative predictor of academic self-concept when the latter refers to a different (non-matching) domain.

The I/E model has been largely supported by empirical research. An extensive cross-cultural study by Marsh and Hau (2004) conducted in 26 countries (N = 55,577) found (a) high correlations between math and verbal achievements, but low correlations between math and verbal self-concepts, (b) positive effects of math achievement on math self-concept and negative effects of math achievement on verbal self-concept and (c) positive effects of verbal achievement on verbal self-concept and negative effects of verbal achievement on math self-concept. A meta-analytic path analysis by Möller, Pohlmann, Köller, and Marsh (2009) found (a) a high correlation between math and verbal achievements (ρ = .67), (b) a positive effect of academic achievement to matching self-concept (ρ = .61 for math; ρ = .49 for verbal) and (c) a negative effect of academic achievement on non-matching self-concept (math achievement → verbal self-concept: β = -.21; verbal achievement → math self-concept: β = -.27). Finally, a recent longitudinal study has further corroborated the validity of the I/E model by taking into account the effect of another important positive self-belief, namely, self-efficacy (Marsh et al., 2019).

The Reciprocal Effect Model (REM)

As shown above, the I/E model is a widely accepted theory in the realm of academic self-concept research. However, it should be noted that the I/E model is silent about the direction of causation between academic self-concepts and academic achievement. Hence, whether academic achievement exerts a stronger influence on academic self-concept than vice versa or whether both have the same size of impact on the other is still debated. In this regard, there are three main positions (see Huang, 2011): The *skill-development model,* the *self-enhancement model* and the *reciprocal effect model*. According to the *skill-development model,* academic achievement affects self-concept and not vice versa, whereas the *self-enhancement model* maintains that self-concept is a significant predictor of academic achievement and not vice versa. However, most longitudinal studies (Arens et al., 2017; Guay, Marsh, & Boivin, 2003; Hoge, Smit, & Crist, 1995; Marsh et al., 2016; Marsh & Yeung, 1997), quantitative meta-analyses (Huang, 2011; Valentine et al., 2004) and reviews (Marsh & Craven, 2006; Marsh & Martin, 2011) support the *reciprocal effect model* (REM), in which academic self-concept and academic achievement "are reciprocally related and mutually reinforcing" (Marsh & Martin, 2011, p. 72).

Integrating the I/E model and the REM: The reciprocal internal/external frame-of-reference model (RI/EM)

Although the I/E model and the REM have been considered two separate models of academic self-concept for many years, they are not mutually exclusive. Indeed, recently some authors have combined the two models into one comprehensive framework, namely the *reciprocal internal/external frame-of-reference model* (RI/EM; Möller et al., 2011, 2014; Niepel, Brunner, & Preckel, 2014; Sewasew & Schroeders, 2019). Accordingly, "the RI/EM postulates positive developmental relations between academic achievement and self-concept within a domain and negative relations across two non-matching domains" (Sewasew & Schroeders, 2019, p. 204).

However, it has been showed that with a long time-lag (e.g., 4 years), the negative cross-correlation may result non-significant or very low in magnitude (Möller et al., 2014). In Figure 1A (see the Appendix), we provide a graphical representation of the relationships between academic self-concepts and achievement based on the RI/EM.

The RI/EM provides a significant step forward in educational research, given that it "focuses on the interplay between students' academic self-concepts and their academic achievement across multiple domains" (Niepel et al., 2014, p. 1186). However, we have noted some gaps in the literature that still need to be addressed. First of all, we have found only four empirical contributions that combine the I/E model and the REM into the RI/EM (Möller et al., 2011, 2014; Niepel et al., 2014; Sewasew & Schroeders, 2019). Moreover, we noted that two of these studies (Möller et al., 2011; Niepel et al., 2014) did not use standardized test scores as achievement indicators; as Marsh et al. (2019) put it "in well-designed longitudinal studies of achievement, standardized tests are designed both to assess achievement with items that are age-appropriate, and to provide a standardized measure of achievement along a common metric by using overlapping sets of items." (Marsh et al., 2019, p. 348). Finally, and most importantly, none of the four studies took account of between-person differences in the within-person change process in academic self-concepts. The latter is a pivotal point to be considered when the aim is to investigate the consequences of enhancing academic self-concept. In what follows, we discuss why it is important to study both the stability and changeability of academic self-concept.

Studying Change Over Time in Academic Self-concept

Academic Self-Concept Change

The study of construct stability and variability is of pivotal importance when the focus is the "change" on a specific construct (Anusic & Schimmack, 2016; Hamaker, 2012) and thus on correlates, predictors or outcomes that are related to within-person change. Indeed, as Grimm, Mazza, and Mazzocco (2016) put it, "Educational researchers are primarily interested in how and why students change over time. In this endeavor, researchers wish to understand the key characteristics of those changes (e.g., shape and timing of those changes), the antecedents of change, and the consequences of change" (Grimm et al., 2016, p. 352). Of interest are contributions highlighting that academic self-concepts may demonstrate a decline across school years, in particular, during junior high school (Green et al., 2012; Hancock, Kuo, & Lawrence, 2001; Preckel, Niepel, Schneider, & Brunner, 2013). Indeed, according to Harter (2006), during junior high school, self-presentations "are likely to be unrealistically *positive* for several reasons" (Harter, 2006, p. 517), such as the lack of cognitive ability to engage in social comparison, the inability to distinguish between their *actual* and *ideal* self-attributes and a tendency to socially desirable responding. Therefore, the emergence of cognitive skills that enhance their ability to use social comparison for the purpose of self-evaluation and to differentiate *actual* and *ideal* self-evaluation "normatively leads many older children to realistically lower their self-evaluations [so that] realistic self-evaluations are more adaptive beginning in middle to late childhood, unlike in early childhood where an overestimation on one's capacities may have a positive motivational function" (Harter, 2006, p. 528). Hence, it would be interesting to examine whether this change has a negative trend and whether it varies across students. Indeed, a significant size in the way students change (technically speaking, a significantly large between-person differences in the within-person change process) would make it possible to find correlates, predictors and outcomes of the "change" factor, with estimates that may be substantially different from those obtained through research on rank-order consistency (see Bleidorn et al., 2019). In their review regarding the determinants and consequences of academic self-concept in school contexts, Trautwein and Möller (2016, paragraph 8.2.2) maintain the importance of studying both stability and malleability (or changeability) in self-concept, however, the lack of studies that use proper psychometric methodologies is evident.

Academic Self-Concept Change and Academic Achievement

The prospective relationship between academic self-concept and academic achievement has only been investigated through a cross-lagged panel model (CLPM; e.g., Sewasew & Schroeders, 2019). However, according to Usami et al. (2019), CLPM belongs to *cross-lagged models that do* *not explicitly model developmental trajectories,* hence leaving the issue of mean-level change in a construct unexplored. A recent study by Jansen, Lüdtke, and Robitzsch (2020) using STARTS found that academic self-concept exhibits significant variance for both stable and temporally changing factors. Thus, modeling the heterogeneity of students' change (i.e., between-person differences in within-person change; Grimm et al., 2016) is important for investigating the prospective relationships among academic self-concept and academic achievement. Indeed, the conceptual meaning of a cross-lagged coefficient in a CLPM is the "prospective effect of individual differences in Construct X on change in individual differences in Construct Y" (Orth et al., 2021, p. 1017). Hence, CLPM takes into account *rank-order change* (i.e., "how people change relative to one another on a trait over a certain period of time"; Bleidorn et al., 2019, p. 1058), but this model is silent regarding *mean-level change* (i.e., "how groups change on average on a trait over a certain period of time", Bleidorn et al., 2019, p. 1058) and *individual-level change* (i.e., "how individuals change differently than the group average over a certain period of time", Bleidorn et al., 2019, p. 1058). Thus, if we are interested in investigating (a) the group-average change in a construct (*mean-level change*); (b) the variability of the group-average change in a construct (between-person differences in within-person change, or simply *individual-level change*); and (c) the relationships of that change factor; then we should utilize cross-lagged models other than CLPM, such as Latent Change Score models.

Latent Change Score (LCS) models (Ferrer et al., 2019; McArdle, 2009; McArdle & Nesselroade, 2014) represent a well-established statistical framework that may be useful in advancing our understanding of academic self-concept. Recently, the LCS framework has been considered as one of the best latent variable approaches for conducting analyses focused on between-person differences in within-person change in educational research (see Grimm et al., 2016). A general LCS model allows researchers to operationalize the longitudinal change in a construct through a latent variable[1] (Δ Self-Concept in Figure 1), as well as to estimate its mean (representing the direction of change; see parameter k_2 in Figure 1), variance (representing between-person differences in the within-person change process; see parameter φ_{22} in Figure 1), covariance with the initial level (see parameter φ_{12} in Figure 1) and relationships with predictors, outcomes and correlates of interest (for a technical discussion of LCS models, see Grimm, Ram, & Estabrook, 2017). Thus, the overarching aim of our study was to contribute to the literature on academic self-concept stability and change: First, by investigating *mean-level* (k_2) and *individual-level* (φ_{22}) change in both math and verbal self-concept during junior high school (Harter, 2006); second, by investigating relationships among academic self-concept levels, academic self-concept changes and academic achievement.

The Present Study

In this study, we applied univariate and multivariate LCS models by means of a two-wave sample of 1674 students (aged 10 at T1 and 13 at T2), including measures of math and verbal self-concepts, as well as measures of math and verbal achievements. We aimed to answer some open questions that may advance the literature on academic self-concept and provide useful and practical information such as whether math and verbal self-concepts are stable constructs or tend to change, whether there is substantial heterogeneity in students' academic self-concept change, whether changes in verbal and math self-concepts are related, as well as highlight the relationships between academic self-concept change and academic achievement.

In more detail, at the *univariate level* (i.e., focusing only on academic self-concepts) we hypothesized that the extent to which students changed in their math and verbal self-concepts would significantly vary across students, according to recent perspectives on the high level of changeability of most psychological individual differences (e.g., Anusic & Schimmack, 2016;

[1] It is reasonable to argue that means are reported in descriptive statistics and thus, they should not be considered information that is unique to the latent change score model. Here, we point out that LCS models analyze the mean-level change through latent variables, while means reported in descriptive statistics usually refer to observed variables. Given that several studies have shown the inaccuracy of parameters estimated through observed composite scores (for measures that are not perfectly reliable; see McNeish & Wolf, 2020, for an updated perspective), the use of LCS models is pivotal for a thorough estimation of mean-level change. Furthermore, with means reported in descriptive statistics, it is not possible to estimate the magnitude and significance of the variance of change.

Bleidorn et al., 2019; Hamaker, 2012; Podsakoff et al., 2019; Wagner et al., 2020). Hence, we hypothesized that students show heterogeneity in the way they change over time in math (H_{1m}) and verbal (H_{1v}) self-concept. Moreover, consistent with previous findings (Green et al., 2012; Hancock et al., 2001; Preckel et al., 2013) and theories (Harter, 2006) discussed above, we hypothesized that across junior high school, the average direction of change is negative: That is, we hypothesized that – on average – students decline in both math (H_{2m}) and verbal (H_{2v}) self-concept.

At the *multivariate level* (i.e., focusing on the relationships among academic self-concepts and academic achievement), we hypothesized that the way students change in one academic self-concept is not related to the way students change in the other (H_3).

Then, drawing on the RI/EM studies attesting that with a large time lag academic self-concepts of different domains are not longitudinally related, we hypothesized that the way students change over time in math self-concept is not related to the initial level of verbal self-concept and that the way students change over time in verbal self-concept is not related to the initial level of math self-concept (H_4). In this way, we drawn a different hypothesis than that of the original RI/EM (see Figure 1A), given that we replaced negative associations (Figure 1 A, Panel B) with non-significant ones.

In the same vein, we hypothesized that academic achievement may contrast the negative trend of self-concept in the matching domain. Thus, we hypothesized that the way students change in one academic self-concept is positively affected by previous achievement levels in the matching domain (H_{5a}) but not significantly affected by previous achievements in the non-matching domain (H_{5b})[2].

At the same time, we hypothesized that those able to maintain a certain degree of stability in self-concepts and those starting high at T1 would benefit in terms of academic achievement in the

[2] A note on hypotheses H_4, H_{5a}, and H_{5b}: A cross-lagged effect between a construct at T1 and a latent change T2 *minus* T1 in another construct is equivalent to a cross-lagged effect between a construct at T1 and another construct at T2, after taking into account its autoregressive path (Kessler & Greenberg, 1981). Thus, whether that cross-lagged effect is investigated through a CLPM or a LSC model, the magnitude is the same. We suggest reading Usami et al. (2015) for other similarities between LCS and CLPM.

matching domain. Thus, we hypothesized that both the way students change in academic self-concept and academic self-concept levels at the beginning of junior high school, positively affects subsequent achievement in the matching domain (H_{6a}) but do not significantly affect subsequent achievement in the non-matching domain (H_{6b}).

Overall, our model drawn from the RI/EM (see Figure 1A) but with one difference: Instead of hypothesizing that academic self-concept and performance are longitudinally negatively related (see Figure 1A, Panel B), given the large time-lag, in line with Möller et al.'s (2014) results, we hypothesized that they are non-significantly associated (see hypotheses H_{5b} and H_{6b}).

In order to strengthen the validity of our results, in the model we also included paths from achievement at T1 to achievement at T2, so that we could probe the effect exerted by self-concept factors (i.e., Δ Self-concept and self-concept at T1) on achievement at T2 after taking into account the autoregressive and cross-lagged effect among achievement. Furthermore, we also controlled for the effect of gender as a control variable. Recent studies on academic self-concept do not seem particularly concerned with the issue of the effect of gender differences on academic self-concept models (e.g., Marsh et al., 2019, 2020), probably because past studies have revealed negligible differences (e.g, Möller et al., 2014). However, we decided to include gender in our model in order to (1) offer further findings about the role of gender on academic self-concepts and achievement (which may be useful for future research, in particular for research synthesis studies) and (2) avoid *omitted variable bias*, that may threaten the consistency of parameters (see Antonakis et al., 2010).

Finally, while in this section we presented our hypotheses in a substantive way, Table 4 outlines the above hypotheses through a methodological lens, in order to demonstrate which parameters of interest are involved for supporting or not supporting each hypothesis.

