

This MS is the final prepublication (open access) version of the published article:

Parker, Philip D., Marsh, Herbert W., Morin, Alexandre J. S., Seaton, Marjorie, & Van Zanden, Brooke. (2015). If One Goes up the Other Must Come Down: Examining Ipsative Relationships between Math and English Self-concept Trajectories across High School. *British Journal of Educational Psychology*, 85(2), 172-191.

This article may not exactly replicate the final published version in the journal. It is not the copy of record and readers are encouraged to obtain the copy of record through their university or local library using the article's DOI (digital object identifier).

If one goes up the other must come down: Examining ipsative relationships between math and English self-concept trajectories across high school

Background

Academic self-concept is not only an important outcome in itself but also an important predictor of other educational outcomes (Guay, Marsh, & Boivin, 2003; Marsh, 1991, 2007). It is now well established that self-concept and achievement are reciprocally related over time (Marsh & Craven, 2006), and a growing body of research indicates that academic self-concept is an important predictor of educational choices made in school and beyond (Marsh, 1991; Nagy, Trautwein, Baumert, Köller, & Garrett, 2006; Nagy *et al.*, 2008; Parker, Marsh, Ciarrochi, Marshall, & Abduljabbar, 2014; Parker, Nagy, Trautwein, & Lüdtke, in press; Parker *et al.*, 2012). Relatively little research, has considered growth in domain specific academic self-concept across key developmental periods such as high-school/adolescence and even less research has considered this growth from a multi-dimensional perspective in which the influence of self-concept in one domain (e.g., English), influences the growth of self-concept in another domain (e.g., math). This is surprising given that one of the key models of self-concept development – the internal/external (IE) frame of reference – considers dimensional comparison between academic domains as a key component in the way individuals develop their sense of self in the academic arena.

Advances in statistical models have allowed for the testing of a range of hypotheses that were not easily addressed previously. In this paper we illustrate three approaches that provide different information about the central hypothesis that English self-concept has a negative effect on the direction of growth in mathematics self-concept, and vice-versa. The three models are: (1) autoregressive cross-lagged models (here after ACLM) to test the

temporal ordering of the associations between English self-concept and mathematics self-concept; (2) latent growth curve (hereafter LGC) models to test hypotheses about the direction of growth in these two domains and the relationship between their growth trajectories; and (3) autoregressive latent trajectory (hereafter ALT) models to explore whether the relationship between English and mathematics self-concept differs at state residual (fluctuations from trajectory) and trait levels. All models include achievement as a time-varying covariate. In addition, we also illustrate how to specify these models in a cohort sequence design. This approach has a number of benefits for longitudinal models but parametrizations of models using this framework is often difficult. This is particularly the case when moving beyond LGC models as we do here. Thus, we provide applied examples and syntax to assist researchers in determining whether such models may be useful for them.

Substantive issues: Academic self-concept formation and growth

Three types of comparison that give rise to the development self-concept

As self-concept develops over the course of schooling it also becomes increasingly differentiated. It is suggested here that frame of reference processes help explain this growth pattern. Indeed, since James (1890/1960), psychologists have stressed that self-concept cannot be understood without recognising the role that comparative processes play in perceptions of the self. Research and theory have focused predominately on the role of social comparison, though there is increasing interest in applied research on dimensional and temporal comparison. Social comparison theory, developed by Festinger (1954), was pivotal in highlighting the importance of external references in self-evaluation. According to Festinger, people are driven to gather information in order to form accurate self-evaluations. Festinger argues that when people do not have objective means to evaluate their own abilities, people will compare themselves against others to form these self-evaluations. Social comparison forms the basis of the so-called big-fish-little-pond effect (or frog-pond effect; Davis, 1966) in which a student's self-concept is paradoxically negatively related to the average ability level of their peers at the school or class level after control for individual achievement differences (for a review, see Marsh, 2007). This external comparison process takes on increasing significance during development as young people become more aware of the abilities of others and adjust their self-concepts accordingly (Marsh, 1989).

Temporal comparison processes are somewhat less well researched, though they also have a long history. Albert's (1977) work on temporal comparisons was significant as it highlighted how internally based comparisons are also critical to the way people self-evaluate. However, when considered alone, social and temporal comparisons fail to explain why academic self-concept is so content-specific (Marsh, 2007).

To explain this content-specificity, Möller and Marsh (2013) focused on dimensional comparisons, in addition to social comparisons, as the core mechanisms behind the formation of differentiated subject-specific academic self-concepts in children and adolescents, as specified in the IE model. According to Möller and Marsh, dimensional comparisons occur when people compare their strengths in different academic domains, which typically results in contrast effects between distant domains (e.g., English and math) and assimilation effects between closely related domains (e.g., math and physics). For example, students may assess their mathematical proficiency by considering how good they are in math relative to other subject areas. This leads to an ipsative process whereby more positive evaluations in one academic domain may lead to lower self-evaluations in other distant domains. Marsh (1986) argues that it is these dimensional comparisons processes that lead to the often-reported lack of cross-sectional correlations between math and English self-concepts. This has received extensive support in cross-cultural support (Marsh & Hau, 2004), meta-analysis (Möller, Pohlmann, Köller, & Marsh, 2009), experimental studies (Möller & Köller, 2001), diary studies (Möller *et al.*, 2009), and IE processes have important influences on long-term educational outcomes (Nagy *et al.*, 2006, 2008; Parker *et al.*, 2012, 2014, in press).

The importance of internal comparison processes and the way they develop across schooling is also central to the Eccles (1994) model of achievement-related choices. Although Eccles and colleagues focus on expectations of success as opposed to self-concept, the two constructs are closely aligned, and empirically indistinguishable (Eccles, 2009; Eccles & Wigfield, 2002). Eccles (2009, 2011) suggests that achievement related choices are influenced by internal rankings of domain specific expectancies and task-values. Furthermore, Eccles and colleagues notes that such hierarchies are evident even in young childhood with domain specificity increasing with age (Eccles & Wigfield, 2002; Eccles, Wigfield, Harold, & Blumenfeld, 1993; Wigfield & Eccles, 2002; Wigfield, Eccles, & Pintrich, 1996). Thus, experience in educational settings from a young age result in the development of this ranking, providing a context for dimensional comparison processes at educational choice points (Eccles, 2009). Put simply, repeated experiences of success in English across school may lead young people to see themselves as verbal rather than mathematically orientated (Parker *et al.*, 2012). This influences decisions at choice points where a young person may choose to do a university major in history or another verbally orientated field, even if they have the requisite ability to succeed in mathematics and science related areas (Parker *et al.*, 2012).

Underlying these dimensional comparison processes are implicit growth assumptions. Namely that self-concept in one domain has a negative effect on change in other self-domains. This assumption is due to the ipsative hypothesis that forms the basis of dimensional comparison theory. This assumption has, to our knowledge, never been

explicitly tested in a large scale longitudinal study. In other words, although we know that self-concept in one academic domain tends to be ipsatively associated with self-concept in other academic domains at a specific point in time, we still do not know whether these ipsative associations hold longitudinally when growth in self-concepts in multiple academic domains is considered. Such an effect would suggest that intervention efforts cannot rely on targeting one domain of academic self-concept alone. Rather they need to take a domain-specific approach, with consideration of how to intervene in the context of the complex and often counter-intuitive processes that give rise to domain specific self-evaluations.

