the british psychological society
promoting excellence in psychology

British Journal of Educational Psychology (2020), 90, 77–91

www.wileyonlinelibrary.com

Re-examining the reciprocal effects model of self-concept, self-efficacy, and academic achievement in a comparison of the Cross-Lagged Panel and Random-Intercept Cross-Lagged Panel frameworks

Richard A. Burns[1]*, Dimity A. Crisp[2] and Robert B. Burns[3]

[1]Centre for Research on Ageing, Health and Wellbeing, Research School of Population Health, The Australian National University, Canberra, Australian Capital Territory, Australia

[2]Centre for Applied Psychology, Faculty of Health, University of Canberra, Australian Capital Territory, Australia

[3]Department of Education, University of Sunshine Coast, Sippy Downs, Queensland, Australia

Background. The cross-lagged panel (regression) model (CLPM) is the usual framework of choice to test the longitudinal reciprocal effects between self-concept and achievement. Criticisms of the CLPM are that causal paths are over-estimated as they fail to discriminate between- and within-person variation. The random-intercept cross-lagged panel model (RI-CLPM) is one alternative that extends the CLPM by partialling out between-person variance.

Aims. We compare analyses from a CLPM and a RI-CLPM which examine the reciprocal relationships between self-concept, self-efficacy, and achievement and determine the extent CLPM estimates are inflated by between-person variance.

Sample(s). Participants (n = 314) were first-year undergraduate psychology students recruited as part of the STudent Engagement with Education and Learning (STEEL) project.

Methods. Participants completed measures of self-efficacy and self-concept prior to completing fortnightly quiz assessments.

Results. Cross-Lagged Panel (regression) Model estimates are likely over-estimated in comparison with RI-CLPM estimates. Cross-Lagged Panel (regression) Model analyses identified a reciprocal effects relationship between self-concept and achievement, confirming established literature. In RI-CLPM analyses, these effects were attenuated and a skill development association between achievement and self-concept was supported. A reciprocal relationship between self-efficacy and achievement was supported. Better model fit was reported for the RI-CLPM analyses.

*Correspondence should be addressed to Richard A. Burns, Centre for Research on Ageing, Health and Wellbeing, Research School of Population Health, The Australian National University, Building 54, Mills Road, Canberra, 0200 ACT, Australia (email: Richard.Burns@anu.edu.au).

DOI:10.1111/bjep.12265

Conclusions. Prior findings relating to the reciprocal effects of self-concept and achievement need to be reconsidered. Whilst such a relationship was supported in a CLPM analysis in this study, within an RI-CLPM framework, only achievement predicted self-concept. However, in both CLPM and RI-CLPM models a reciprocal effects model of self-efficacy and achievement was supported.

In educational research, establishing cause and effect relationships between self-attitudes and achievement has long been a primary focus (Arens *et al.*, 2016, 2017; Becht *et al.*, 2017; Calsyn & Kenny, 1977; Guay, Marsh, & Boivin, 2003; Marsh, Craven, & Debus, 1999; Marsh & Yeung, 1997; Pietsch, Walker, & Chapman, 2003; Valentine & DuBois, 2005). Currently, the most ubiquitous model for examining dynamic relationships between constructs is the cross-lagged panel model (CLPM), also referred to as the cross-lagged regression model. The CLPM framework derives its foundations from a seminal critique of the use of cross-lagged correlations to describe longitudinal changes (Rogosa, 1980). Rogosa (1980) demonstrated that spurious conclusions about the longitudinal relationship between two or more constructs could be derived if the pattern and mechanisms underlying the temporal nature in those constructs differed. Rogosa (1980) proposed the incorporation of autoregressive parameters, in which each construct under investigation was regressed on its prior state, in order to control for the within-construct temporal relationship. Since Rogosa's critique, and of particular relevance for the substantive area of focus in the current paper, the CLPM framework has become the main analytical tool educational researchers have used to describe longitudinal relationships between attitudes of self (e.g., self-concept, self-efficacy) and academic achievement.

Four main theoretical propositions have been proposed to describe the mechanisms by which self-concept and self-efficacy might relate to achievement outcomes. The *self-enhancement model* argues that self-attitudes are a major determinant of achievement, whilst the *skill development model* posits that achievement, driven by strong instrumental academic skills, drives self-concept (Calsyn & Kenny, 1977). Valentine and DuBois (2005) suggest a *null model*, perhaps more appropriately described a *common-cause* hypothesis since any purported association between self-concept and achievement is spurious and due to unmeasured and unidentified factors which drive both variables. Perhaps the most pervasive framework is the *reciprocal effects model* (Arens *et al.*, 2017; Guay *et al.*, 2003; Marsh, 1990; Marsh & Yeung, 1997; Pinxten, Marsh, De Fraine, Van Den Noortgate, & Van Damme, 2014; Seaton, Marsh, Parker, Craven, & Yeung, 2015; Seaton, Parker, Marsh, Craven, & Yeung, 2014) which combines the *skill-enhancement* and *skill development models.* The testing of these alternative hypotheses with CLPM methods in observational studies, where randomization of participants is precluded, allows, to some degree, the comparison of predictive regression parameters between constructs which are assessed on repeated occasions. But primarily, the benefit of the CLPM to test these mechanisms was the incorporation of autoregression parameters which address the longitudinal relationship within each construct which Rogosa (1980) so strongly advocated.