Method

Preliminary Information

The dataset used for investigating our hypotheses consisted of students that participated to a project (titled *Scuolinsieme*) that took place in two northern urban Italy regions (Liguria and

Piemonte). The aim of the project, among others, was to assess and investigate the relationship between non-cognitive skills and cognitive skills. Different schools were invited to join the project by means of the direct involvement of the *Fondazione San Paolo-Fondazione per la Scuola*, which is the financing entity of the project. In the beginning of 2013, The *Fondazione San Paolo-Fondazione per la Scuola* asked to local school system public authorities (Liguria and Piemonte) to share the project participation call among the local schools. 83 schools (with high levels of complexity in terms of special education needs and migration status of their students) have been involved in a participatory process in which to present and share the project main features. The available schools (50 out of 83) were then randomly assigned to different conditions, such as “experimental” (i.e., 26 were invited to participate to some seminars and activities about non-cognitive skills, but two schools withdrew) vs. “control” (i.e., 24 schools that did not participate to those seminars and activities): In this study we only used the 24 schools that participated to seminars and activities. Thus, the final group of schools is just a small part of the total number of schools of the same kind present in the two regions (e.g., Piemonte region alone has more than 300 school of the same kind considered in the present study) and this has not to be considered as statically representative of the schools and students’ population from the two regions. Schools are from different geographical areas in the two regions: Both urban and rural areas are represented in the sample used for this study. About the students, all the students in the 1st year of junior high school at the beginning of the project have been involved, based on the availability of the teachers. In some cases, only a small proportion of students has been involved, due to the role of teachers in managing the participation inside the study (teachers were directly involved in the administration of the students’ questionnaires). Basically, the majority of the students formally assigned to all the classes involved, participated into the study (see Participants section). The project lasted from June 2013 (research design) to July 2018 (report deployment) and the whole survey was administered by INVALSI officials by means of a paper-and-pencil method.

Participants

The participants were 1674 junior high school students (50.4% females) from 24 schools in two northern Italian regions (the initial pool of invited students consisted of 2341 units, hence data were available for 71.51% of them). The number of students from each school ranged from 28 to 125 (M = 69.75, SD = 26.55; in regards to the initial pool, retention for each school varied from 58.86% to 93.33%). In the subsequent analyses standard errors were corrected for clustering with an appropriate procedure (McNeish et al., 2017; see Model Evaluation section for more information) and no two-level analysis was performed. Students' math and verbal self-concept and achievements were assessed at T1 between the end of fifth grade and the beginning of the first year of junior high school[3] (mean age: about 10; school year: 2014-2015), whereas T2 was assessed at the end of the third year of junior high school (mean age: about 13; school year: 2016-2017). The retention rate was high (93.74%, n = 1571), thus attesting to a low likelihood of attrition-related problems. We handled missing data with the Full Information Maximum Likelihood (FIML) estimation procedure (Enders, 2010).

Measures

Academic Self-concepts. Math and verbal self-concept were assessed by means of 8 items (4 for each subject) with a Likert-type scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Items were gathered from a self-concept measure widely used in Italy (Alivernini & Manganelli, 2014). We selected the four items that most resembled those used in Marsh et al. (2019, Supplemental Materials Section 1). The items were "*Compared to the other students in my class, I'm good at math/Italian*" (Item 1), "*I've always been good at math/Italian*" (Item 2), "*I quickly learn math/Italian*" (Item 3), "*It is easy for me to study math/Italian*" (Item 4). Cronbach alphas for math self-concept at T1-T2 and verbal self-concept at T1-T2 were .85, .90, .79, and .82, respectively.

[3] We point out that in the Italian context there is no 6th grade, thus the difference between the end of fifth grade and the beginning of the first year of junior high school is no more than 3 months. For this reason, we did not collect age, as all students were around 10 years-old at T1

Academic Achievement. Math and verbal academic achievement at T1 and T2 were measured in terms of standardized scores obtained from the Italian National Testing system issued by INVALSI (*Istituto Nazionale di Valutazione del Sistema Educativo di Istruzione e Formazione* [National Institute for the Evaluation of Education and Training System]). These scores were standardized accordingly to a Rasch model-based approach (Bond & Fox, 2007; INVALSI, 2017), with M = 200 and SD = 40. Thus, the measurement for each academic achievement consisted of one objective measure, which was the final result of a procedure consisting of (a) collection, (b) calculation, and (c) standardization of different indicators of achievement. This procedure was carried out by INVALSI, which then sent to researchers the final standardized composite score for each student. The same procedure and measurement have been used at T1 and T2. Given that a high discrepancy in terms of variance would impact on parameter estimation (Kline, 2016, p. 81-82), the four academic achievement variables were divided by 100.

Data Analytic Strategy

In order to test our hypotheses, we adopted a latent variable approach, with latent variables composed of observed items (see Marsh, Lüdtke, Nagengast, Morin, & von Davier, 2013). The data analysis strategy moved through three steps.

In the first step, we investigated the tenability of the longitudinal measurement invariance assumption (Little, 2013; Vandenberg & Lance, 2000), given that a lack of invariance could generate severe biases in parameter estimates when performing a LCS model (Clark et al., 2018). To this end, we started by testing *configural invariance*, simply assuming that the same observed variables loaded onto the same latent variable across time. In this model, and in the subsequent ones, we modeled a covariance path linking the latent variables at T1 and T2, four residual covariances linking each observed variable with its longitudinal counterpart, and we used the first observed indicator as the marker variable (loading fixed at 1). We then moved to *metric invariance*, where we tested the assumption that factor loadings would not significantly vary in their magnitude over time (i.e., $\lambda_{21} = \lambda_{62}$, $\lambda_{31} = \lambda_{72}$, $\lambda_{41} = \lambda_{82}$). Thereafter, we examined *scalar invariance*, in which we tested the assumption that intercepts of observed indicators would not significantly vary in their magnitude over time (i.e., $\tau_2 = \tau_6$, $\tau_3 = \tau_7$, $\tau_4 = \tau_8$). Finally, we tested *strict invariance*, in which we tested the assumption that residual variances of observed indicators would not significantly vary in their magnitude over time (i.e., $\theta_{11} = \theta_{55}$, $\theta_{22} = \theta_{66}$, $\theta_{33} = \theta_{77}$, $\theta_{44} = \theta_{88}$).

In the second step, we tested univariate and unconditional LCS models, separately for math and verbal self-concepts. Albeit for several years change between two constructs (either cross-sectionally [Edwards, 2001] or longitudinally but with only two waves of data [Gu et al., 2018]) was computed using raw scores, recent advances in latent variable analyses demonstrated the possibility of modeling latent change with only two waves of data, if multiple indicators are available (Alessandri et al., 2017; Finch & Shim, 2018; Miyazaki, 2017). In particular, after constructing measurement models according to the level of invariance obtained in the previous step, following Alessandri et al. (2017), we compared a *no-change model* with a *latent change model*[4]. In our case, a *no-change model* had an intercept factor that loaded on self-concept latent variables at T1 and T2, with loadings fixed at one and the intercepts of latent variables fixed at zero. This model assumed that no mean-level change occurs between T1 and T2 in the examined construct, so that the construct was stable over time. Instead, a *latent change model* consisted of a latent change factor that operationalized the mean-level change in the construct from T1 to T2. A graphical representation is provided in Figure 1. In order to build our *latent change model*, we first regressed the latent variable self-concept at T2 on self-concept at T1 fixing this autoregressive path at 1. Then, we fixed as zero residual variance (ψ_{11}) and the intercept of the latent variable self-concept at T2. Thereafter, we regressed the latent variable self-concept at T2 on a latent variable that represented the change factor (Δ Self-Concept, in Figure 1) fixing this path to 1. This model had the following 5 parameters of interest: (a) mean (k_1) and variance (φ_{11}) of the latent variable self-concept at T1, which represented the initial level and its variability, respectively; (b) mean (k_2) and variance (φ_{22}) of the latent variable Δ Self-Concept, which represented mean-level change and its

[4] The latter, with $T = 2$, is equivalent to a second-order latent growth curve model (see Alessandri et al., 2017)

variability, respectively; and (c) a covariance between self-concept at T1 and Δ Self-Concept (φ_{12}), which represented the extent to which the initial status at T1 covaried with the subsequent change in the construct. Note that a positive k_2 value means that the sample tended to increase in self-concept from T1 to T2; a negative k_2 value means that the sample tended to decrease in self-concept from T1 to T2.

In the third step, we tested a multivariate and conditional LCS model. In this model, we included both LCS models of self-concepts, together with math and verbal achievement measures at T1 and T2, and gender. In particular, the change factors (Δ Self-Concepts) covaried and were regressed on (a) math and verbal self-concept at T1 (b) math and verbal achievement at T1, and (c) gender. Math and verbal achievement at T2 were both regressed on (a) math and verbal achievement at T1, (b) math and verbal self-concept at T1, (c) math and verbal Δ Self-concept, and (d) gender. Regarding residual covariances, we added those among each pair of observed indicators (i.e., the residual variance of math self-concept item *i* at T*n* was allowed to covary with the residual variance of verbal self-concept item *i* at T*n*). Finally, at T1 we estimated all pairs of covariances, and at T2 we estimated the residual covariance between academic achievements.

Model Evaluation

M*plus* 8 statistical software (Muthén & Muthén, 1998-2017) was used to estimate all models using the robust maximum likelihood estimator (MLR). Moreover, since we are not dealing with multilevel hypotheses but students were in any case clustered within schools, in all models the M*plus* `type=complex` command was used to adjust standard errors appropriately (see McNeish et al., 2017): This strategy is useful when using hierarchical dataset but cluster differences are not of interest (see Marsh et al., 2019, for a similar application; see McNeish et al., 2017, for technical details). The goodness of fit of each model was evaluated using the Yuan-Bentler scaled χ^2 statistic (YBχ^2; Yuan & Bentler, 2000), the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root-Mean-Square Error of Approximation (RMSEA) and the Standardized Root Mean Square Residual (SRMR). We accepted CFI and TLI values > 0.90, RMSEA values < 0.05 and SRMR values < 0.08 as indicators of adequate fit (Kline, 2016). In invariance routine, we compared nested models using the scaled difference chi-square devised by Satorra and Bentler (2001; SB$\Delta\chi^2$) and difference in CFI (ΔCFI). The more parsimonious model was considered to have a similar fit to the comparison model if SB$\Delta\chi^2$ was non-significant and ΔCFI < .01 (Cheung & Rensvold, 2002); however, "if the sample sizes are large […] even a small difference […] may result in a significant value of $\Delta\chi^2$, indicating that the null hypothesis of no difference should be rejected even when the difference is trivial" (Cheung & Rensvold, 2002, p. 239). Hence, if SB$\Delta\chi^2$ and ΔCFI returned inconsistent results (i.e., if SB$\Delta\chi^2$ was significant but ΔCFI < .01), we preferred the second. Finally, the comparison between the *no-change* and *latent change* models was conducted through inspection of SB$\Delta\chi^2$ and difference in Akaike Information Criterion (AIC): A significant value of SB$\Delta\chi^2$ and a low AIC (Burnham & Anderson, 2004) would indicate the need to include a change factor (i.e., the *latent change model*).

Results

Zero-order Correlations

In Table 1 we report zero-order correlations and descriptive statistics for all study variables. The inter-item correlations ranged from medium to high (all p-values < .001). In particular, for math self-concept the inter-item correlation ranged from .53 to .67 at T1 and from .65 to .75 at T2; for verbal self-concept, it ranged from .42 to .59 at T1 and from .44 to .64 at T2. Instead, correlations among items of different self-concepts ranged from low to medium (Cohen, 1992). In particular, the correlations among items of math self-concept and verbal self-concept at T1 ranged from .03 (p > .05; SC4_Verb_T1 with SC2_Math_T1) to .26 (p < .001; SC3_Verb_T1 with SC3_Math_T1), and at T2 ranged from .05 (p > .05, SC2_Verb_T2 with SC4_Math_T2) to .31 (p < .001, SC1_Verb_T2 with SC1_Math_T2). Regarding the relationship between academic self-concept items and academic achievement, correlations were medium-sized within the same domain (e.g., math self-concept and math achievement) and low-sized among different domains (e.g., math self-concept and verbal achievement; see Table 1). At the longitudinal level, the patterns of correlations were similar to those observed at the cross-sectional level, but with slightly smaller effect sizes (see Table 1).

Measurement Invariance

In Table 2 we report the results of invariance analyses for both math and verbal self-concept. Prior to starting with configural invariance, we added a residual covariance between items 3 and 4 (θ_{34} at T1 and θ_{78} at T2; see Figure 1), given their content is similar (at least with respect to the student's perception), in that it regards learning/studying (see, Brown, 2015, p. 38 and p. 245). As shown in Table 2, both instruments achieved strict invariance, thus attesting good psychometric properties[5].

Univariate and Unconditional Latent Change Score Models

In Table 3, we report results of the comparison between the *no-change* and *latent change* models, together with estimates of the parameters of interest. As shown by the significant SB$\Delta\chi^2$ value and by the large difference in AIC, the *latent change model* significantly fit the data better than the *no-change model* for both math and verbal self-concepts. Regarding the parameters of interest (see also Figure 1), the average initial level was similar for math (k_1 = 2.809, p < .001) and verbal (k_1 = 2.746, p < .001) self-concepts, and there was significant variability among students (for math, φ_{11} = 0.396, p < .001; for verbal, φ_{11} = 0.284, p < .001). Change factors showed similar results between math and verbal self-concepts; indeed, our sample showed a significant mean-level decrease in math self-concept (k_2 = –0.314, p < .001) as well as a significant decrease in verbal self-concept (k_2 = –0.075, p = .005), thus supporting both H_{2m} and H_{2v}. It is worth noting that the variance of change factor was significant for both academic self-concepts (for math, φ_{22} = 0.308, p < .001; for verbal, φ_{22} = 0.219, p < .001); these results attest to a substantial variability in the extent to which the students of our sample changed in academic self-concepts from T1 to T2, thus supporting H_{1m} and H_{1v}. Finally, the covariance between initial level and factor of change was

[5] In order to further support the good psychometric properties of the instrument adopted to measure academic self-concept, we also attested *gender*longitudinal* strict invariance. See Table 1A in Appendix.

negative and significant for both academic self-concepts (for math, standardized $\varphi_{12} = -.20$, $p < .001$; for verbal, standardized $\varphi_{12} = -.44$, $p < .001$). Hence, (on average) the higher a student was at T1, the lower his/her score on the change factor. Given the negative sign of the change factor, this result implies that those starting high in academic self-concept at T1 were more likely to decrease, whereas those starting low were more likely to remain fairly stable (or slightly grow). Figure 2 presents a plot of the mean-level change representing our results.