In addition to this assumption, it is critically important to note that recent research has not only suggested that self-beliefs are hierarchical and domain specific but self-beliefs possess both trait and state residual (fluctuations from trait) components (Morin, Maïano, Marsh, Janosz, & Nagengast, 2011). Such research has shown, for example, that lower levels of self-esteem are associated with greater fluctuations in self-esteem (i.e., greater state residual; Morin, Maïano, Marsh, Nagengast, & Janosz, 2013). This research has not been conducted on domain specific self-concept however, and it is thus unclear whether such factors consist of both trait and state residual components and whether relationships between domain specific self-concept are consistent or different at trait and state-residual levels.

Academic self-concept growth trajectories

Self-concept has been posited as a critical aspect of the social and emotional development of children by a number of researchers (Davis-Kean & Sandler, 2001; Marsh, Debus, & Bornholt, 2005; Marsh, Ellis, & Craven, 2002). Researchers have been particularly interested in charting how self-concept develops with age. Marsh *et al.* (2002) demonstrated that preschool aged children were able to distinguish between multiple dimensions of self-concept at a younger age than suggested by previous research. However, measures of English and math self-concepts were highly correlated ($r = .73$) in comparison to the typical small correlations observed between math and English self-concepts evidenced in older adolescence (Marsh, 1986; Marsh & Hau, 2004; Möller *et al.*, 2009). This suggests that growth in self-beliefs over the course of development results in self-concept factors which are increasingly differentiated as children move into high-school (see also Eccles & Wigfield, 2002; Eccles *et al.*, 1993; Wigfield & Eccles, 2002; Wigfield *et al.*, 1996). However, Marsh and Ayotte (2003) found that increasing differentiation did not explain all results. Instead, they proposed a differential distinctiveness hypothesis, which argued that with increasing age and cognitive development, there are counterbalancing processes of self-concept integration and differentiation. According to Marsh and Ayotte, integration occurs when closely related areas of self-concept (e.g., cognitive and affective components of math self-concept) become amalgamated, while differentiation refers to the increasing differentiation of disparate areas of self-concept (e.g., math and English self-concepts). This

suggests that dimensional comparison processes are likely to become more important as children get older. However, it is not clear how this dual process develops, or at what ages it reaches its final equilibrium. Researchers have also investigated how self-concept varies with age. Marsh (2007) summarises the available research suggesting there is a general and worrying decline across schooling, though more complex (e.g., initial decline and then a trend towards recovery in late adolescence early adulthood) patterns of growth have been observed (Marsh, 1989).

Hypotheses

The literature on dimensional comparison and self-concept growth during adolescence suggest a number of hypotheses.

Hypothesis 1: Dimensional comparison theory suggests that self-concept levels in one domain will have a negative effect on change in self-concept in another domain. According to research which shows increasing differentiation between self-concept domains as a function of age, we can also expect these negative effects to increase with age up to an unknown point where a state of equilibrium between differentiation and integration processes have been reached. We explore these hypotheses via a series of ACLM.

Hypothesis 2: Both math and English self-concept will display a decline over the course of high school. We leave as a research question whether this decline is linear or whether there is evidence of more complex growth patterns as suggested by Marsh (1989). We also leave as a research question whether dimensional comparison processes will lead to growth trajectories in math and English being negatively related: That is, that a more significant decline in one self-concept dimension dampens or reverses declines in the other domain. These hypotheses will be explored via a series of LGC models.

Hypothesis 3: In relation to recent research which suggests that self-beliefs have both state and trait structures (e.g., Morin *et al.*, 2013), we estimate whether the hypothesised negative relationship between self-concepts across multiple domains is consistent across state residual and trait components of self-concept. Indeed, it is possible that the relationship between state residual components and trait components of two constructs can be quite distinct and even different in sign.

In all cases, we consider the effect of achievement in math and English as a time varying covariate. This is not only a critical control variable but allows us to test whether the IE model hold for state and state residual components of academic self-concept.

Methodological issues: Longitudinal modelling

Models of growth and change

In the current research we test a previously untested assumption of the IE model. This assumption however, provides a useful context for illustrating different approaches to the analysis of longitudinal data. Thus, we use autoregressive cross-lag (ACL) models to provide evidence of how self-concept in one domain can be used to predict the degree and direction of change in another component of self-concept; latent growth curve models (LGC) to explore the relationship between growth trajectories in math and English self-concept; and autoregressive latent trajectory models (ALT) to explore growth and change in English and mathematical self-concept at the trait and state residual level. We stress that each of these approaches focuses on and provides different information about growth and change processes.

ACL

Autoregressive cross-lag models are common methods used to consider temporal ordering of constructs in order to distinguish between alternative causal hypotheses, or directionality of the associations between constructs (i.e., a predicts changes in b; b predicts changes in a; or a and b are reciprocally related). This model's focus is on the relations between one construct at a time point T on change in another construct observed to occur between time point T and $T + 1$. In this way ACL models explore patterns as a series of longitudinal associations between constructs rather than to investigate the pattern of growth over time that characterizes any one construct, or the associations between growth patterns of multiple constructs. However, an increasingly frequent approach when multiple waves are present is to test whether the interrelationships between constructs in an ACL model have reached a developmental equilibrium – that is, whether the effect of one variable on another is consistent across time lags (see Figure 1). For example, Marshall, Parker, Ciarrochi, and Heaven (2013) recently showed that self-esteem predicted perceptions of social support (but not the other way around) and that this system had attained equilibrium at least by junior high school.

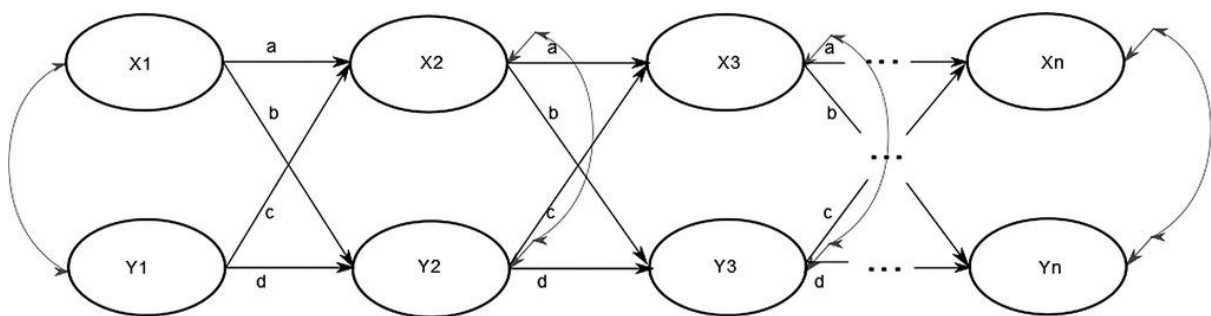


Figure 1

Autoregressive cross-lagged model. Letters represent paths that are to be constrained to equality in a cohort sequential model (for overlapping lags) or developmental equilibrium model (for all lags). Paths in black represent regression estimates and paths in grey equal covariances.