However, analyses even within a CLPM framework have their limitations. Recently, Hamaker, Kuiper, and Grasman (2015) have proposed the random-intercept cross-lagged panel model (RI-CLPM) as one alternative that includes the important features of the CLPM, but extends the CLPM by partialling out between- and within-person variance in repeatedly observed manifest indicators. That is, the between-person component reflects variance due to differences that exist between persons whilst the within-person component reflects variances due to changes which vary within individuals over time. An assumption of the CLPM model is that the temporal relationships or within-person change in the constructs of interest is consistent over time, and on reflection, this is unlikely to be a reasonable assumption to hold, particularly since the outcome of interest is very much an individual characteristic. To make comparison between the CLPM and RI-CLPM models easier for the reader, the more commonly implemented CLPM is displayed in Figure 1 whilst the RI-CLPM is displayed in Figure 2. It can be seen that several parameters are consistent between models. These include the autoregressive parameters α and δ for the quiz and self-referent manifest indicators, respectively, and the cross-lagged parameters β and γ in which self-referent indicators are regressed on quiz score$_{t\text{-}1}$ and where quiz scores are regressed on self-referent indicators$_{t\text{-}1}$.

The most notable differences between these models relate to the estimation of latent factors which partial out the between-person stability in the manifest indicators. However, other differences relate to the interpretation of the structural parameters. In the CLPM, the autoregressive parameters reflect the rank order stability of persons from one measurement occasion to the next. In contrast, for the RI-CLPM, the autoregressive parameters reflect the amount of within-person carry-over effect. That is, a positive autoregressive parameter reflects the likelihood that when a person scores above (or below) their average at one occasion, then their following score at the next occasion will again be above (or below) their average score. In contrast, negative autoregressive parameters reflect the scenario where a person scores above (or below) their mean at one occasion and reverses their score relative to their mean at the next occasion by scoring below (or above) their expected score. The interpretation of the cross-lagged parameters also changes. In the RI-CLPM, they now reflect whether changes from an individual's expected score on one variable are predicted from preceding deviations on a second variable and are an average of the within-person change. We note that in some respects, the interpretation of cross-lagged parameters in CLMP analyses, including earlier *reciprocal effects models,* are often poorly described. Researchers do not often explicitly state that the cross-lagged parameters within the CLMP framework are measures of change in variable x_t regressed on subjects' deviation from the group mean on variable y_{t-1}. For a full detailed discussion of these issues, we recommend Hamaker *et al.* (2015).

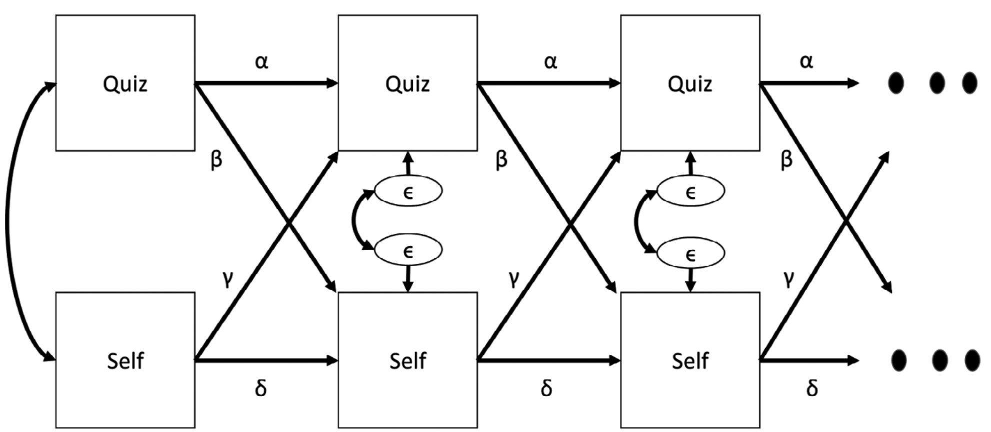

Figure 1. Cross-lagged panel model reflecting the longitudinal relationship between achievement and self-concept and self-efficacy. *Note.* Quiz reflects achievement scores; Self reflects the self-referent variables (self-efficacy) and two self-concept factors (Competence and Affect) modelled simultaneously.