Bivariate and Conditional Latent Change Score Model

Figure 3 shows the bivariate and conditional LCS model, together with standardized parameters. This model fit the data well, according to the above-mentioned criteria: YBχ^2 = 402.157, df = 158, $p < .001$ [SCF = 1.0721]; CFI = .982; TLI = 0.976; RMSEA = 0.030; SRMR = 0.029.

As hypothesized, change factors (a) were not correlated ($\psi = .06$, $p = .108$), thus supporting H_3; (b) were not significantly affected by self-concept at T1 in the non-matching domain (math self-concept T1 → Δ verbal self-concept: .09, $p = .070$; verbal self-concept T1 → Δ math self-concept: .03, $p = .417$), thus supporting H_4; and (c) were significantly and positively affected by previous achievements in the matching domain (math achievements T1 → Δ math self-concept: .21, $p < .001$; verbal achievements T1 → Δ verbal self-concept: .17, $p < .001$) and not affected by achievement at T1 in the non-matching domain (math achievements T1 → Δ verbal self-concept: -.02, $p = .703$; verbal achievements T1 → Δ math self-concept: .04, $p = .397$), thus supporting H_{5a} and H_{5b}, respectively.

Interestingly, the explained percentage of variance of academic self-concept change factors was significant ($R^2_{\Delta\text{Math\_SC}} = .245$ [24.5%], $R^2_{\Delta\text{Verb\_SC}} = .095$ [9.5%]), but at the same time this leaves room for potential antecedents other than previous levels of academic self-concepts, academic achievement and gender.

In regards to the prediction of academic achievement at T2, both showed a medium-sized autoregressive path (for math achievement, $\beta = .32$, $p < .001$; for verbal achievement, $\beta = .45$, $p <$

.001). As hypothesized, math achievement at T2 was positively and significantly affected by the self-concept factor in the matching domain (math self-concept T1 → math achievement T2: .28, p < .001; Δ math self-concept → math achievement T2: .21, p < .001) and not significantly affected by the self-concept factor in the non-matching domain (verbal self-concept T1 → math achievement T2: .01, p = .722; Δ verbal self-concept → math achievement T2: .07, p = .084). As hypothesized, verbal achievement at T2 was positively and significantly affected by the self-concept factor in the matching domain (verbal self-concept T1 → verbal achievement T2: .25, p < .001; Δ verbal self-concept → verbal achievement T2: .18, p < .001). Contrary to our expectations, verbal achievement at T2 was also significantly affected by the self-concept factor in the non-matching domain, but the effect sizes were small (math self-concept T1 → verbal achievement T2: .13, p < .001; Δ math self-concept → verbal achievement T2: .11, p < .001). Thus, we can partially support H_6.

The model explained approximately half of the variance in academic achievement ($R^2_{\text{Math\_Ach\_T2}}$ = .546 [54.6%], $R^2_{\text{Verb\_ Ach\_T2}}$ = .537 [53.7%]).

Regarding gender (0 = *males*, 1 = *females*), all paths were significant but very low; indeed, only two paths were above .10, one affecting Δ verbal self-concept (β = .13), and the other affecting verbal self-concept at T1 (β = .15). Thus, it seems that females were slightly higher than males on these two variables.

A summary of the hypotheses and results regarding both univariate and multivariate models is provided in Table 4.

Ancillary Analyses

In the sequel, we reported results from further analyses, in order to check the robustness of our findings.

First of all, given that in our analyses we controlled for the effect of gender through a Multiple Indicators and Multiple Causes (MIMIC) approach, to further investigate potential differences between males and females, we re-ran our model using a multiple-group approach by specifying *gender*longitudinal* strict invariance for the measurement models of both academic self-concepts, given that it was supported by a previous analysis (see Table 1A, in Appendix). The model fit the data well, according to previously mentioned criteria: YBχ^2 = 567.130 [YBχ^2_{male} = 275.303, YBχ^2_{female} = 291.827], df = 308, p < .001 [SCF = 1.0691]; CFI = 0.980; TLI = 0.976; RMSEA = 0.032; SRMR = 0.039. Overall, estimates did not show substantial differences between male and female models (see Figure 2A, in Appendix). Furthermore, after constraining all regression coefficients to be equal across groups, the model still fit the data well (YBχ^2 = 590.874 [YBχ^2_{male} = 286.739, YBχ^2_{female} = 304.135], df = 328, p < .001 [SCF = 1.0649]; CFI = 0.980; TLI = 0.977; RMSEA = 0.031; SRMR = 0.041) and Satorra-Bentler scaled difference chi-square test attested that constraints did not worsen the unconstrained model (SB$\Delta\chi^2$ = 22.989, Δdf = 20, p = .2938).

Second, given that academic self-concept latent factors were composed of observed items, we re-ran the model using an estimator for ordinal variables, specifying the categorical nature of each academic self-concept item. We found few contributions with the categorical variables on the LCS model, hence we merged the recommendations on how to parametrize an LCS model with categorical variables (McArdle, Hamagami, Chang, & Hishinuma, 2014; McArdle & Nesselroade, 2014: Chapter 20) and the recommendations on fit indices for categorical indicators (e.g., DiStefano et al., 2018). Thus, we used weighted least squares means and variance adjusted (WLSMV) as estimator, theta parametrization for handling residual variances, and both factor loadings and thresholds imposed to be equal across time for each categorical variable. Furthermore, we evaluated model fit as the following: non-significant WLSMV-basedχ^2; CFI and TLI > 0.90, RMSEA < 0.05, and Weighted Root Mean Square Residual (WRMR) < 1 (for WRMR, see DiStefano et al., 2018). Again, the model showed a good fit to the data: WLSMV-basedχ^2 = 273.342, df = 166, p < .001; CFI = 0.994; TLI = 0.992; RMSEA = 0.020; WRMR = 0.984. Again, parameters obtained with WLSMV did not show substantial differences in comparison to parameters estimated with MLR, as can be seen by comparing Figures 3 and 3A (the latter is reported in Appendix). Indeed, the largest observed difference is only |.08| (i.e., path ‘verbal self-concept T1 → Δ verbal self-concept’, which is β = -.52 for MLR estimator and β = -.44 for WLSMV estimator).

Third, we ran a factor CLPM (Usami et al., 2019) with no-mean-structure, metric invariance, and MLR + cluster robust-standard errors estimation method, in order to probe the difference between the LCS model and CLPM parameters. The model fit the data well: YBχ^2 = 354.818, df = 144, p < .001 [SCF = 1.0678]; CFI = 0.984; TLI = 0.977; RMSEA = 0.030; SRMR = 0.027. Estimates are reported in Figure 4A (in Appendix). We point out that univariate results obtained with LCS models are not replicable with a factor CLPM approach, given that the latter does not take into account the mean-structure. As expected, all autoregressive academic self-concept paths were positive, whereas they were negative in the LCS model (due to regression to the mean; see Figure 4A and Figure 3). Another difference regards the explained variance of academic self-concepts at T2, which was higher than the ones found in Δ Self-Concept factors. Instead, as previously discussed, cross-lagged relations predicting academic self-concepts at T2 were mostly equivalent to those obtained in the LCS model (Kessler & Greenberg, 1981). However, cross-lagged relations for academic self-concepts T1 → academic achievement T2 were lower than those reported in the LCS model, given that the latter approach also took into account the effect of Δ Self-Concept factors on academic achievement at T2. None of the negative relationships hypothesized by RI/EM (Figure 1A, Panel B) were supported. Indeed, all paths were non-significant or positive in our factor CLPM (see Figure 4A). This finding is consistent with RI/EM studies that utilized a large time-lag.

Discussion

In this study, we used univariate and multivariate Latent Change Score models in order to further our knowledge on (a) academic self-concept change and (b) its relationship with academic achievement.

Our study has made a novel contribution to the literature on academic self-concept in several ways. First of all, our focus on “change” has shown that both math and verbal self-concept have a significant changeable component that should be taken into account when studying its predictors, correlates and outcomes (Bleidorn et al., 2019; Grimm et al., 2016). After all, educational psychologists are interested in the extent to which a construct is open to interventions. The study of the nomological network of change factors is therefore of pivotal importance in order better to address interventions and manipulations of a specific construct; hence, it should be as relevant as the study of the nomological network of its stable components (e.g., Anusic & Schimmack, 2016; Bleidorn et al., 2019; Hamaker, 2012; Podsakoff et al., 2019; Wagner et al., 2020).

Univariate Level Results: Self-concept Change

At the *univariate level*, we demonstrated that math and verbal self-concepts significantly decline over time (*mean-level change*; Bleidorn et al., 2019). While this finding is not new (Green et al., 2012; Hancock et al., 2001; Harter, 2006; Preckel et al., 2013), our study sheds light on the between-person differences in the within-person change process of this trend (*individual-level change*; Bleidorn et al., 2019), as attested by the significant variance of the latent change factors. This result implies that the overall trend in self-concept is negative, but at the same time, it also indicates that not all students have the same trend of change (i.e., there is variability in the amount and direction of change). Moreover, these findings are also consistent with recent perspectives on the high level of changeability of most psychological individual differences (e.g., Anusic & Schimmack, 2016; Bleidorn et al., 2019; Hamaker, 2012; Podsakoff et al., 2019; Wagner et al., 2020).

Multivariate Level Results: Self-concept, Self-concept Change, and Academic Achievements

At the *multivariate level*, we found that changes in academic self-concepts are not related, consistent with the I/E model (Marsh et al., 2019), and that change in academic self-concept in a domain is negatively affected by previous self-concept levels in the matching domain, but not by previous self-concept levels in the non-matching domain. We also found a reciprocal relationship between self-concept, change in self-concept and achievement within matching domains, hence supporting the mutual reinforcement exerted by self-concept and achievement (Marsh & Martin, 2011). Of importance, our results evidence a positive and significant longitudinal path between the academic self-concept change factor and matching achievement. This means that (within the same domain) the higher the score on the self-concept change factor, the greater the achievement, whereas the lower the score on the self-concept change factor, the lower the achievement. This finding is consistent with research on mean-level changes in personality traits attesting that changes in individual differences have a significant impact on important life outcomes, such as work adjustment (Alessandri et al., 2020) and career success (Hoff et al., 2021). Therefore, it has been argued that positive mean-level changes "can have functional value" (Alessandri et al., 2020, p. 1229). In our case, the overall negative trend in academic self-concept hypothesized by previous research and supported by our study should stimulate researchers and practitioners to devise schemes aimed at reinforcing academic self-concepts over time (but see Trautwein & Möller, 2016, paragraph 8.3.2). Finally, we found that gender does not exert a strong effect in any of the variables involved, which is consistent with previous studies (e.g, Möller et al., 2014): The only two effects above |.10| are the effects on verbal self-concept stability (.15) and change (.13) that attested higher levels for females (see Figure 3).

Summary of Results

To summarize, to our knowledge this is the first study that draws on a combination of research on individual difference change (e.g., Anusic & Schimmack, 2016; Bleidorn et al., 2019; Hamaker, 2012; Podsakoff et al., 2019; Wagner et al., 2020) and research on the RI/EM (Möller et al., 2011, 2014; Niepel et al., 2014; Sewasew & Schroeders, 2019), thus including both the stability and changeability of academic self-concepts across junior high school. As shown by our findings, some estimates of the relationships between academic self-concepts and academic achievement differ from those obtained in classical cross-lagged panel studies. While those studies have shown a negative mutual impact between academic self-concept and achievement in the non-matching domain (see Figure 1A, Panel B, in Appendix), our study has attested to non-significant or low-positive effects of academic self-concept change factors on subsequent non-matching academic achievement. Hence, increasing levels of academic self-concept will not impact subsequent non-matching academic achievement. Given that these results are not consistent with the RI/EM (e.g., Niepel et al., 2014), they should be thoroughly investigated in future studies. In more detail, while we explained that in the long run those negative hypothesized paths approach zero, we noticed that math self-concept demonstrated a positive and significant effect on verbal achievement T2 in both the factor CLPM and LCS model. While the factor CLPM showed that the *initial level* of math self-concept positively and significantly affected subsequent verbal achievement, our LCS model added pivotal information, namely that both math self-concept *stability* and *change* affected subsequent verbal achievement. Thus, this effect – albeit low in terms of Cohen's guidelines (Cohen, 1992) – is not negligible in our sample. We have two explanations for this inconsistent finding. The first regards the need for more studies that investigate the tenability of the RI/EM hypothesized links; indeed, most primary and meta-analytical studies have investigated the relationships among academic self-concepts and achievement with a reduced number of predictors or outcomes and only few studies have investigated the entire RI/EM utilizing an adequate methodological procedure (i.e., use of latent variables, controlling the effects of all predictors on all outcomes, use of proper estimators and adequate sample size) given that this comprehensive model is relatively new (Möller et al., 2011). Thus, the actual estimates of the RI/EM deserve to be further investigated. Second, a possible explanation is cross-cultural or sample specific (Simons, Shoda, & Lindsay, 2017); future studies with a hierarchical structure (e.g., in terms of countries and age) may reveal that not all paths hypothesized by the RI/EM are consistent across levels. Therefore, it is possible that some characteristics of our sample may have contributed to the above-mentioned differences. In this regard, we try to provide a culture-related explanation. In the Italian context, math is perceived as one of the most important subjects; this is attested, for example, by the attention payed by Italian researchers about the construct of "math anxiety" (e.g., Caviola et al., 2017; Passolunghi et al., 2020; Primi et al., 2014) while there is no presence of an "Italian/Verbal anxiety" scale). As such, it is possible that math self-concept variance also includes elements of general (or higher-order) academic-self-concept/positive-self-belief feelings, which are naturally positively related to academic achievement (e.g., Marsh, Martin, Yeung, & Craven, 2017). This component of variance may in turn be the reason why we found a positive significant effect exerted by math self-concept to verbal achievement. While we cannot be sure of the validity of our speculation, future studies could adopt two strategies for investigating the pure effect of math self-concept on verbal achievements: First, by including in the model measures related to self-worth or positive self-beliefs, such as global self-esteem (Trautwein et al., 2006), academic self-efficacy (Huang, 2012), or academic (contingent) self-worth (Crocker & Luhtanen, 2003). Second, by using bifactor models (see Morin et al., 2020) that disentangle specific factor variance (e.g., unique variance due to verbal and math self-concept) vs. general factor variance (i.e., what is shared by verbal and math self-concept). In this way, it is possible to take into account potential differences in the perceived value of the subjects under examination.

Finally, the significant percentage of explained variance for math (9.5%) and verbal (24.5%) self-concept change factors does not exclude the need to study other predictors of within-person change in order to enhance the development of interventions that aim at tackling academic self-concept change.