LGC

Autoregressive cross-lag models are limited to considering temporal ordering and generally give fairly limited indications of individual growth over time. Conversely, LGC models provide considerable flexibility in estimating growth trajectories over time including linear, various polynomial, and other complex growth patterns (see Figure 2; see Diallo & Morin, 2015; Diallo, Morin, & Parker, 2015; Ram & Grimm, 2007). These models decompose variance in the repeated measures of a construct into intercept and slope components respectively, reflecting the initial level and growth in the estimated developmental trajectories (see Duncan, Duncan, & Strycker, 2006 for an applied overview). Tests of significance for both the mean and the variance of the intercept and slope components are typically provided as indication of the significance of the observed growth component (mean) and of the presence of inter-individual variability on this growth component (variance) present within the sample. Additionally, comparisons of nested models can be used to compare trajectories, allowing for a test of whether self-concept shows a linear or non-linear trajectory. For the cohort sequential design used here (see below), instead of specifying a linear growth component as it is typically done in classical LGC models through fixing the loadings of the repeated measures on the slope factor to be equal to 0, 1, 2, 3 for each cohort (reflecting the passage of time); successive models were fitted in each cohort to reflect the overlap in grade levels, such that loadings of the repeated measures on the slope factors were specified to be: 0, 1, 2, 3 in Year 7; 2, 3, 4, 5 in Year 8; 4, 5, 6, 7 in Year 9; and 6, 7, 8, 9 in Year 10 (see Brodbeck, Bachmann, Croudace, & Brown, 2012 for an example).

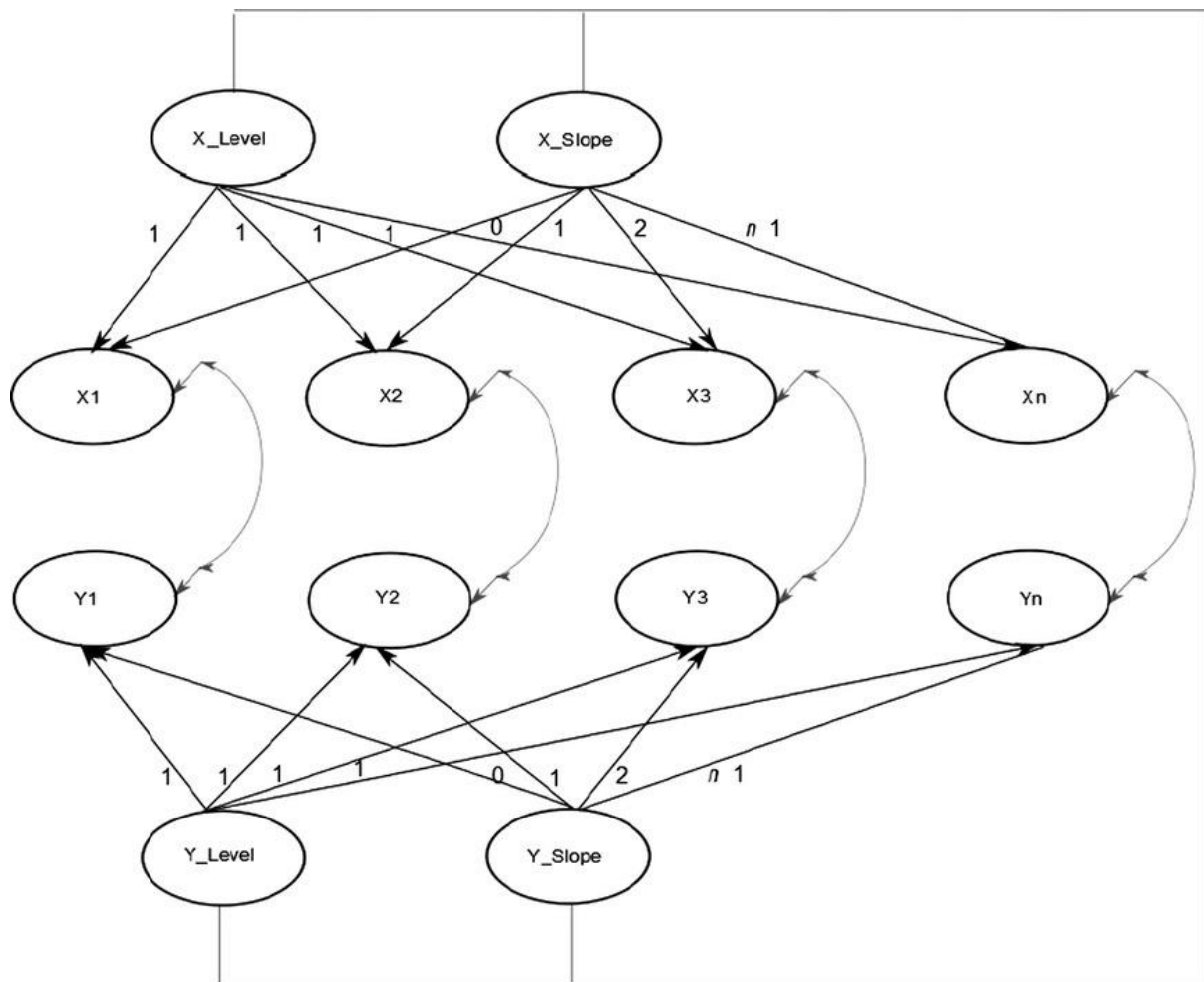


Figure 2

Latent growth curve model. Numbers represent specified loadings for level/intercept and the linear trajectory components of the model (see Duncan *et al.*, 2006 for representations of more complex growth patterns). Paths in black represent regression estimates and paths in grey equal covariances.

ALT

Both ACL and LGC models potentially conflate state and trait components of change in a construct over time. ACL models estimate the relationship between state variables. That is the effects of one construct at a given point in time on levels of the other construct at a later time points, over and above the stability of this other construct over time (i.e., change in state levels from one time wave to another). In contrast, LCM models focus on trait components. That is the estimated initial levels and growth trajectory and the relations between constructs fully at the trait level. ALT models combine elements of both ACL and LGC models (Figure 3; see Bollen & Curran, 2004, 2006; for an overview). Here observed variables are decomposed into three latent components (1) a trait, (2) a state residual, and (3) a measurement error. An individual's given state level (e.g., their level of self-concept at a given point in time) is equal to the trait component plus the state residual component, where

the state residual refers to movements at a given time-point above or below trait trajectories (Steyer, Geiser, & Fiege, 2012). The ALT model relies on an LGC specification to estimate the trait component. Then, an ACL specification is modelled directly from the state residual component of the time-specific measures over and above their trait component (see Morin *et al.*, 2011). This allows us to address questions like “is there a relationship between variables in terms of their time specific movements away from the trait trajectory”. In this case we ask questions such as: (1) is there a negative relationship between initial levels of mathematics self-concept on trajectories in English self-concept (and vice versa); and (2) does this relationship differ in size and/or direction for the time-wave to time-wave fluctuations from this overarching trajectory. The use of achievement as a time-varying covariate also needs to be interpreted in this light. For the ACLM the time-varying covariates test the effect of achievement on self-concept fully at the state level. For the LGC and ALT models however, the effect of state achievement is used to predict state residual components of self-concept (i.e., that achievement at a given time point predicts fluctuations from tend in self-concept).