Practical Implications

In their review on the current state of knowledge about personality trait change in terms of its implications for public policy, Bleidorn and colleagues (2019) maintained that “the success of specific practices, interventions, and laws designed to improve the human condition depends, at least in part, on an informed understanding of *when*, *what*, *who*, and *how* to intervene” (Bleidorn et al., 2019, p. 1063). Our results at the *univariate level* clearly indicated that the mean-level change (across junior high school) in math and verbal self-concepts is negative, thus empirically supporting Harter’s (2006) theoretical arguments on academic self-concept change discussed in the introductory section. Furthermore, we also attested a significant degree of individual-level change variance, which means that students varied in the degree to which they change over junior high school. Finally, we also found that those starting higher (or lower) in self-concept are more prone to decline (or increase) across years. Thus, drawing on Bleidorn et al.'s (2019) insights, our results may suggest paying particular attention to academic self-concept "individual" change across junior high school ("*when*") and implement interventions not only on the basis of the rank-order level, but also on the basis of individual change. That is, if a student has an acceptable rank-order level of academic self-concept, but in the last years or months his/her academic self-concept level has decreased substantially, then he/she deserves attention ("*who*"). In a similar scenario, it is important to understand which types of behaviors have led the student to a significant decrease in academic self-concept ("*what*"). Finally, albeit the literature has largely supported the utility of interventions on academic self-concepts (e.g., O'Mara et al., 2006; but see Parker, Dicke, Guo, & Marsh, 2019), our focus on the "change" may advise interventions whose outcomes are not only tracked at the rank-order level but also at the individual change level ("*how*").

At the *multivariate level,* we found support for the hypothesis that math and verbal self-concept change are not related. From a practical point of view, this finding is consistent with the view of academic self-concept as a multifaceted construct (Marsh et al., 2019, 2020), and hence it further supports the notion that interventions should be targeted to a specific academic self-concept domain (e.g., math self-concept). Finally, our results (a) confirmed the positive effect of academic achievement in a domain on subsequent academic self-concept change within the same domain, (b) found a positive effect of academic self-concept change in a domain on subsequent academic achievement within the same domain, and (c) found no negative effect of self-concept change in a domain to academic achievement in the other domain. The latter result is particularly important because it differs from RI/EM principles (see Figure 1A, Panel B). Specifically, it suggests that practitioners should not fear that increasing academic self-concept in a domain will decrease achievement in the other domain. As mentioned in the Discussion section, this finding is also supported by the recent body of research attesting the "functional value" of positive mean-level changes in personality traits (Alessandri et al., 2020; Bleidorn et al., 2019; Hoff et al., 2021).

That said, considering the combination of both *univariate* and *multivariate* results, there are two main ways to sustain academic self-concept levels. The first strategy is mainly cognitive and regards the promotion of competences in subject matters of interest (Harter, 1999). Instead, the second strategy is non-cognitive and regards the support provided by significant "others" (e.g., teachers, classmates, parents; Hamre & Pianta, 2006). In both cases, a crucial aspect is the external support available to students. As far as we know, self-concept is closely related to the labels and judgments attributed by others (e.g., teachers' assessment of students). Interventions should aim at the development of students' ability to "metabolize" negative judgments, and on the other hand, internalize and identify with favorable judgments.

Interventions for enhancing student's self-concept may be applied at both the individual and classroom levels. At the *individual level*, working on social and socio-emotional skills is crucial to support student's self-concept in specific domains. This can be done through tools stemming from clinical psychology (e.g., cognitive behavioral therapy or mindfulness techniques; Hattie, 1992) and positive psychology (e.g., techniques that stimulate optimism, hope, resilience, and determination in dealing with negative situations; see Duckworth et al., 2011; Yeager et al., 2014, 2019). At the *classroom level*, all activities strongly characterized by cooperation, socialization, positive feedback, peer tutoring, praise, recognition, and gratitude are highly recommended, given the impact of the classroom environment on student's self-concept (Elbaum & Vaughn, 2001; Hamre & Pianta, 2006). In the same vein, teaching strategies that break down or cancel social confrontation in the classroom, which contain the effects of the "point of reference" that students develop in judging their abilities, can be just as effective (Manning et al., 2006). Finally, the classroom climate is another positive element; indeed, the perception of classroom as a caring community is positively related to student's academic, social, and global self-concept (Battistich et al., 1995; Manning, 2007).

Limitations

Despite several strengths, such as the use of objective measures of achievement, a large sample size and the use of two waves of data, our study has various limitations. First of all, although it is possible to run a Latent Change Score model with only two data points (Alessandri et al., 2017; Finch & Shim, 2018; Miyazaki, 2017), in order to better track differences in within-person changes, more waves of data would allow the specification of more complex and informative Latent Change Scores models (see Ferrer et al., 2019). Second, in this study we investigated linear relationships among variables, whereas an interesting point is the study of the non-linear association between academic self-concepts and academic achievement. Indeed, while research has attested that low levels of academic self-concept may lead to low levels of academic achievement, it is also likely that excessively high levels of academic self-concepts may lead to an unrealistic trust in oneself that in turn, may lead to low levels of academic achievement (Trautwein & Möller, 2016, paragraph 8.3.2). Future studies may probe the tenability of a non-linear (e.g., inverted U-shaped) relationship between academic self-concept and achievement, with the aim of finding the functional levels (neither too low, nor unrealistically high) of academic self-concepts. Third, our study was conducted on a sample consisting of Italian early adolescent students; thus, we do not know whether our results could be generalized to other countries that are different from a cultural point of view, as well as to other age stages (Simons et al., 2017). In particular, it would be interesting to find the age at which academic self-concepts become more stable; future studies may investigate this trend by using a non-linear modeling of academic self-concept, but in samples that span from early adolescence to young adulthood and with several time points (e.g., McNeish & Dumas, 2017). Fourth, our study did not take into account the effect of shared environmental influences, such as teachers', classes', or socioeconomic characteristics. Future studies adopting hierarchical models may identify potential 2-level variables that may moderate the relationship among our study variables. In a similar vein, students that participated to this study at T1 moved from elementary to junior high school, and thus experienced a compelling "shift" in terms of educational context. According to the stage-environment-fit theory (Eccles & Midgley, 1989;

Eccles et al., 1993), school transition may have a significant impact on adolescent development, in particular on motivational and behavioral spheres (e.g., Eccles & Roeser, 2009). Hence, future studies should also consider the perceived levels of shift in certain characteristics (e.g., task demands) and then include them in the model, in order to probe the potential impact on academic self-concept change and on academic achievements. Fifth, in our study we could not take into account achievement changes across time, given that we used a single observed variable and given that the measures used for achievement at T1 and T2 shared the same analytical framework and unit of analysis, but not the same content. This content difference allowed us to investigate rank-order change (by estimating the autoregressive component), but not mean-level change (given that a student's score at T2 was not comparable to the score of the same student at T1). However, throughout the paper we noted that our focus was on academic self-concept change. Nonetheless, since there are recent studies that center on academic achievement changes (Wolff et al., 2019, 2020, 2021), future studies may investigate the tenability of our results with those of recent achievement studies, in order to achieve a more comprehensive model of stability/changeability in both academic self-concept and academic achievement. Sixth, the measurement of academic self-concepts and achievement at T2 occurred synchronously, thus reducing the strength of the causal claim about the effect of self-concept change on achievement. However, we point out that (a) the change operationalized by the latent factor regards a process that developed across three years and (b) in several latent change studies, the outcomes were predicted by a change factor computed through a variable synchronously measured (with outcomes) at the last time-point (e.g., Castro- Schilo, Ferrer, Hernández, & Conger, 2016; Stein et al., 2019). Finally, it may be noted that we did not discuss the negative relationship between academic self-concept at T1 and subsequent change. Despite the fact that this relationship could be interpreted as a finding supporting the notion that academic self-concept may become "more realistic over time" (Trautwein & Möller, 2016, p. 191; see also Harter, 2006), we cannot exclude the possibility that the negative effect exerted by academic self-concept levels at T1 on subsequent change may be due to regression toward the mean

(Castro-Schilo & Grimm, 2018; Marsh & Hau, 2002). Indeed, change scores typically correlate negatively with T1 scores when $T = 2$ (Linn & Slinde, 1977). Albeit the use of change scores is not necessarily inappropriate (Gu et al., 2018), one must be cautious in interpreting this result: As Campbell and Kenny (1999) put it "[…] if change scores are computed, it is inadvisable to correlate them with initial status. That correlation would likely be negative and must be negative if the variances do not increase" (p. 99). Thus, in order to better investigate the relationship between initial level and subsequent academic self-concept change, we call for studies with more waves.

Conclusion

In this contribution, we investigated some unexplored characteristics of academic self-concept (ASC) over the course of junior high school. First, we demonstrated that both math and verbal self-concept, on average, tend to decline from the beginning to the end of junior high school. Second, we found heterogeneity in this change: While the overall mean is negative, students showed different trajectories in the way they changed. Third, we found that the way students change in one ASC (e.g., math) is not related to how students change in the other ASC (e.g., verbal). Fourth, we found that those students who were able to contrast this negative trend were also those that reported higher academic achievement at the end of junior high school. Taken together, these findings suggest that the implementation of interventions on ASC during the beginning (e.g., first year) of junior high school may have subsequent beneficial effects. Indeed, despite previous studies that found a negative between-domain effect of academic self-concept on academic achievement, the present study demonstrated that the "personal change" variability exerts a non-significant or even positive effect on academic achievement of the other domain. Related to this point, the significant heterogeneity found in the way students may change in one academic self-concept suggests the need to constantly investigate and assess self-concept at the individual-level (e.g., how each student has changed from one time to another) instead of solely focusing on the aggregate-level (e.g., how the whole class has changed from one time to another).

References

Alessandri, G., Perinelli, E., Robins, R. W., Vecchione, M., & Filosa, L. (2020). Personality trait change at work: Associations with organizational socialization and identification. *Journal of Personality*, *88*(6), 1217-1234. https://doi.org/10.1111/jopy.12567

Alessandri, G., Zuffianò, A., & Perinelli, E. (2017). Evaluating intervention programs with a pretest-posttest design: A Structural Equation Modeling approach. *Frontiers in Psychology: Quantitative Psychology and Measurement, 8*, Article 223. https://doi.org/10.3389/fpsyg.2017.00223

Alivernini, F., & Manganelli, S. (2014). La misurazione a livello nazionale delle variabili collegate al successo scolastico e formativo degli studenti: Situazione attuale e prospettive di sviluppo [National-level measurement of variables connected to scholastic and educational success of students: Current situation and developmental prospects]. *RicercAzione*, *6*(1), 71-84.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086-1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology*, *110*(5), 766-781. https://doi.org/10.1037/pspp0000066

Archambault, I., Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology*, *102*(4), 804–816. https://doi.org/10.1037/a0021075

Arens, A. K., Jansen, M., Preckel, F., Schmidt, I., & Brunner, M. (2021). The structure of academic self-concept: A methodological review and empirical illustration of central models. *Review of Educational Research*, *91*(1), 34–72. https://doi.org/10.3102/0034654320972186

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*(5), 621-634. https://doi.org/10.1037/edu0000163

Battistich, V., Solomon, D., Kim, D., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, *32*(3), 627–658. https://doi.org/10.3102/00028312032003627

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons. https://doi.org/10.1002/9781118619179

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Lawrence Erlbaum Associates Publishers.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*(3), 425-440. https://doi.org/10.1007/s11336-006-1447-6

Bleidorn, W., Hill, P. L., Back, M. D., Denissen, J. J. A., Hennecke, M., Hopwood, C. J., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Wagner, J., Wrzus, C., Zimmermann, J., & Roberts, B. (2019). The policy relevance of personality traits. *American Psychologist*, *74*(9), 1056–1067. https://doi.org/10.1037/amp0000503

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261-304. https://doi.org/10.1177/0049124104268644

Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts*. Guilford Press.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. Cambridge University Press.

Castro- Schilo, L., Ferrer, E., Hernández, M. M., & Conger, R. D. (2016). Developmental outcomes of school attachment among students of Mexican origin. *Journal of Research on Adolescence, 26*(4), 753-768. https://doi.org/10.1111/jora.12223

Castro-Schilo, L., & Grimm, K. J. (2018). Using residualized change versus difference scores for longitudinal research. *Journal of Social and Personal Relationships, 35*(1), 32-58. https://doi.org/10.1177/0265407517718387

Caviola, S., Primi, C., Chiesi, F., & Mammarella, I. C. (2017). Psychometric properties of the Abbreviated Math Anxiety Scale (AMAS) in Italian primary school children. *Learning and Individual Differences, 55*, 174-182. https://doi.org/10.1016/j.lindif.2017.03.006

Chen, S., Yeh, Y., Hwang, F., & Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences, 23*, 172–178. https://doi.org/10.1016/j.lindif.2012.07.021

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Clark, D. A., Nuttall, A. K., & Bowles, R. P. (2018). Misspecification in latent change score models: Consequences for parameter estimation, model evaluation, and predicting change. *Multivariate Behavioral Research, 53*(2), 172-189. https://doi.org/10.1080/00273171.2017.1409612

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Crocker, J., & Luhtanen, R. K. (2003). Level of self-esteem and contingencies of self-worth: Unique effects on academic, social, and financial problems in college students. *Personality and Social Psychology Bulletin, 29*(6), 701-712. https://doi.org/10.1177/0146167203029006003

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 453-466. https://doi.org/10.1080/10705511.2017.1390394

Duckworth, A. L., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. M. (2011). Self- regulation strategies improve self- discipline in adolescents: Benefits of mental contrasting and implementation intentions. *Educational Psychology, 31*(1), 17-26. https://doi.org/10.1080/01443410.2010.506003

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78-89. https://doi.org/10.1080/00461520902832368

Eccles, J. S., & Midgley, C. (1989). Stage/environment fit: Developmentally appropriate classrooms for early adolescents. In R. Ames & C. Ames (Eds.), *Research on motivation in education* (Vol. 3, pp. 139-181). Academic Press.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist, 48*(2), 90–101. https://doi.org/10.1037/0003-066X.48.2.90

Eccles, J. S., & Roeser, R. W. (2009). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology: Individual bases of adolescent development* (pp. 404–434). John Wiley & Sons Inc. https://doi.org/10.1002/9780470479193.adlpsy001013

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215-225. https://doi.org/10.1177/0146167295213003

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*(3), 265-287. https://doi.org/10.1177/109442810143005

Elbaum, B., & Vaughn, S. (2001). School-based interventions to enhance the self-concept of students with learning disabilities: A meta-analysis. *Elementary School Journal, 101*(3), 303–329. https://doi.org/10.1086/499670

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Ferrer, E., Boker, S. M., & Grimm, K. J. (2019). *Longitudinal multivariate psychology*. Routledge.