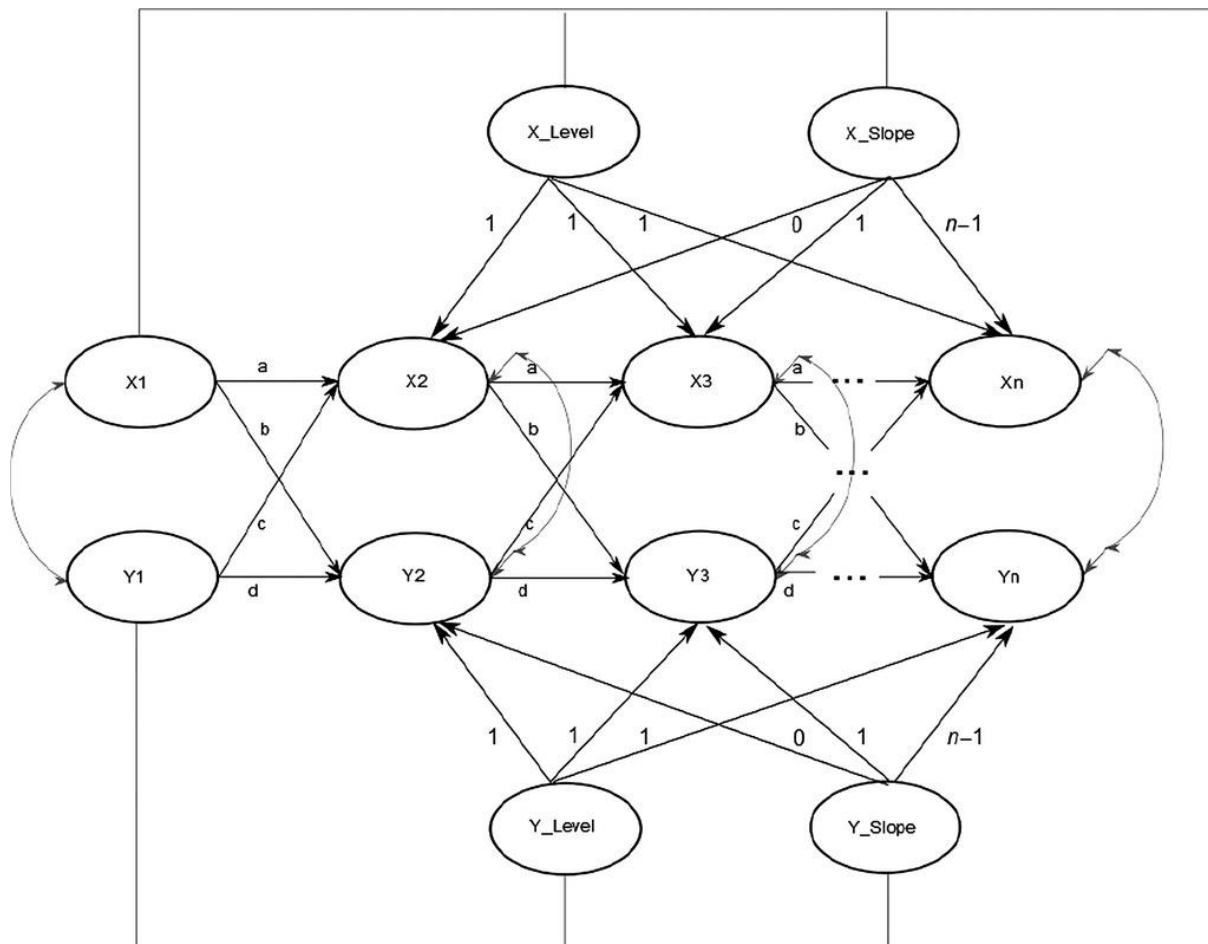


Figure 3
Autoregressive latent trajectory model. Numbers represent specified loadings for level/intercept and the linear trajectory components of the model (see Duncan *et al.*, 2006 for

representations of more complex growth patterns). Paths in black represent regression estimates and paths in grey equal covariances. Letters represent paths that are to be constrained to equality in a cohort sequential model (for overlapping lags) or developmental equilibrium model (for all lags).

Cohort sequential designs

Increasingly, educational research is relying on large scale longitudinal datasets of school age children. Due to the nature of research funding, however, many of these databases extend only over three or fewer years, making it difficult to fully explore the growth of key educational constructs over the course of major developmental periods like high-school/adolescence. The use of cohort sequential designs, however, in which multiple waves of data are simultaneously collected from multiple age cohorts, provides researchers with a feasible and cost-effective alternative to explore growth over the course of an entire developmental period (see Brodbeck *et al.*, 2012; Enders, 2010; Graham, 2012). A cohort sequential, or accelerated design provides many practical advantages for applied longitudinal research (and for growth modelling in particular; see Brodbeck *et al.*, 2012; Enders, 2010; Graham, 2012). For example, simulation studies have shown that cohort sequential designs have greater power than standard longitudinal designs when the same number of time waves are collected in each cohort (Graham, 2012). In the current research four waves of data were collected, 6 months apart, for four age cohorts. This cohort sequential design thus provides a total of 10 waves of data covering all but the final year of high school where two waves of data overlap between each successive cohort (see Figure 4 for structure).

	Year 7		Year 8		Year 9		Year 10		Year 11	
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
Year 7 cohort										
Year 8 cohort										
Year 9 cohort										
Year 10 cohort										

Figure 4
Cohort sequential design. Grey squares = Collected data. White square = Missing by design.

While cohort sequence designs have many advantages they also have important limitations. One such limitation is that models are often difficult to parametrize. One of the aims of this paper is to provide an applied example of how to estimate common longitudinal models in the context of this complexity. A second major limitation is the need to deal with the inevitable large amount of missing data (white cells in Figure 4). In cohort-sequential designs missing time points in all cohorts are missing due to the design of the study and not

as a function of participants' characteristics – thus fully corresponding to missing-completely-at-random assumptions of modern missing data techniques (Enders, 2010). This suggests that modern missing data techniques can provide unbiased parameter estimates even in the presence of missing data (Enders, 2010). There are essentially two approaches to estimating growth models with cohort sequential data that aim to overcome this missing by design component. A common approach is to use full information maximum likelihood estimation on data that is stacked and merged across cohorts. In other words, this approach involves re-organizing the dataset so that each participant (each line) is specified as having 10 measurement points, with 6 of those being missing (see Figure 4). In this approach however, some cells of the variance covariance matrix have zero coverage and thus multiple imputations and the estimation of some covariances becomes problematic; though full information maximum likelihood has no such concerns (see Enders, 2010).