Finch, W. H., & Shim, S. S. (2018). A comparison of methods for estimating relationships in the change between two time points for latent variables. *Educational and Psychological Measurement, 78*(2), 232-252. https://doi.org/10.1177/0013164416680701

Gaspard, H., Wigfield, A., Jiang, Y., Nagengast, B., Trautwein, U., & Marsh, H. W. (2018). Dimensional comparisons: How academic track students' achievements are related to their expectancy and value beliefs across multiple domains. *Contemporary Educational Psychology, 52*, 1-14. https://doi.org/10.1016/j.cedpsych.2017.10.003

Green, J., Liem, G. A. D., Martin, A. J., Colmar, S., Marsh, H. W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of Adolescence, 35*(5), 1111-1122. https://doi.org/10.1016/j.adolescence.2012.02.016

Grimm, K. J., Mazza, G. L., & Mazzocco, M. M. M. (2016). Advances in methods for assessing longitudinal change. *Educational Psychologist, 51*(3-4), 342-353. https://doi.org/10.1080/00461520.2016.1208569

Grimm, K. J., & Ram, N. (2018). Latent growth and dynamic structural equation models. *Annual Review of Clinical Psychology, 14*, 55-89. https://doi.org/10.1146/annurev-clinpsy-050817-084840

Grimm, K. J., Ram, N., & Estabrook, R. (2017). *Growth Modeling: Structural equation and multilevel modeling approaches*. Guilford Press.

Gu, Z., Emons, W. H., & Sijtsma, K. (2018). Review of issues about classical change scores: A multilevel modeling perspective on some enduring beliefs. *Psychometrika, 83*(3), 674-695. https://doi.org/10.1007/s11336-018-9611-3

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology, 95*(1), 124-136. https://doi.org/10.1037/0022-0663.95.1.124

Guo, J., Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A longitudinal study. *American Educational Research Journal*, *52*(2), 371-402. https://doi.org/10.3102%2F0002831214565786

Hamaker, E. L. (2012). Why researchers should think “within-person”: A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43-61). The Guilford Press.

Hamre, B. K., & Pianta, R. C. (2006). Student-teacher relationships. In G. G. Bear & K. M. Minke (Eds.), *Children’s needs III: Development, prevention, and intervention* (pp. 59–72). National Association of School Psychologists.

Hancock, G. R., Kuo, W.-L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling: A Multidisciplinary Journal, 8*(3), 470-489. https://doi.org/10.1207/s15328007sem0803_7

Harter, S. (1999). *The construction of the self: A developmental perspective*. Guilford Press.

Harter, S. (2006). The self. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 505-570). Wiley & Sons.

Harter, S. (2012). *The construction of the self: Developmental and sociocultural foundations*. Guilford Press.

Hattie, J. (1992). *Self-concept*. Erlbaum.

Hattie, J., & Marsh, H. W. (1996). Future directions in self-concept research. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (p. 421–462). John Wiley & Sons.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Hoff, K. A., Einarsdóttir, S., Chu, C., Briley, D. A., & Rounds, J. (2021). Personality changes predict early career outcomes: Discovery and replication in 12-year longitudinal studies. *Psychological Science*, *32*(1), 64-79. https://doi.org/10.1177/0956797620957998

Hoge, D. R., Smit, E. K., & Crist, J. T. (1995). Reciprocal effects of self-concept and academic achievement in sixth and seventh grade. *Journal of Youth and Adolescence, 24*(3), 295-314. https://doi.org/10.1007/BF01537598

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

Huang, C. (2012). Discriminant and incremental validity of self-concept and academic self-efficacy: A meta-analysis. *Educational Psychology*, *32*(6), 777–805. https://doi.org/10.1080/01443410.2012.732386

INVALSI (2017). *Rilevazioni nazionali degli apprendimenti 2016- 17* [National surveys of learning]. Rome, Italy: INVALSI. Retrieved from https://www.invalsi.it/invalsi/doc_eventi/2017/Rapporto_tecnico_SNV_2017.pdf

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self- competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*(2), 509-527. https://doi.org/10.1111/1467-8624.00421

Jansen, M., Lüdtke, O., & Robitzsch, A. (2020). Disentangling different sources of stability and change in students' academic self-concepts: An integrative data analysis using the STARTS model. *Journal of Educational Psychology, 112*(8), 1614–1631. https://doi.org/10.1037/edu0000448

Kessler, R. C., & Greenberg, D. F. (1981). *Linear panel analysis: Models of quantitative change*. Academic Press.

Kline, R. B. (2016). *Principles and practice of Structural Equation Modeling* (4th ed.). Guilford Press.

Klopack, E. T., & Wickrama, K. (2020). Modeling latent change score analysis and extensions in Mplus: A practical guide for researchers. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 97-110. https://doi.org/10.1080/10705511.2018.1562929

Linn, R. L., & Slinde, J. A. (1977). The determination of the significance of change between pre- and posttesting periods. *Review of Educational Research, 47*(1), 121–150. https://doi.org/10.3102/00346543047001121

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

Manning, M. A. (2007). Self-concept and self-esteem in adolescents. *Student Services, 2*, 11-15.

Manning, M. A., Bear, G. G., & Minke, K. M. (2006). Self-concept and self-esteem. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 341–356). National Association of School Psychologists.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35*(2), 63–78. https://doi.org/10.1037/0022-3514.35.2.63

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38*, 299-337. https://doi.org/10.1146/annurev.ps.38.020187.001503

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*(1), 129-149. https://doi.org/10.3102/00028312023001129

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. (2002). Multilevel modeling of longitudinal growth and change: Substantive effects or regression toward the mean artifacts? *Multivariate Behavioral Research, 37*(2), 245-282. https://doi.org/10.1207/S15327906MBR3702_04

Marsh, H. W., & Hau, K. T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96*(1), 56-67. https://doi.org/10.1037/0022-0663.96.1.56

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J. S., & von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods, 18*(3), 257-284. https://doi.org/10.1037/a0032773

Marsh, H. W., & Martin, A. J. (2011). Academic self- concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59-77. https://doi.org/10.1348/000709910X503501

Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. C. (2017). Competence self-perceptions. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (2nd ed., pp. 85–115). Guilford Press.

Marsh, H. W., Parker, P. D., Guo, J., Pekrun, R., & Basarkod, G. (2020). Psychological comparison processes and self-concept in relation to five distinct frame- of- reference effects:

Pan- human cross-cultural generalizability over 68 countries. *European Journal of Personality*, *34*(2), 180-202. https://doi.org/10.1002/per.2232

Marsh, H. W., Pekrun, R., Lichtenfeld, S., Guo, J., Arens, A. K., & Murayama, K. (2016). Breaking the double-edged sword of effort/trying hard: Developmental equilibrium and longitudinal relations among effort, achievement, and academic self-concept. *Developmental Psychology*, *52*(8), 1273-1290. https://doi.org/10.1037/dev0000146

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology*, *111*(2), 331-353. https://doi.org/10.1037/edu0000281

Marsh, W. H, & Scalas, L. F. (2010). Self-concept in learning: Reciprocal effects model between academic self-concept and academic achievement. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., pp. 660-667). Elsevier. https://doi.org/10.1016/B978-0-08-044894-7.00619-9

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self- concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397-416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality*, *74*(2), 403-456. https://doi.org/10.1111/j.1467-6494.2005.00380.x

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: The effects of academic self-concept and achievement. *American Educational Research Journal*, *34*(4), 691-720. https://doi.org/10.3102/00028312034004691

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McArdle, J. J., Hamagami, F., Chang, J. Y., & Hishinuma, E. S. (2014). Longitudinal dynamic analyses of depression and academic achievement in the Hawaiian High Schools Health Survey using contemporary latent variable change models. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(4), 608-629. https://doi.org/10.1080/10705511.2014.919824

McArdle, J. J., & Nesselroade, J. R. (2014). *Longitudinal data analysis using structural equation models*. American Psychological Association. https://doi.org/10.1037/14440-000

McNeish, D., & Dumas, D. (2017). Nonlinear growth models as measurement models: A second-order growth curve model for measuring potential. *Multivariate Behavioral Research, 52*(1), 61-85. https://doi.org/10.1080/00273171.2016.1253451

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114-140. https://doi.org/10.1037/met0000078

McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods, 52*(6), 2287-2305. https://doi.org/10.3758/s13428-020-01398-0

Miyazaki, Y. (2017, July 18-21). *Evaluation of quality of estimates for change in two-wave data and its implications on interpretations* [Paper presentation]. International Meeting of the Psychometric Society (IMPS), Zürich, Switzerland.

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research, 79*(3), 1129-1167. https://doi.org/10.3102/0034654309337522

Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal, 48*(6), 1315-1346. https://doi.org/10.3102/0002831211419649

Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology, 84*(4), 591-611. https://doi.org/10.1111/bjep.12047

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376-419. https://doi.org/10.3102%2F0034654320919354

Morin, A. J., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, Exploratory Structural Equation Modeling (ESEM), and bifactor- ESEM. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (Vol. 1, 4th ed., pp. 1044-1073). John Wiley & Sons, Inc.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Author.

Niepel, C., Brunner, M., & Preckel, F. (2014). The longitudinal interplay of students' academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology, 106*(4), 1170–1191. https://doi.org/10.1037/a0036307

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181-206. https://doi.org/10.1207/s15326985ep4103_4

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013-1034. http://doi.org/10.1037/pspp0000358

Parker, P., Dicke, T., Guo, J., & Marsh, H. (2019, June 10). A macro context theory of academic self-concept: Ability stratification and the big-fish-little-pond effect. https://doi.org/10.31234/osf.io/bwy59

Passolunghi, M. C., De Vita, C., & Pellizzoni, S. (2020). Math anxiety and math achievement: The effects of emotional and math strategy training. *Developmental Science*, *23*(6), Article e12964. https://doi.org/10.1111/desc.12964

Pekrun, R., Murayama, K., Marsh, H. W., Goetz, T., & Frenzel, A. C. (2019). Happy fish in little ponds: Testing a reference group model of achievement and emotion. *Journal of Personality and Social Psychology*, *117*(1), 166–185. https://doi.org/10.1037/pspp0000230

Plante, I., O'Keefe, P. A., & Théorêt, M. (2013). The relation between achievement goal and expectancy-value theories in predicting achievement-related outcomes: A test of four theoretical conceptions. *Motivation and Emotion*, *37*(1), 65-78. https://doi.org/10.1007/s11031-012-9282-9

Podsakoff, N. P., Spoelma, T. M., Chawla, N., & Gabriel, A. S. (2019). What predicts within-person variance in applied psychology constructs? An empirical examination. *Journal of Applied Psychology*, *104*(6), 727-754. https://doi.org/10.1037/apl0000374

Preckel, F., Niepel, C., Schneider, M., & Brunner, M. (2013). Self-concept in adolescence: A longitudinal study on reciprocal effects of self-perceptions in academic and social domains. *Journal of Adolescence*, *36*(6), 1165-1175. https://doi.org/10.1016/j.adolescence.2013.09.001

Primi, C., Busdraghi, C., Tomasetto, C., Morsanyi, K., & Chiesi, F. (2014). Measuring math anxiety in Italian college and high school students: Validity, reliability and gender invariance of the Abbreviated Math Anxiety Scale (AMAS). *Learning and Individual Differences*, *34*, 51-56. https://doi.org/10.1016/j.lindif.2014.05.012

Quílez-Robres, A., Moyano, N., & Cortés-Pascual, A. (2021). Motivational, Emotional, and Social Factors Explain Academic Achievement in Children Aged 6–12 Years: A Meta-Analysis. *Education Sciences*, *11*(9), Article 513. https://doi.org/10.3390/educsci11090513

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. https://doi.org/10.1007/BF02296192

Scalas, L. F., & Fadda, D. (2019). The weight of expectancy-value and achievement goals on scientific career interest and math achievement. In D. G. Sampson, D. Ifenthaler, P. Isaias, & M. L. Mascia (Eds.), *16th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2019)* (pp. 281-288). http://www.iadisportal.org/components/com_booklibrary/ebooks/201911L035.pdf

Sewasew, D., & Schroeders, U. (2019). The developmental interplay of academic self-concept and achievement within and across domains among primary school students. *Contemporary Educational Psychology*, *58*, 204-212. https://doi.org/10.1016/j.cedpsych.2019.03.009

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407-441. https://doi.org/10.3102/00346543046003407

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123-1128. https://doi.org/10.1177/1745691617708630

Stein, G. L., Castro-Schilo, L., Cavanaugh, A. M., Mejia, Y., Christophe, N. K., & Robins, R. (2019). When discrimination hurts: The longitudinal impact of increases in peer discrimination on anxiety and depressive symptoms in mexican-origin youth. *Journal of Youth and Adolescence*, *48*(5), 864-875. https://doi.org/10.1007/s10964-019-01012-3

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept.