The second approach is to make use of a multi-group approach to model estimation. Modern structural equation modelling packages are becoming increasingly powerful and flexible in estimating complex models. For models like those proposed here, we utilise a multigroup approach to fit differing models in each cohort. Essentially, models are fitted in each cohort (treated as separate groups in a multiple group design) that reflects their relative position in the developmental period of interest. In the current research, models were specified such that the first cohort reflected growth over Years 7 and 8, the second Years 8 and 9, the third Years 9 and 10, and the last Years 10 and 11. Through the inclusion of invariance constraints across these multiple groups for the overlapping time points, the resulting full model covers the high school period despite any one participant only completing data across 2 years. In the online supplemental materials that accompany this article, we provide annotated Mplus scripts for the estimation of growth curve models and autoregressive latent trajectory models using a cohort sequential design (see also Brown, Croudace, & Heron, 2011).

Methodology

Participants

Participants were 2,781 (51% females; $n = 1,432$) students from eight co-educational high schools (including six comprehensive and two academically selective government high schools; mean N per school = 348 [range = 151–772]) in comparatively wealthy suburbs of Sydney Australia. For ease of interpretation, cultural background was operationalised as Anglo and non-Anglo. The Anglo category (50.5%) included parentage from Australia, Europe, and the USA. The non-Anglo category (49.5%) reported being mainly of Asian heritage (87.5% of the non-Anglo category), but also included parentage from Melanesia,

Africa, and the Middle East. Thirty-two percent of participants were from non-English speaking backgrounds.

Measures

Self-concept

The math and English subscales of the Academic Self Description Questionnaire II (ASDQII; Marsh, 1990) were used to measure self-concept. Items were measured on a six-point Likert scale with poles of *strongly disagree* and *strongly agree*. In the current research reliability was estimated using the Omega coefficient from the restrictive MIM5 model (see below) and thus estimates were constrained to be consistent over the waves of the study. Reliability was .78 for English self-concept and .80 for math self-concept.

Achievement

Math and English achievement were measured using alternative forms of the Wide Rang Achievement Test 4 (WRAT4; Wilkinson & Robertson, 2006). Alternate time waves used different forms of the test. For the purpose of this study the spelling test was used to measure English achievement and math computation were used to measure math achievement. The WRAT4 is reliable, with coefficient alphas for spelling at .95 and math computation at .94 and is used widely to assess academic achievement (Wilkinson & Robertson, 2006).

Analyses and assumptions

Across all models, self-concept constructs were specified as latent variables. For identification purposes we used a non-arbitrary metric for item loadings and intercepts allowing results to be interpreted according to the original 6-point Likert scale (see Little, Slegers, & Card, 2006; see supplementary material). All models were fitted using the robust Maximum Likelihood estimator (MLR) available in Mplus 7.11 in conjunction with Full Information Maximum Likelihood procedures to handle missing data (Muthén & Muthén, 1998–2012). In each model, math and English achievement were used as a time-varying covariate with only within wave relationships estimated.

Model fit was evaluated using three goodness-of-fit indices; RMSEA, CFI and TLI. A good fit using the incremental fit CFI and TLI is indicated by values of .95 and greater and an acceptable fit is indicated by values between .90 and .94. A close fit using the absolute fit RMSEA was indicated by values of $\leq .06$ and an acceptable fit is indicated by values between .06 and .08. In addition, we provide a series of information criteria (AIC, BIC and sample size adjusted BIC).

Longitudinal models have clear invariance assumptions. ACL models assume weak factorial invariance (model MIM2) in which factor loadings are assumed to be equivalent across time waves. Both LGC and ALT models assume a minimum of strong factorial invariance (MIM3) in which loadings and item intercepts are assumed to be equal across time waves. In a cohort sequential design, these parameters are thus assumed to be invariant across both time waves and cohorts. Furthermore, to estimate the growth trajectories from all time waves and cohorts, cohort sequential designs also assume that overlapping latent means (e.g., T3 and T4 of the Year 7 cohort with T1 and T2 of the Year 8 cohort) are invariant across cohorts. Thus model MIM4 constrains loadings and intercepts to be invariant over time, and latent means for overlapping time points to be invariant across time points – providing an additional test of cohort-specific historical effects. This is an inherent assumption of cohort sequential designs for models involving latent means but is rarely tested in applied research. Finally, given the complexity of estimation for ACL, LGC, and ALT models, more parsimonious models significantly aid convergence. Thus, we also tested for strict measurement invariance in which loadings, intercepts, overlapping latent means, and item residuals were constrained to be equivalent (MIM5). Support is demonstrated for successive nested models if the fit is reduced by $<.01$ for CFI and TLI and $<.015$ for the RMSEA (Chen, 2007; Cheung & Rensvold, 2002). We note that while these cut-off criteria have often been used irrespective of models types in the applied literature, simulation research suggests that different fit indices are more sensitive to misspecification of some model components than others. For example, Fan and Sivo (2005) suggest that the CFI and TLI may not be as sensitive to misspecification of measurement structures than they are for factor loadings. Clearly this has different implications for the various models under investigation here. For this reason we include a variety of fit information in addition to the fit indices including delta χ^2 values, AIC, BIC, and Adjusted BIC to aid readers in assessing our judgments relating to comparing model fit.

As can be seen from Table 1, as the model increased in parsimony from MIM1 up to MIM5, the fit of the models remained fully satisfactory and changes in fit indices were minimal. As such MIM5 was used as a basis for all models estimated later as this was the most parsimonious measurement model and thus provided particular advantages in terms of convergence (see Diallo *et al.*, 2014) and computational speed.

Table 1. Measurement invariance

Model	χ^2	Df	$\Delta\chi^2$	RMSEA	CFI	TLI	AIC	BIC	A-BIC
MIM1: Configural	7,368	3,920	-	.036	.950	.943	277,470	283,258	280,157
MIM2: Loadings (Time + Cohort)	7,623	4,070	255 (18) ^a	.035	.949	.943	277,483	282,382	279,758
MIM3: MIM2 + Intercepts (Time + Cohort)	8,230	4,220	607 (150) ^a	.037	.942	.938	277,857	281,866	279,718
MIM4: MIM3 + Matching Cohort Means	8,257	4,232	27 (12)	.037	.942	.938	277,861	281,799	279,689
MIM5: MIM4 + Invariant Item Residuals	8,587	4,412	330 (180) ^a	.037	.940	.938	278,181	281,052	279,514

Note

a $p < .001$. For $\Delta\chi^2$ degrees of freedom in brackets. A-BIC is sample size adjusted BIC. For all information criteria (i.e., AIC, BIC, and A-BIC) smaller values are associated with better fitting models.