Journal of Personality and Social Psychology, 90(2), 334–349. https://doi.org/10.1037/0022-3514.90.2.334

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777. https://doi.org/10.1037/a0027470

Trautwein, U., & Möller, J. (2016). Self-concept: Determinants and consequences of academic self-concept in school contexts. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the twenty-first century: Theory, research, and applications* (pp. 187-214). Springer. https://doi.org/10.1007/978-3-319-28606-8_8

Usami, S., Hayes, T., & McArdle, J. J. (2015). On the mathematical relationship between latent change score and autoregressive cross-lagged factor approaches: Cautions for inferring causal relationship between variables. *Multivariate Behavioral Research, 50*(6), 676-687. https://doi.org/10.1080/00273171.2015.1079696

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods, 24*(5), 637-657. http://dx.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

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70. https://doi.org/10.1177/109442810031002

Wagner, J., Orth, U., Bleidorn, W., Hopwood, C. J., & Kandler, C. (2020). Towards an integrative model of sources of personality stability and change. *Current Directions in Psychological Science, 29*(5), 438–444. https://doi.org/10.1177/0963721420924751

Wan, S., Lauermann, F., Bailey, D. H., & Eccles, J. S. (2021). When do students begin to think that one has to be either a "math person" or a "language person"? A meta-analytic review. *Psychological Bulletin, 147*(9), 867–889. https://doi.org/10.1037/bul0000340

Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology, 48*(6), 1594–1611. https://doi.org/10.1037/a0027838

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Wolff, F., Nagy, G., Retelsdorf, J., Helm, F., Köller, O., & Möller, J. (2019). The 2I/E model: Integrating temporal comparisons into the internal/external frame of reference model. *Journal of Educational Psychology, 111*(7), 1131-1161. https://doi.org/10.1037/edu0000319

Wolff, F., Sticca, F., Niepel, C., Götz, T., Van Damme, J., & Möller, J. (2021). The reciprocal 2I/E model: An investigation of mutual relations between achievement and self-concept levels and changes in the math and verbal domain across three countries. *Journal of Educational Psychology, 113*(8), 1529–1549. https://doi.org/10.1037/edu0000632

Wolff, F., Wigfield, A., Möller, J., Dicke, A.-L., & Eccles, J. S. (2020). Social, dimensional, and temporal comparisons by students and parents: An investigation of the 2I/E model at the transition from elementary to junior high school. *Journal of Educational Psychology, 112*(8), 1644–1660. https://doi.org/10.1037/edu0000440

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., ... Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature, 573*(7774), 364-369. https://doi.org/10.1038/s41586-019-1466-y

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, *107*(4), 559–580. https://doi.org/10.1037/a0037637

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*, 165-200. https://doi.org/10.1111/0081-1750.00078

Table 1

Correlations and Descriptive Statistics of Study Variables

	1)	2)	3)	4)	5)	6)	7)	8)	9)	10)	11)	12)	13)	14)	15)	16)	17)	18)	19)	20)
1) SC1_Math_T1	1																			
2) SC2_Math_T1	.64***	1																		
3) SC3_Math_T1	.54***	.57***	1																	
4) SC4_Math_T1	.54***	.53***	.67***	1																
5) SC1_Verb_T1	.25***	.09***	.11***	.08**	1															
6) SC2_Verb _T1	.11***	.15***	.12***	.08**	.56***	1														
7) SC3_Verb _T1	.14***	.13***	.26***	.14***	.44***	.49***	1													
8) SC4_Verb _T1	.08**	.03$^{\text{n.s.}}$	.13***	.23***	.42***	.45***	.59***	1												
9) Math_Ach_T1	.34***	.34***	.31***	.24***	.12***	.08**	.12***	.06*	1											
10) Verb_Ach_T1	.23***	.21***	.23***	.17***	.18***	.19***	.19***	.14***	.68***	1										
11) SC1_Math_T2	.48***	.44***	.39***	.38***	.16***	.13***	.12***	.08**	.37***	.28***	1									
12) SC2_Math_T2	.51***	.53***	.43***	.40***	.11***	.08**	.09***	.04$^{\text{n.s.}}$	.40***	.28***	.71***	1								
13) SC3_Math_T2	.40***	.42***	.41***	.37***	.07**	.06*	.13***	.07**	.32***	.22***	.66***	.65***	1							
14) SC4_Math_T2	.41***	.41***	.39***	.39***	.08**	.04$^{\text{n.s.}}$	.09***	.08**	.33***	.22***	.67***	.66***	.75***	1						
15) SC1_Verb _T2	.14***	.13***	.15***	.13***	.36***	.37***	.30***	.29***	.15***	.25***	.31***	.18***	.11***	.09**	1					
16) SC2_Verb _T2	.11***	.10***	.12***	.08**	.35***	.43***	.31***	.28***	.12***	.22***	.16***	.18***	.08**	.05$^{\text{n.s.}}$	.61***	1				
17) SC3_Verb _T2	.13***	.12***	.21***	.17***	.29***	.28***	.33***	.32***	.13***	.18***	.16***	.14***	.21***	.16***	.49***	.49***	1			
18) SC4_Verb_T2	.08**	.08**	.14***	.15***	.27***	.28***	.31***	.32***	.11***	.18***	.09**	.07**	.09***	.18***	.44***	.52***	.64***	1		
19) Math_Ach_T2	.34***	.37***	.34***	.28***	.11***	.09***	.11***	.06*	.63***	.57***	.47***	.48***	.42***	.41***	.21***	.16***	.14***	.09***	1	
20) Verb_Ach_T2	.25***	.25***	.28***	.21***	.24***	.26***	.25***	.20***	.52***	.64***	.36***	.32***	.28***	.27***	.36***	.33***	.31***	.26***	.64***	1
M	2.79	2.95	2.96	2.94	2.73	2.85	2.97	2.90	2.10	2.10	2.52	2.59	2.63	2.50	2.68	2.75	2.87	2.81	1.98	2.01
SD	0.80	0.91	0.92	0.93	0.73	0.82	0.86	0.88	0.41	0.41	0.91	0.99	1.03	1.03	0.73	0.80	0.84	0.86	0.40	0.40
Sk	-0.28	-0.44	-0.50	-0.48	-0.28	-0.31	-0.48	-0.43	0.40	0.12	-0.12	-0.07	-0.19	-0.04	-0.35	-0.26	-0.44	-0.41	0.42	0.11
Ku	-0.35	-0.71	-0.65	-0.68	-0.06	-0.44	-0.48	-0.56	0.26	0.16	-0.78	-1.03	-1.11	-1.14	-0.03	-0.33	-0.34	-0.41	0.56	0.43

Note. SC*i*_T*n* = Self-Concept Item *i* at Time *n*; Verb = Verbal; Ach = Achievement; *M* = Mean; *SD* = Standard Deviation; *Sk* = Skewness; *Ku* = Kurtosis. $^{\text{n.s.}}$ $p > .05$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table 2

Longitudinal Measurement Invariance

Self-Concept	Invariance step	NFP	YBχ^2	*df*	*p*	SCF	CFI	TLI	RMSEA	SRMR	CD	SB$\Delta\chi^2$	Δdf	*p*	ΔCFI
Math	configural	31	29.294	13	.006	1.1388	.997	0.994	0.027	0.012					
	metric	28	35.290	16	.004	1.0887	.997	0.994	0.027	0.018	0.872	5.806	3	.121	0
	scalar	25	62.347	19	< .001	1.1070	.993	0.989	0.037	0.028	1.205	25.401	3	< .001	.004
	strict	**21**	**67.242**	**23**	**< .001**	**1.1465**	**.993**	**0.991**	**0.034**	**0.027**	**1.334**	**6.053**	**4**	**.195**	**0**
Verbal	configural	31	46.805	13	< .001	1.1192	.992	0.983	0.039	0.021					
	metric	28	50.268	16	< .001	1.0871	.992	0.986	0.036	0.023	0.948	2.386	3	.496	0
	scalar	25	53.513	19	< .001	1.0918	.992	0.988	0.033	0.022	1.117	3.384	3	.336	0
	strict	**21**	**68.354**	**23**	**< .001**	**1.0847**	**.989**	**0.987**	**0.034**	**0.030**	**1.051**	**14.956**	**4**	**.005**	**.003**

Note. Estimation method: MLR + Cluster robust-standard errors. NFP = Number of Free Parameters; YBχ^2 = Yuan-Bentler scaled chi-square; *df* = degrees of freedom; SCF = Scaling Correction Factor; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; CD = Difference Test Scaling Correction; SB$\Delta\chi^2$ = Satorra-Bentler scaled chi-square difference; Δdf = difference in degrees of freedom; ΔCFI = difference in CFI.
The level of measurement invariance obtained by the instrument is reported in bold.

Table 3

Comparison of No-Change vs. Latent Change Models and Estimate of Parameters of Interest

Self-Concept	Model	NFP	YBχ^2(*df*)	SCF	CFI	TLI	RMSEA	SRMR	AIC	SB$\Delta\chi^2$(Δdf)	ΔAIC	k_1	k_2	φ_{11}	φ_{22}	$\varphi_{12(STD)}$
Math	no-change	20	347.052(24)***	1.157	.947	0.938	0.090	0.076	28560.897	-	-	-	-	-	-	-
	latent change	**21**	**67.242(23)*****	**1.147**	**.993**	**0.991**	**0.034**	**0.027**	**28238.614**	**233.879(1)*****	**322.283**	2.809***	-0.314***	0.396***	0.308***	-.20***
Verbal	no-change	20	84.295(24)***	1.171	.986	0.983	0.039	0.029	26454.743	-	-	-	-	-	-	-
	latent change	**21**	**68.354(23)*****	**1.085**	**.989**	**0.987**	**0.034**	**0.030**	**26432.175**	**7.784(1)****	**22.568**	2.746***	-0.075**	0.284***	0.219***	-.44***

Note. Estimation method: MLR + Cluster robust-standard errors. NFP = Number of Free Parameters; YBχ^2 = Yuan-Bentler scaled chi-square; *df* = degrees of freedom; SCF = Scaling Correction Factor; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion; SB$\Delta\chi^2$ = Satorra-Bentler scaled chi-square difference; Δdf = difference in degrees of freedom; ΔAIC = difference in AIC; k_1 = latent mean of self-concept at T1; k_2 = latent mean of self-concept change factor (Δ Self-Concept); φ_{11} = latent variance of self-concept at T1; φ_{22} = latent variance of self-concept change factor (Δ Self-Concept); $\varphi_{12(STD)}$ = standardized covariance between self-concept at T1 and self-concept change factor (Δ Self-Concept).
Best models are reported in bold.
$^{**}p < .01$. $^{***}p < .001$.

Table 4

Summary of Hypotheses and Parameter Results

Level	*H#*	Hypothesis	Parameter	Supported
Univariate	H_{1m}	Δ Math Self-concept has a variance significantly different from zero	$\varphi_{22} = 0.308^{***}$	Yes
	H_{2m}	The latent mean of Δ Math Self-concept is negative and significantly different from zero.	$k_2 = -0.314^{***}$	Yes
	H_{1v}	Δ Verbal Self-concept has a variance significantly different from zero	$\varphi_{22} = 0.219^{***}$	Yes
	H_{2v}	The latent mean of Δ Verbal Self-concept is negative and significantly different from zero.	$k_2 = -0.075^{**}$	Yes
Multivariate	H_3	The covariance between Δ Math Self-concept and Δ Verbal Self-concept is not significantly different from zero	$\psi = .06^{\text{n.s.}}$	Yes
	H_4	Δ Math Self-concept is not significantly affected by Verbal Self-concept T1	$\beta = .03^{\text{n.s.}}$	Yes
	"	Δ Verbal Self-concept is not significantly affected by Math Self-concept T1	$\beta = .09^{\text{n.s.}}$	Yes
	H_{5a}	Δ Math Self-concept is significantly and positively affected by Math Achievement T1	$\beta = .21^{***}$	Yes
	"	Δ Verbal Self-concept is significantly and positively affected by Verbal Achievement T1	$\beta = .17^{***}$	Yes
	H_{5b}	Δ Math Self-concept is not significantly affected by Verbal Achievement T1	$\beta = .04^{\text{n.s.}}$	Yes
	"	Δ Verbal Self-concept is not significantly affected by Math Achievement T1	$\beta = -.02^{\text{n.s}}$	Yes
	H_{6a}	Δ Math Self-concept and Math Self-concept T1 significantly and positively affect Math Achievement T2	$\beta = .21^{***}, .28^{***}$	Yes
	"	Δ Verbal Self-concept and Verbal Self-concept T1 significantly and positively affect Verbal Achievement T2	$\beta = .18^{***}, .25^{***}$	Yes
	H_{6b}	Δ Math Self-concept and Math Self-concept T1 do not significantly affect Verbal Achievement T2	$\beta = .11^{***}, .13^{***}$	NO
	"	Δ Verbal Self-concept and Verbal Self-concept T1 do not significantly affect Math Achievement T2	$\beta = .07^{\text{n.s.}}, .01^{\text{n.s.}}$	Yes

Note. Parameters for univariate level are reported in non-standardized form; parameters for multivariate level are reported in standardized form. $^{\text{n.s.}}$ $p > .05$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

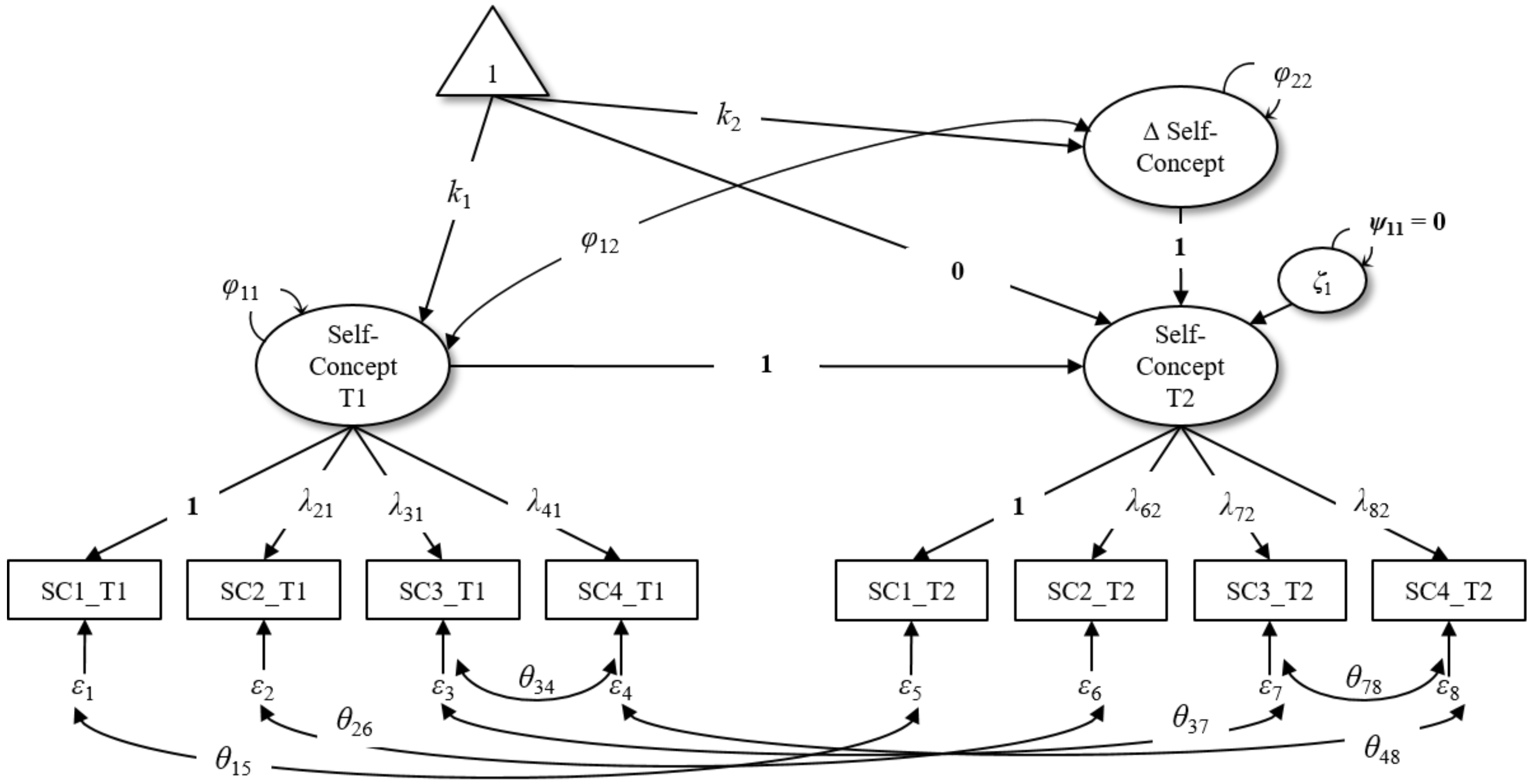

Figure 1. Univariate and unconditional Latent Change Score Model. Strict measurement invariance is imposed. Parameters in bold have been fixed. The “1” inside a triangle represents a constant (used for estimating latent means and intercepts). Δ Self-Concept = self-concept change factor; SC*i*_T*n* = self-concept Item *i* at Time *n*; *k* = latent mean; φ = latent variance or latent covariance; ψ = residual variance of latent variable (in this case, it is fixed to zero); λ = factor loading; ε = residual score of observed variable; θ = residual covariance between observed variables.