Results

Growth/change model

ACL

Three nested ACL models were fitted to the data. These models explored the effect of one self-concept factor at one time point (time T) on the level of growth in the other self-concept factor occurring between this same time point and the next (between time T and $T + 1$). ACLM1 included no constraints across time or cohort. This model would be appropriate if there were major differences in the effects observed across cohorts and time waves. If this model was substantially better fitting than the other models this would suggest problems in treating the data in a cohort sequential manner. However, as Table 2 illustrates the considerably more parsimonious models ACLM2 (with the effects estimated between overlapping time lags across cohorts constrained to be the same) and the even more parsimonious ACLM3 (the developmental equilibrium model with predictions constrained to be the same across time waves and cohorts) did not result in substantial change in fit (nevertheless please see supplementary material for major parameters for models ACLM2). Indeed, there was little change between ACLM1 and ACLM3 suggesting the relationship between math and English self-concept had reached developmental equilibrium in secondary school. From ACLM3 it was clear that the within time effects of achievement on self-concept

closely followed IE relationships with math self-concept positively predicted by math achievement ($\beta = .336, p < .001$) and negatively predicted by English achievement ($\beta = -.131, p < .001$). Likewise, English self-concept was positively predicted by English achievement ($\beta = .172, p < .001$) and negatively predicted by math achievement ($\beta = -.09, p < .001$). For the longitudinal relationships between self-concept, however, there was only significant relationship between matching domains. Thus, math self-concept at time T positively predicted math self-concept at time $T + 1$ ($\beta = .720, p < .001$) but not English self-concept ($\beta = .02, ns$). Likewise, only English self-concept at time T predicted English self-concept at time $T + 1$ ($\beta = .771, p < .001$).

Table 2. Growth models

Model	χ^2	Df	$\Delta\chi^2$	RMSEA	CFI	TLI	AIC	BIC	A-BIC
Autoregressive Cross-lag Models									
ACLM1: Unconstrained Structure	10,695	5,710	-	.035	.933	.929	317,101	322,426	319,573
ACLM2: Matching Cohort Structure Constrained	10,996	5,756	301 *** (46)	.036	.929	.926	317,376	322,429	319,721
ACLM3: Constrained Structure	11,255	5,824	269 *** (80)	.037	.927	.924	317,555	322,204	319,713
Latent Growth Curve Model									
LGCM1: Unconstrained Growth	11,300	5,945	-	.036	.928	.927	317,636	321,568	319,461
LGCM2: Linear Growth	11,344	5,961	44 *** (16)	.036	.928	.926	317,691	321,528	319,473
Autoregressive Latent Trajectory Model									
ALTM0: Growth Only	11,807	5,944	-	.038	.921	.920	318,275	322,213	320,103
ALTM1: Unconstrained State Structure	11,535	5,908	272 *** (36)	.037	.924	.922	318,024	322,175	319,951
ALTM2: Constrained State Structure	11,718	5,940	183 *** (32)	.037	.922	.921	318,171	322,133	320,011

Note

The $\Delta\chi^2$ for ALTM0 versus ALTM2 is 89 (4), $p < .001$. The $\Delta\chi^2$ for ACLM1 versus ACLM3 is 560 (126), $p < .001$. *** $p < .001$. For $\Delta\chi^2$ degrees of freedom in brackets. A-BIC is sample size adjusted BIC. For all information criteria (i.e., AIC, BIC, and A-BIC) smaller values are associated with better fitting models.

Taken together the ACLM with time-varying covariates provide evidence of a consistent IE pattern for within time achievement predicting self-concept but, against out

hypothesis, only stability effects were for self-concept predicting itself at subsequent time waves was observed with no evidence of ipsative relationships.

LGC

We fitted two multigroup cohort sequential LGC models to the data. These models explored growth in the trait trajectory of both math and English self-concept over the 10 time waves of the study, and the interrelations between the trait trajectories of these constructs. The first LGCM1 model estimated a free (latent basis) growth trajectory (see Morin *et al.*, 2013; Ram & Grimm, 2009) with few constraints and under which all possible polynomial growth models were nested. In this model, the loading of the first time wave on the slope factor (in the Year 7 cohort) was constrained to be zero, the loading of the last time wave on the slope factor (in the Year 10 cohort) was constrained to be one, and all other loadings were freely estimated, allowing for the estimation of a purely non-linear growth trajectory against which to compare alternative model specifications. The only other constraints included in this model involved the loadings of the overlapping time points on the slope factor, which were constrained to be equal across cohorts. This model was compared to LGCM2 which tested a purely linear growth model. As can be seen from Table 2, there was little difference between the two models and thus the more parsimonious linear growth model was retained.

For both math and English self-concept the mean slope factor was significantly negative (math: $\mu = -.05$, $p < .001$; English: $\mu = -.03$, $p < .001$). This translates to an approximate .12 and .07 standard deviation decline in math and English self-concept for each wave across high-school. This translates to over one standard deviation unit decline in math and three quarters of a standard deviation in English achievement from the beginning to the end of high school. There was also evidence of significant inter-individual variability in the level of linear growth over time (math: $\sigma = .01$, $p < .001$; English: $\sigma = .01$, $p < .001$). For this model there was no evidence for the IE model for the relationships between the growth components of self-concept (i.e., math self-concept intercept with English self-concept growth). Likewise, only the matching paths from achievement predicting self-concept were significant (Math: $\beta = .06$, $p < .05$; English: $\beta = .05$, $p < .05$). Finally, there were significant within domain correlations between intercept and slope factors (Math: $r = -.429$, $p < .001$; English: $r = -.301$, $p < .001$) consistent with the regression to the mean phenomena.

Taken together, the results from the growth model suggested a steady and worrying decline in self-concept in both domains. There was, however, little evidence in support of the central ipsative hypothesis for either the time-varying covariate or growth factors.

ALT

On the basis of the previous results, we estimated ALT models assuming a linear growth process and invariance of the predictive paths across cohorts. These models tested the relationship between growth in the trait trajectory of both English and math self-concept, while also considering interrelations between the state residual associated with these two constructs (fluctuations from trait trajectory) occurring over time. A LGC model reparametrized to be nested under the ALT models was also fitted to the data (ALTM0). The difference between ALTM0 and ALTM1 was within the cut-off criteria outlined in the methodology. However, given that the ALT and LGC models focus on different research questions and provide different information we choose to explore the ALT models in detail. ALTM1 allowed the relationship between the state components to differ across time. The second model (ALTM2) hypothesised a developmental equilibrium model between the state components consistent with suggestions by Bollen and Curran (2004, 2006); see also Morin *et al.*, 2011) and constrained the autoregressive cross lagged predictions between the state components to be invariant across time waves. The time varying covariates were also constrained to be invariant for each wave. The fit of ALTM2 was not substantially worse than ALTM1 and thus this simpler model was retained as the final model (Parameter estimates for model ALTM1 can be found in the supplementary material).

The results from this model provide a particularly interesting picture. Firstly, all but one of the relationships between the growth components was non-significant. The only significant correlation was positive and between the intercept for English self-concept and the slope for math self-concept ($r = .635, p < .001$). These results should be interpreted with caution, as while the linear downward trend remain significantly negative, and of comparable size to LGCM2 (math: $\mu = -.050, p < .001$; English: $\mu = -.039, p < .001$), the variance components were extremely small (math: $\sigma = .002, p < .05$; English: $\sigma = .001, p = .211$). Evidence for the central hypothesis for the state residual components was stronger. Indeed, the within wave IE model for achievement predicting state residual self-concept was partially supported in ALTM2 for the matching (Math \rightarrow Math: $\beta = .369, p < .001$; English \rightarrow English: $\beta = .121, p < .001$) and for the path from math to English self-concept ($\beta = -.063, p < .05$). The effect of English achievement on math self-concept was not, however, significant ($\beta = .037, ns$). There was consistent evidence of IE relationships between the state residuals of math and English self-concept for both the matching domain (Math \rightarrow Math: $\beta = .085, p < .001$; English \rightarrow English: $\beta = .078, p < .001$) and the non-matching domain (Math \rightarrow English: $\beta = -.067, p < .001$; English \rightarrow Math: $\beta = -.104, p < .001$).

Taken together, the results for the ALT models suggested considerable complexity in the growth patterns for self-concept. In particular, there was some evidence of positive

relationships between initial levels of self-concept in one domain and trajectories in the other domain. Alternatively, however, the results for the state residual components of the model and for the time-varying covariates were in keeping with the ipsative assumptions of the IE model.

Discussion

The current research explored three alternative models of growth to test a series of hypotheses and research questions about the relationship between growth in math and English self-concept. Consistent with previous research on academic self-concept (see Marsh, 2007), both English and math self-concept showed a statistically and practically significant decline over high-school. Importantly, evidence suggested that this trend was linear with little evidence of recovery in late high-school. There was no evidence in favour of the hypothesised negative relationship between the two self-concepts over time in LGC and ACL models was mixed. For the time varying covariates, the ACL model had evidence in favour of the IE model for achievement predicting self-concept but there was no evidence in favour of achievement predicting state residual self-concept in the growth curve models. The ALT models did produce evidence in favour of the hypothesised relationships. In particular, there was evidence of IE ipsative relationships between state residual math and later state residual English self-concept. Likewise, the relationships between state achievement and state residual self-concept were consistent with the IE model. Surprisingly and against hypotheses, there was a significant positive relationship between the trait components of self-concept between initial English self-concept and growth in math self-concept. However, we choose not to consider this significant effect further given that it is based on an extremely small near-zero variance in the slope of math self-concept (.002).

Implications for theory

The negative relationship between growth in English and math self-concept domains is an implicit assumption of dimensional comparison theory but has not been explicitly tested. Likewise, to date no theory or research has considered differential functioning for state and trait components of academic self-concept. The results here provide some support for the implicit negative association of dimensional comparison theory but only for the state residual components of self-concept. These results for the state residual components of self-concept may suggest that dimensional comparisons relating to fluctuations from trait trajectories in self-concept may act as a mechanism to return overall academic self-concept back to a stable growth after events (e.g., doing much better than expected on a math test, or failing an English test that you believed you would do well on) create changes in self-perceptions. Put simply, an event that leads to a student doubting their ability in math may result in a rise in their English self-concept, thus helping to promote a stable growth trajectories in academic

self-concept. In this way, the extreme multi-dimensionality of academic self-concept and dimensional comparisons may help act as resilience mechanisms. This is interesting when interpreted in the light of Morin *et al.*'s (2013) self-equilibrium hypothesis in which individuals with low self-esteem tend to be characterised by highly unstable growth trajectories, while high levels of self-esteem tend to be highly stable. Morin *et al.*'s study also suggested the presence of a sub-group of low-unstable self-esteem students that did change to a high-stable self-esteem trajectory over time, with this switch occurring approximately in the middle of the high school period. The authors suggest this may be due to physiological and psychological maturation, consolidation of identity, self-acceptance, or positive changes in life circumstances. More specifically, among the school-related factors considered in their study, they note that GPA and positive perceptions of the school's educative climate predicted a greater likelihood of experiencing such a positive switch, particularly for boys, whereas more positive perceptions of the school's justice and bonding climates and lower levels of loneliness proved more important for girls. Our research may also suggest that development in dimensional comparison processes may also play a role. Thus, future research would both need to replicate the results found here and should explore whether those low in self-concept are characterised by a relative absence of dimensional comparison processes in state residual components of academic self-concept.

Implications for practice

One of the disconcerting results from this study was the relatively drastic decline in self-concept across adolescence for both math and English self-concept. While this is consistent with previous research (see Marsh, 2007 for a review), the size of the decline and its presence in a sample that consisted of both selective students and comprehensive students from largely high SES backgrounds is surprising. This slide in self-concept is of particular concern given a wealth of research, which notes the effect of self-concept on later achievement (see Marsh, 2007 for a review) and educational and occupational choices (Davis, 1966; Marsh, 1991; Nagy *et al.*, 2008, 2006; Parker *et al.*, 2012, 2014, in press).

The findings relating to dimensional comparison at the state residual level also suggest a need to consider implications for applied settings. Indeed, if dimensional comparisons do act as mechanisms to protect against rapid declines or increases in academic self-concept, this may suggest that they play a positive role in the lives of young people. However, given that the long-term trajectories are on average negative, there may be a need to disrupt dimensional comparison processes in order to raise young people's stable self-perceptions. However, this remains speculative until additional research first verifies the presence of negative dimensional comparison processes at the state residual level and then considers whether it operates as a regulatory framework for long-term academic self-concept trajectories.

Limitations

It is important to note that there are limitations in the present study, which suggest directions for future research. In particular, the multigroup approach to cohort sequential designs used here illustrates a means for applied researchers to study development over larger time periods in a cost effective manner. Indeed, the approach taken here can easily be extended to a range of other growth models and can handle both continuous and categorical data and thus is a viable approach for applied longitudinal studies in educational research. However, it is important to note that there are limitations. In particular, this approach is largely suited to growth models and leads to a great deal of complexity in model development. In addition, it is presently difficult to address some issues that are likely to be of interest to educational and developmental psychologists such as the effect of prior covariates (e.g., achievement levels in primary school). Thus, there is a need for further development in this area.

Another limitation is that the distinction between state and trait components in this research is based on 6-month time lags between waves. Shorter time frames would have provided greater scope for considering dynamic processes in the development and maintenance of academic self-concept. In addition, an index cataloguing major events in the participants' educational experiences would have provided greater details on the processes that lead to fluctuations in state self-concept.

Conclusions

The current research explored the relationship between growth in math and English self-concept to test implicit assumptions in dimensional comparison theory. We compared and contrasted three models of growth. ALT models suggested the presence of dimensional comparison in fluctuations in self-concept. This has potential implications for self-concept theory and provides new avenues for future research. The research also provides an illustration of estimating complex growth models using cohort sequential data in an applied educational setting that has wide applicability for educational research in relation to other academic outcomes.

Acknowledgement

This research was funded in part by grants from the Australian Research Council (DP0877868, DP130102713, and DE140100080).