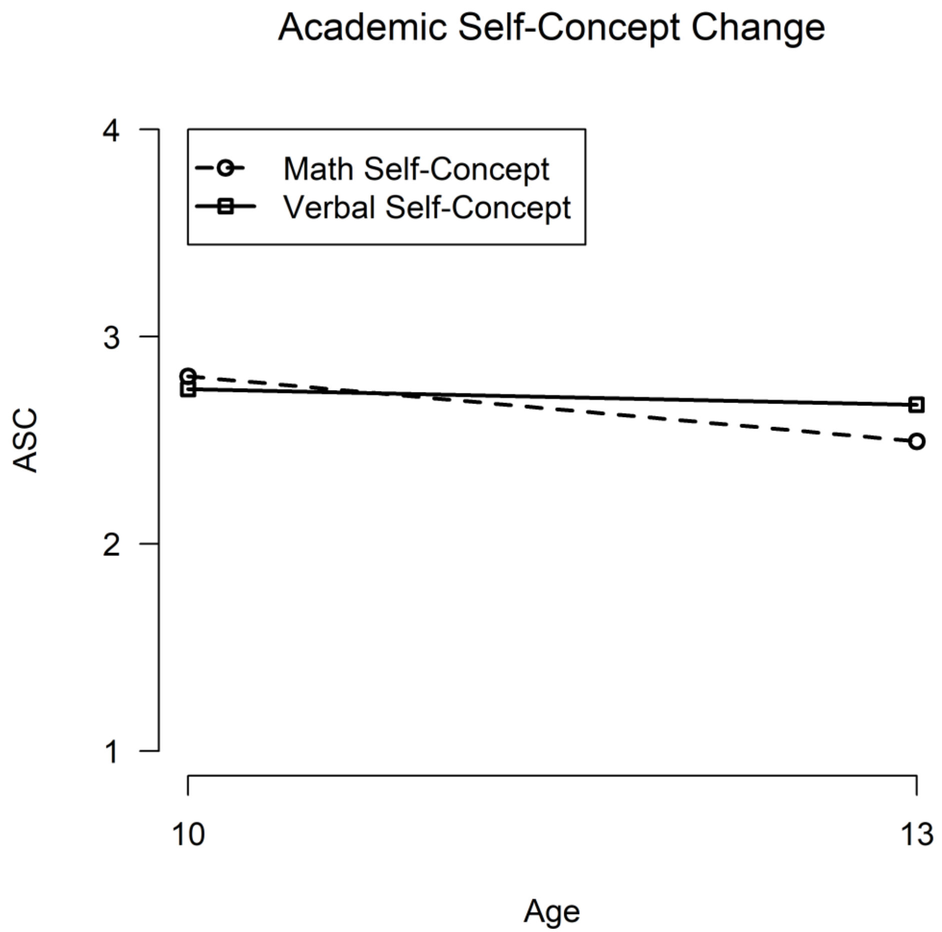

Figure 2. Academic Self-Concept (ASC) change for Math and Verbal Self-Concept from 10 to 13.

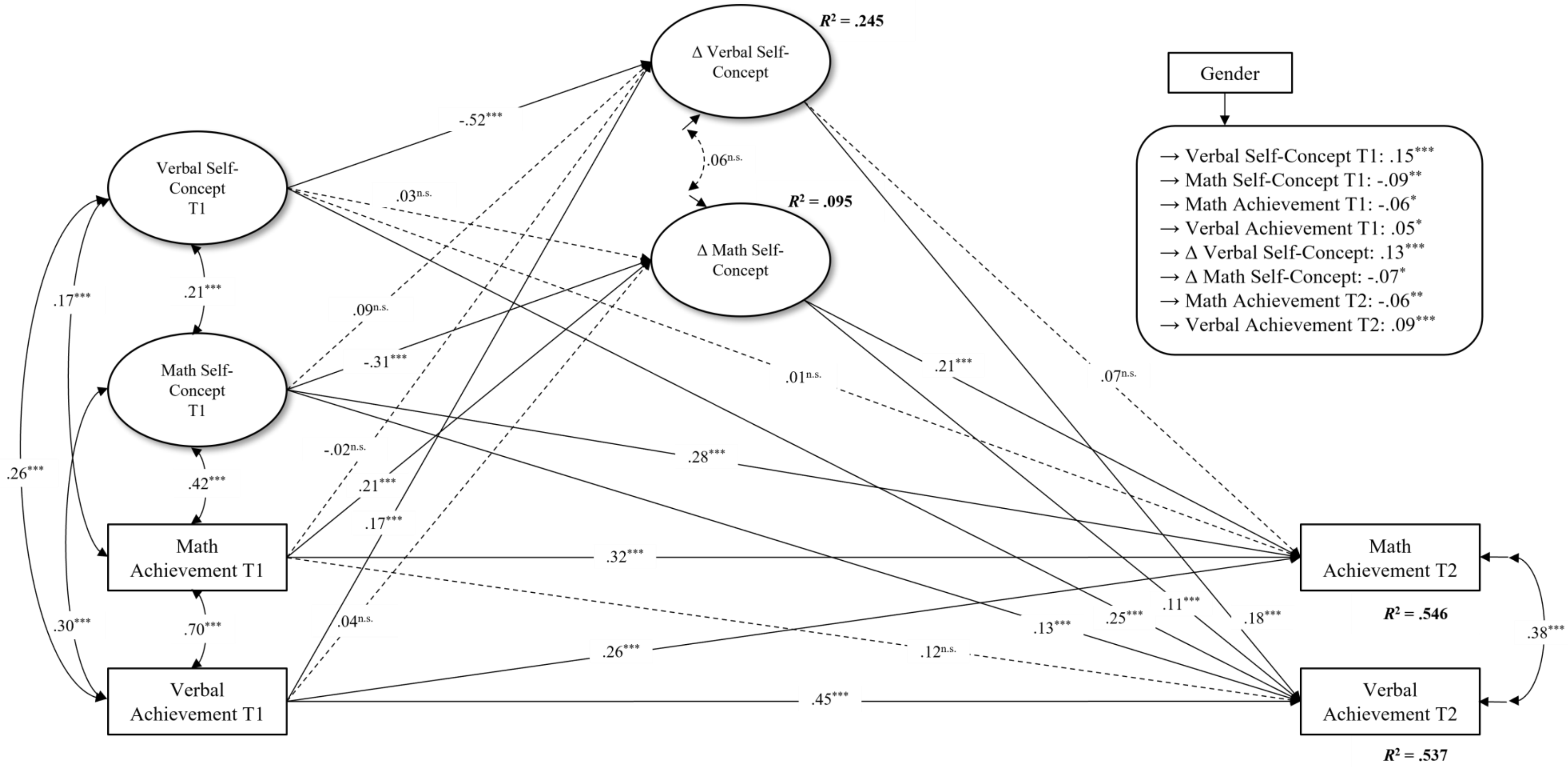

Figure 3. Bivariate and Conditional Latent Change Score Model. Estimation method: MLR + Cluster robust-standard errors. Parameters are reported in standardized form. Measurement models are omitted for the sake of clarity (see Figure 1). Paths from Gender (0 = *male*, 1 = *female*) to study variables are reported on the top right of the Figure for the sake of clarity.

n.s. $p > .05$. *$p < .05$. **$p < .01$. ***$p < .001$.

APPENDIX

Table 1A

*Gender*Longitudinal Measurement Invariance*

Self-Concept	Gen.Inv.	Long.Inv.	NFP	YBχ^2	*df*	*p*	SCF	CFI	TLI	RMSEA	SRMR	CD	SB$\Delta\chi^2$	Δdf	*p*	ΔCFI
Math	configural	configural	62	45.150	26	.011	1.1880	0.997	0.993	0.030	0.016					
	metric	configural	56	55.236	32	.007	1.1009	0.996	0.993	0.030	0.025	0.723	9.912	6	.128	.001
	scalar	configural	50	74.157	38	< .001	1.1071	0.994	0.991	0.034	0.029	1.140	18.873	6	.005	.002
	strict	configural	42	87.701	46	< .001	1.0729	0.993	0.992	0.033	0.030	0.910	13.175	8	.106	.001
	strict	metric	39	93.600	49	< .001	1.0630	0.993	0.992	0.033	0.033	0.911	5.929	3	.115	0
	strict	scalar	36	120.583	52	< .001	1.0654	0.989	0.988	0.040	0.038	1.105	26.229	3	< .001	.004
	strict	**strict**	**32**	**125.644**	**56**	**< .001**	**1.0836**	**0.988**	**0.988**	**0.039**	**0.038**	**1.320**	**5.816**	**4**	**.213**	**.001**
Verbal	configural	configural	62	56.422	26	< .001	1.0674	0.992	0.983	0.037	0.024					
	metric	configural	56	58.367	32	.003	1.0665	0.993	0.988	0.031	0.026	1.063	1.904	6	.928	-.001
	scalar	configural	50	71.194	38	< .001	1.0718	0.992	0.988	0.032	0.030	1.100	12.779	6	.047	.001
	strict	configural	42	105.586	46	< .001	1.0866	0.985	0.982	0.039	0.047	1.157	33.213	8	< .001	.007
	strict	metric	39	108.375	49	< .001	1.0791	0.985	0.983	0.038	0.048	0.964	2.300	3	.512	0
	strict	scalar	36	111.746	52	< .001	1.0823	0.985	0.984	0.037	0.047	1.135	3.521	3	.318	0
	strict	**strict**	**32**	**127.500**	**56**	**< .001**	**1.0773**	**0.982**	**0.982**	**0.039**	**0.053**	**1.012**	**16.214**	**4**	**.003**	**.003**

Note. Estimation method: MLR + Cluster robust-standard errors. Gen.Inv. = Gender Invariance Step; Long.Inv. = Longitudinal Invariance Step; NFP = Number of Free Parameters; YBχ^2 = Yuan-Bentler scaled chi-square; *df* = degrees of freedom; SCF = Scaling Correction Factor; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; CD = Difference Test Scaling Correction; SB$\Delta\chi^2$ = Satorra-Bentler scaled chi-square difference; Δdf = difference in degrees of freedom; ΔCFI = difference in CFI.

The level of measurement invariance obtained by the instrument is reported in bold.

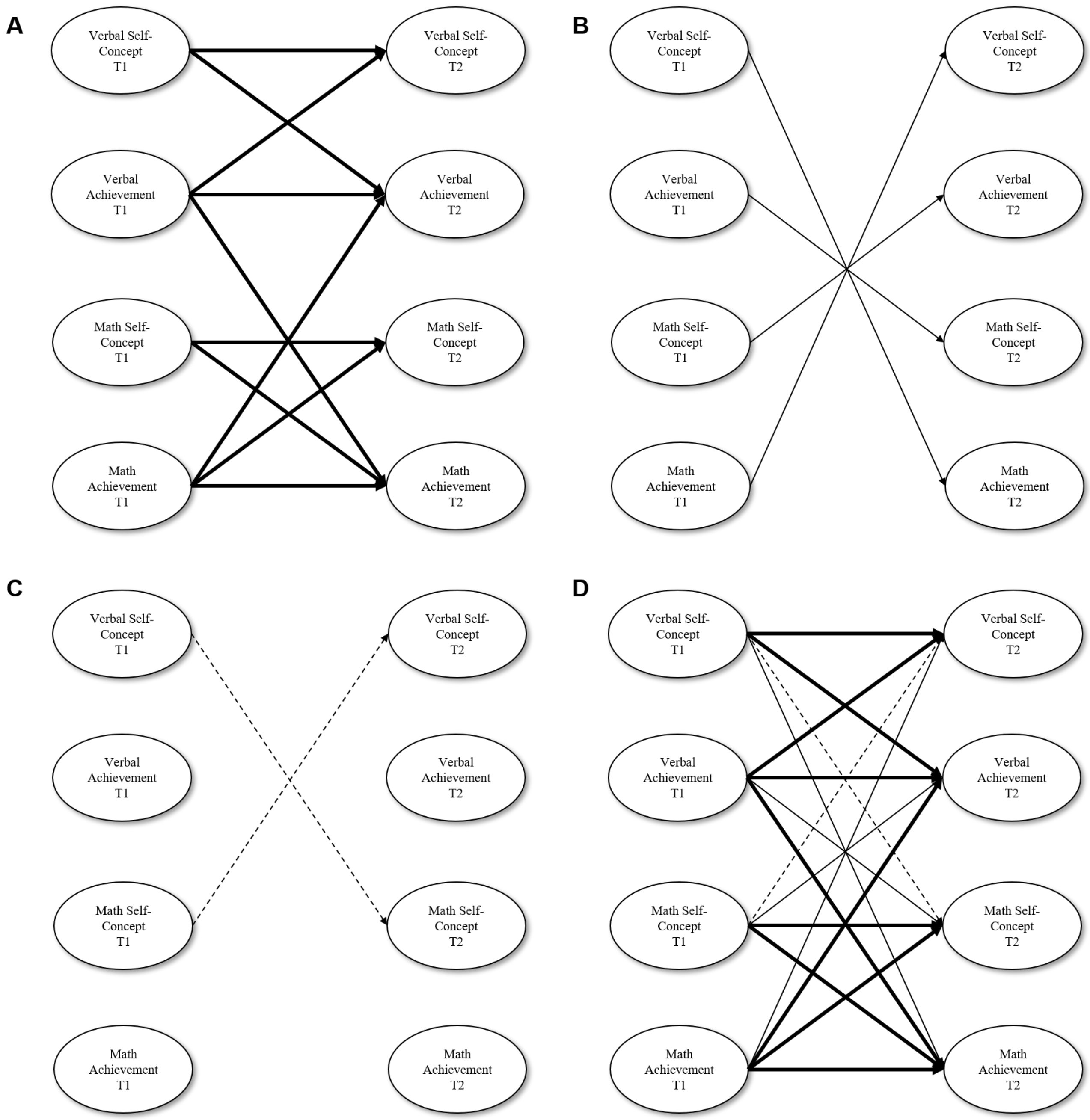

Figure 1A. A graphical representation of the RI/EM. Panel A = positive relationships (bolded lines); Panel B = negative relationships (thin lines); Panel C = non-significant relationships (dotted lines); Panel D = full RI/EM.