References

- Albert, S. U. (1977). Temporal comparison theory. *Psychological Review*, 84, 485–503. doi:10.1037/0033-295X.84.6.485
- Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models a synthesis of two traditions. *Sociological Methods & Research*, 32, 336–383. doi:10.1177/0049124103260222
- Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
- Brodbeck, J., Bachmann, M. S., Croudace, T. J., & Brown, A. (2012). Comparing growth trajectories of risk behaviors from late adolescence through young adulthood: An accelerated design. *Developmental Psychology*, 49, 1732–1738. doi:10.1037/a0030873
- Brown, A., Croudace, T., & Heron, J. (2011). Longitudinal data modelling: Sequential cohort design [Course handout]. Retrieved from <http://www.restore.ac.uk/appliedpsychometrics/Tutorial.materials/Intro.to.longitudinal.modelling/Part%203%20Accelerated%20cohort.pdf>
- Chen, F. F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504. doi:10.1080/10705510701301834
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. doi:10.1207/S15328007SEM0902_5
- Davis, J. (1966). The campus as a frog pond: An application of the theory of relative deprivation to career decisions of college men. *American Journal of Sociology*, 72, 17–31. doi:10.1086/224257
- Davis-Kean, P. E., & Sandler, H. M. (2001). A meta-analysis of measures of self-esteem for young children: A framework for future measures. *Child Development*, 72, 887–906. doi:10.1111/1467-8624.00322
- Diallo, T. M. O., & Morin, A. J. S. (2015). Power of latent growth curve models to detect piecewise linear trajectories. *Structural Equation Modeling*. Advance online publication. doi:10.1080/10705511.2014.935678
- Diallo, T. M. O., Morin, A. J. S., & Parker, P. D. (2014). Statistical power of latent growth curve models to detect quadratic growth. *Behavioural Research Methods*. Advanced online publication. doi: 10.3758/s13428-013-0395-1
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. New York, NY: Lawrence Erlbaum.

- Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly*, 18, 585–609. doi:10.1111/j.1471-6402.1994.tb01049.x
- Eccles, J. S. (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, 78–89. doi:10.1080/00461520902832368
- Eccles, J. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, 35, 195–201. doi:10.1177/0165025411398185
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132. doi:10.1146/annurev.psych.53.100901.135153
- Eccles, J. S., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's self and task perceptions during elementary school. *Child Development*, 64, 830–847. doi:10.2307/1131221
- Enders, C. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.
- Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling*, 12, 343–367. doi:10.1207/s15328007sem1203_1
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. doi:10.1177/001872675400700202
- Graham, J. W. (2012). *Missing data: Analysis and design*. New York, NY: Springer. doi:10.1007/978-1-4614-4018-5
- 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, 124–136. doi:10.1037/0022-0663.95.1.124
- James, W. (1890/1960). *The principles of psychology* (Vol. 1). New York, NY: Dover Publications.
- Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling*, 13, 59–72. doi:10.1207/s15328007sem1301_3
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23, 129–149. doi:10.3102/00028312023001129
- Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, 81, 417–430. doi:10.1037/0022-0663.81.3.417

- Marsh, H. W. (1990). ASDQII: Academic self-description questionnaire II [Psychometric scale]. Retrieved from <http://www.uws.edu.au/cppe/research/instruments/asdqii>
- Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal*, 28, 445–480. doi:10.3102/00028312028002445
- Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester: British Psychological Society.
- Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology*, 95, 687–706. doi:10.1037/0022-0663.95.4.687
- 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, 133–163. doi:10.1111/j.1745-6916.2006.00010.x
- Marsh, H. W., Debus, R., & Bornholt, L. (2005). Validating young children's self-concept responses: Methodological ways and means to understand their responses. In D. M. Teti (Ed.), *Handbook of research methods in developmental science* (pp. 138–160). Oxford: Blackwell Publishers. doi:10.1002/9780470756676.ch8
- Marsh, H. W., Ellis, L. A., & Craven, R. G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental Psychology*, 38, 376–393. doi:10.1037/0012-1649.38.3.376
- 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, 56–67. doi:10.1037/0022-06663.96.1.56
- Marshall, S. L., Parker, P. D., Ciarrochi, J., & Heaven, P. C. (2013). Is self-esteem a cause or consequence of social support? A 4-year longitudinal study. *Child Development*. Advanced online publication. doi: 10.1111/cdev.12176
- Möller, J., & Köller, O. (2001). Frame of reference effects following the announcement of exam results. *Contemporary Educational Psychology*, 26, 277–287. doi:10.1006/ceps.2000.1055
- Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, 120, 544–560. doi:10.1037/a0032459
- 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, 1129–1167. doi:10.3102/0034654309337522

- Morin, A. J., Mañano, C., Marsh, H. W., Janosz, M., & Nagengast, B. (2011). The Longitudinal interplay of adolescents' self-esteem and body image: A conditional autoregressive latent trajectory analysis. *Multivariate Behavioral Research*, 46, 157–201. doi:10.1080/00273171.2010.546731
- Morin, A. J. S., Mañano, C., Marsh, H. W., Nagengast, B., & Janosz, M. (2013). School life and adolescents' self-esteem trajectories. *Child Development*, 84, 1967–1988. doi:10.1111/cdev.12089
- Muthén, L. K. & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.) Los Angeles, CA: Muthén & Muthén
- Nagy, G., Garrett, J., Trautwein, U., Cortina, K., Baumert, J., & Eccles, J. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In H. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 115–143). Washington, DC: American Psychological Association.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, 12, 323–345. doi:10.1080/13803610600765687
- Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, 34, 29–48. doi:10.1080/01443410.2013.797339
- Parker, P. D., Nagy, P. D., Trautwein, U., & Lüdtke, O. (in press) The internal/external frame of reference as predictors of career aspirations and university majors. In J. Eccles & I. Schoon (Eds.), *Gender differences in aspirations and attainment: A life course perspective*. Cambridge University Press.
- Parker, P. D., Schoon, I., Tsai, Y., Nagy, G., Trautwein, U., & Eccles, J. (2012). Achievement, agency, gender, and socioeconomic background as predictors of post-school choices: A multi-context study. *Developmental Psychology*, 48, 1629–1642. doi:10.1037/a0029167
- Ram, N., & Grimm, K. J. (2007). Using simple and complex growth models to articulate developmental change: Matching theory to method. *International Journal of Behavioral Development*, 31, 303–316. doi:10.1177/0165025407077751
- Ram, N., & Grimm, K. J. (2009). Methods and measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, 33(6), 565–576. doi:10.1177/0165025409343765
- Steyer, R., Geiser, C., & Fiege, C. (2012). Latent state-trait models. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol 3: Data analysis and research publication* (pp. 291–308). Washington, DC: American Psychological Association. doi:10.1037/13621-014

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs and values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 92–120). San Diego, CA: Academic Press. doi:10.1016/B978-012750053-9/50006-1

Wigfield, A., Eccles, J. S., & Pintrich, P. R. (1996). Development between the ages of 11 and 25. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 148–185). New York, NY: Macmillan.

Wilkinson, G. S., & Robertson, G. J. (2006). *WRAT4: Wide Range Achievement Test [Standardised Achievement Test]*. Lutz, FL: Psychological Assessment Resources.