Note that our Hypotheses H_{5b} and H_{6b} (see Table 4) differ from relationships reported in Panel B, because given the large time lag we hypothesized those relationships to be non-significant instead of negative.

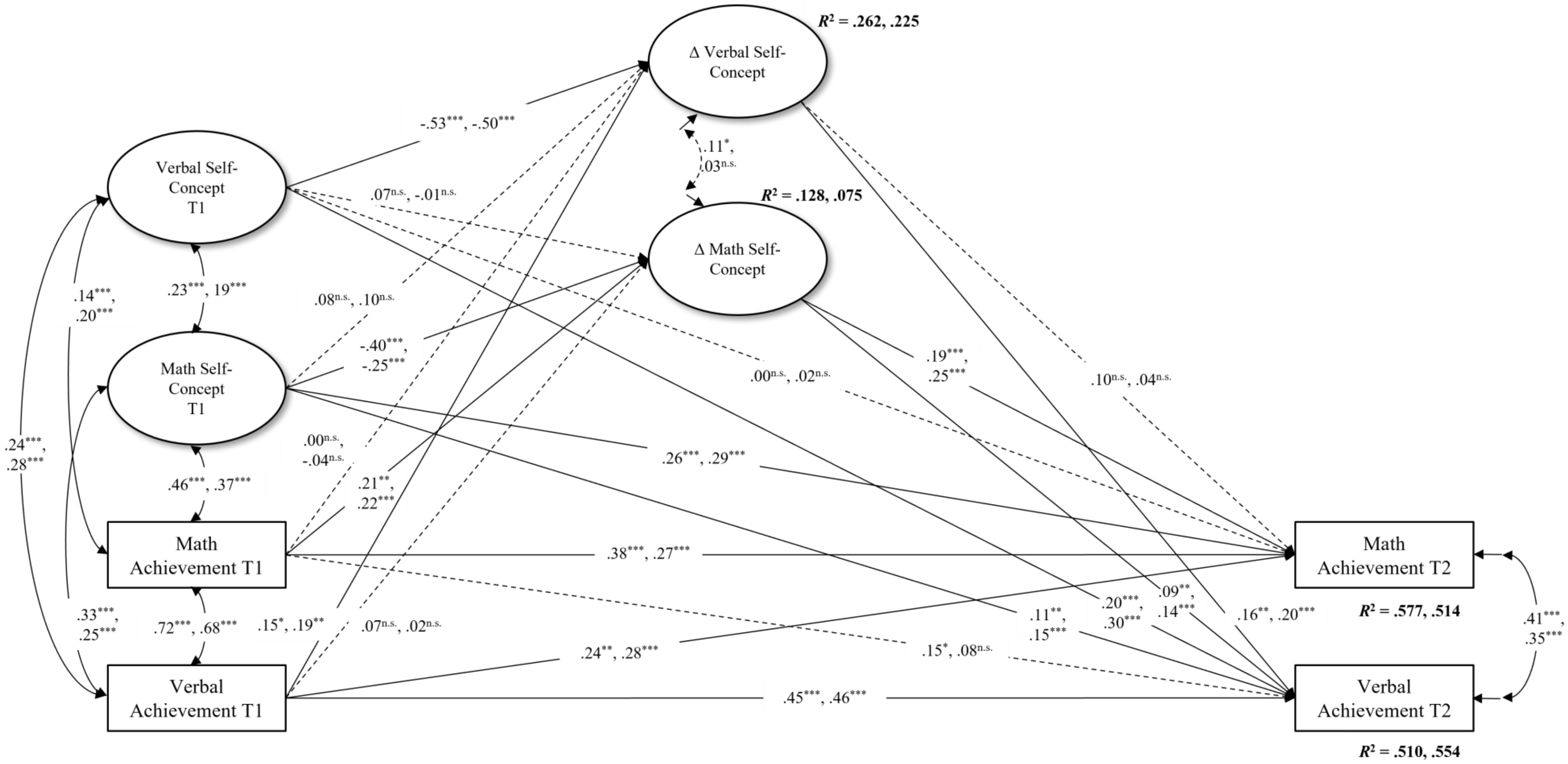

Figure 2A. Multiple Group Bivariate and Conditional Latent Change Score Model.
Note. Gender*Longitudinal strict factorial invariance imposed for math and verbal self-concept (See Table 1A). Estimation: MLR + Cluster robust-standard errors. Measurement models are omitted for the sake of clarity. The first estimate refers to “Male model”, the second estimate refers to “Female model”.
Fit indices: $YB\chi^2$ = 567.130 [$YB\chi^2_{male}$ = 275.303, $YB\chi^2_{female}$ = 291.827], df = 308, $p < .001$ [SCF = 1.0691]; CFI = 0.980; TLI = 0.976; RMSEA = 0.032; SRMR = 0.039.

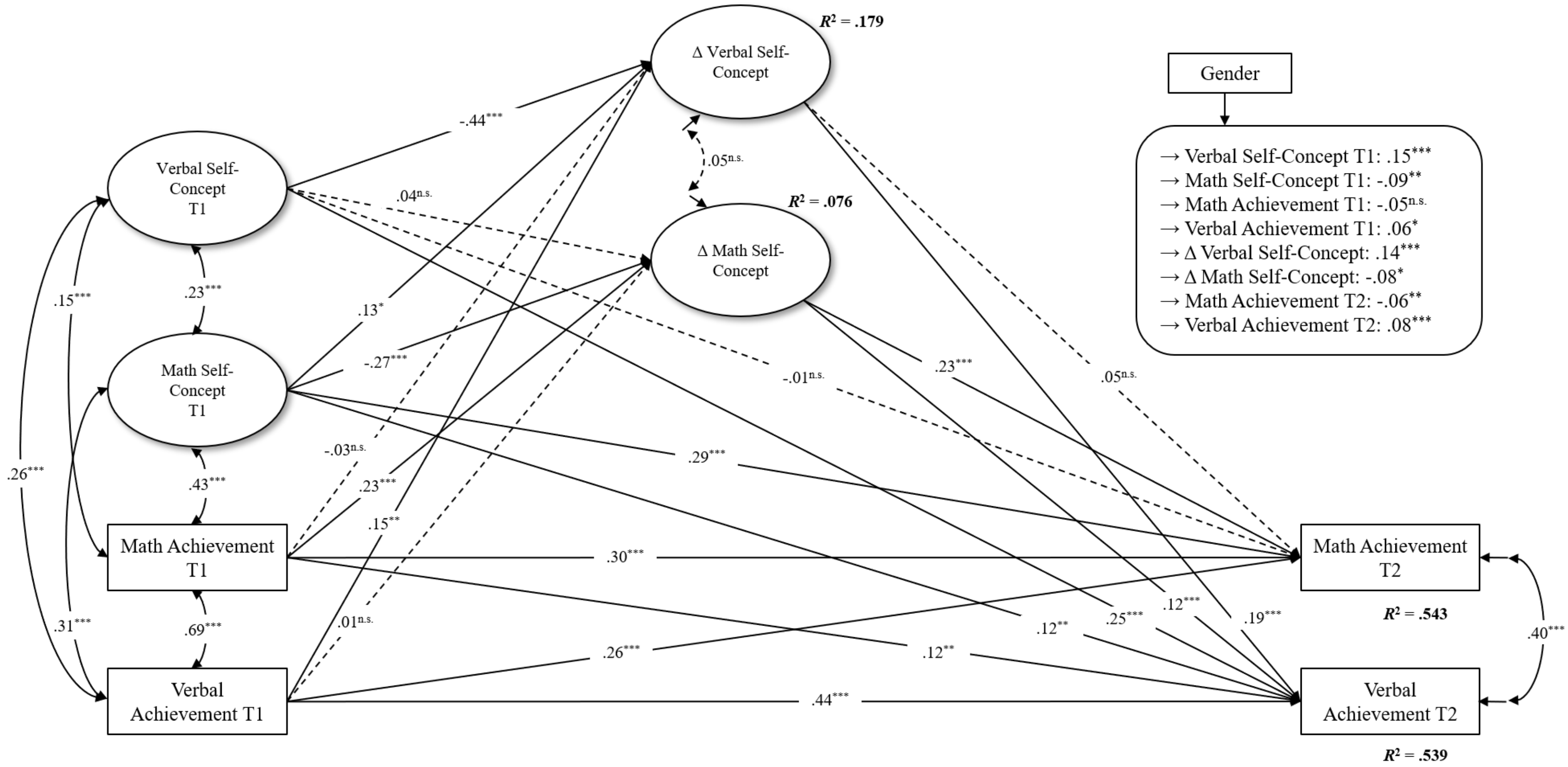

Figure 3A. Bivariate and Conditional Latent Change Score Model with WLSMV estimator and theta parameterization.
Note. Longitudinal invariance among factor loadings and thresholds was imposed. Estimation: WLSMV (theta parameterization) + Cluster robust-standard errors. Measurement models are omitted for the sake of clarity.
Fit indices: WLSMV-based χ^2 = 273.342, *df* = 166, *p* < .001; CFI = 0.994; TLI = 0.992; RMSEA = 0.020; WRMR = 0.984.

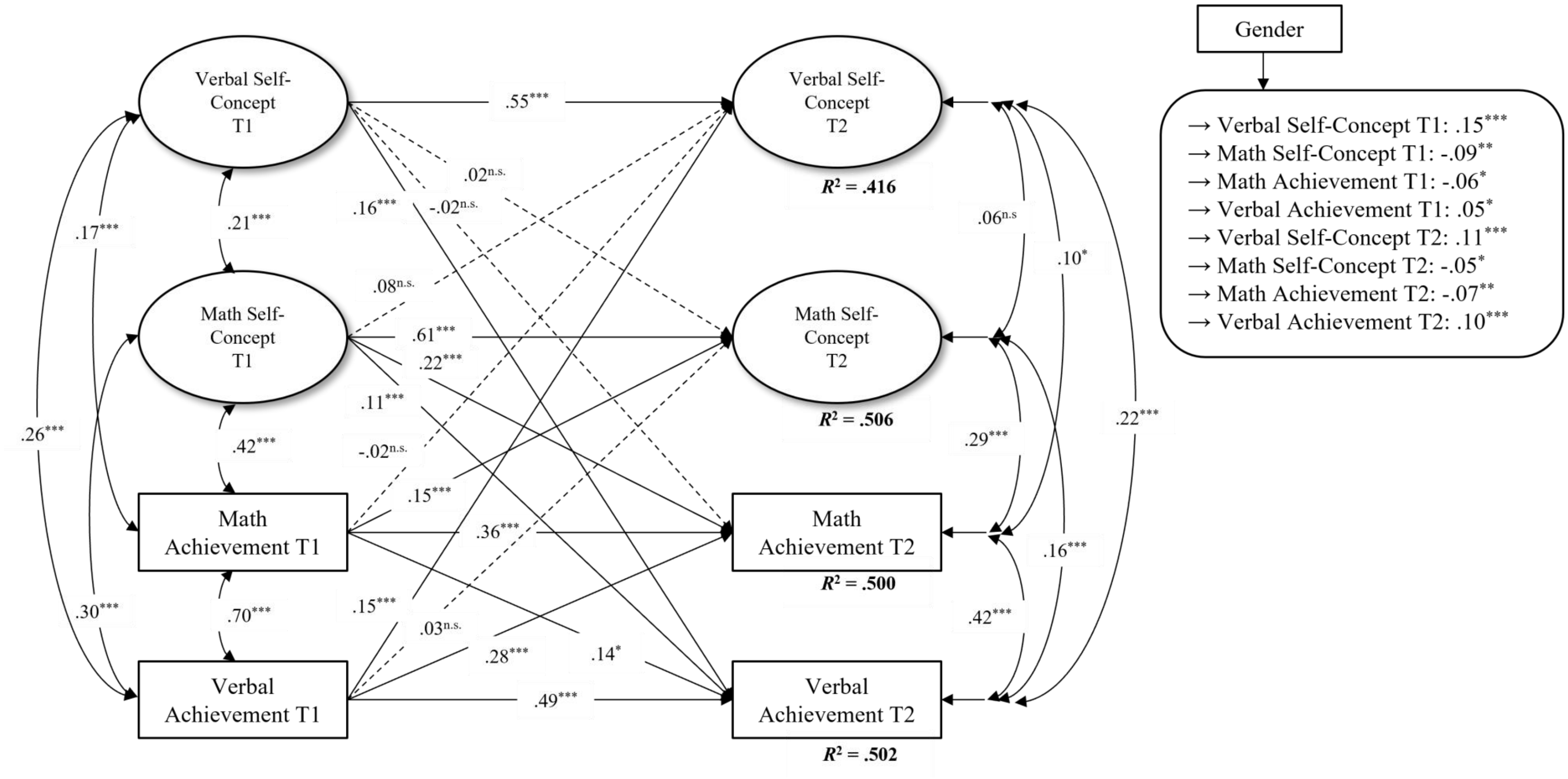

Figure 4A. Factor Cross-Lagged Panel Model.
Note. Metric factorial invariance imposed (after passing invariance routine). Estimation: MLR + Cluster robust-standard errors. Measurement models are omitted for the sake of clarity.
Fit indices: YBχ^2 = 354.818, *df* = 144, *p* < .001 [SCF = 1.0678]; CFI = 0.984; TLI = 0.977; RMSEA = 0.030; SRMR = 0.027.