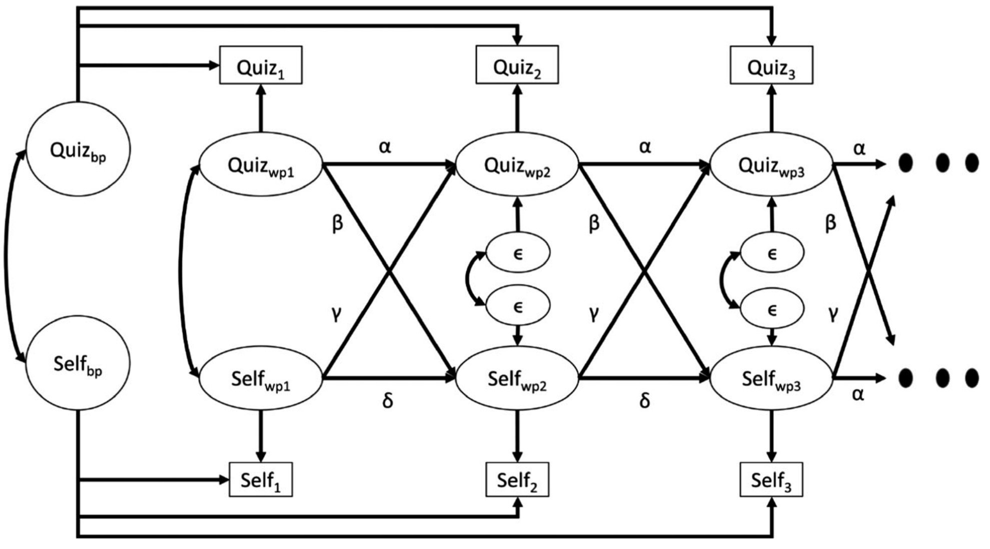

Figure 2. Random intercept cross-lagged panel model reflecting the longitudinal relationship between achievement and self-concept and self-efficacy with latent factors partialling out between-person variance. *Note*. Quiz reflects achievement scores; Self reflects the self-referent variables (self-efficacy) and two self-concept factors (Competence and Feeling) modelled simultaneously.

At the time of writing, few educational or development studies had considered the potential for erroneous conclusions drawn from CLPM which might better be described within an RI-CLPM framework. Whilst the findings of CLPM analyses are not necessarily incorrect, the interpretation of CLPM findings warrants careful consideration. Clearly, when considering individual level characteristics, it may be more important to decompose variance into between- and within-person components since it is often the improvement of individuals' self-referency and their personal achievement on which practitioners are focused. Recently, Becht *et al.* (2017) examined the cross-lagged relationship between social relationship quality with parents and peers with a child's self-concept, whilst Vangeel, Vandenbosch, and Eggermont (2018) examined the within-person variation between different aspects of self-objectification during and after adolescent. However, neither study specifically compared results from a CLPM or RI-CLPM analyses to examine the extent to which conclusions varied according to the model implemented. A literature review further failed to identify any current research specifically focusing on self-concept and achievement comparing these analytical frameworks. Therefore, the aim of the current paper is to compare the findings of CLPM and RI-CLPM analyses which examine the longitudinal relationship between self-concept, self-efficacy, and academic achievement.

Methods

Participants

Participants were recruited as part of the STudent Engagement with Education and Learning (STEEL) Project. Data were obtained from one STEEL study that comprised responses from first-year undergraduate psychology students. All students enrolled in a first-year introductory psychology course were invited to participate. Participants completed quizzes fortnightly as part of normal course assessment at which time they were also invited to complete a brief survey to assess their self-efficacy and self-concept prior to the assessment task. Following course administration, quizzes were open for students to complete their assessment for a week at a time. Participants (n = 314; response rate = 58.4%) provided between 1 and 7 (M = 4.7) observations (Nobs = 1,115) of data. The surveys were administered online through the Qualtrics Survey software. This study was approved by the University of Canberra Human Research Ethics Committee.

Measures

Measures pertinent to the current study are detailed below.

Academic self-concept

Students' academic self-concept in relation to their psychology studies was assessed using the 10-item General Academic Self-Concept Scale (Marsh & O'Neill, 1984). The wording of the scale was adapted to reflect their current psychology unit (e.g., 'I enjoy doing work for the course Psychology 102'). Items were scored on an 8-point scale from 0 (*Definitely false*) to 7 (*Definitely true*). Internal reliability for the scale was very high (α = .86). Previous literature (Abu-Hilal, Abdelfattah, Alshumrani, Abduljabbar, & Marsh, 2013; Arens, Yeung, Craven, & Hasselhorn, 2011; Arens & Hasselhorn, 2015; Arens *et al.*, 2016; Burns, Crisp, & Burns, 2018; Marsh *et al.*, 1999; Pietsch *et al.*, 2003; Pinxten *et al.*, 2014; Yang, Arens, & Watkins, 2014) has identified two oblique self-concept factors reflecting (1) Competence: perceptions of academic competence in the subject and (2) Affect: their affect or feeling about the subject. We have previously found support for a two-factor model in this sample and other student populations within the STEEL project and found that this measurement model was invariant over time (Burns *et al.*, 2018). These separate self-concept factors were used in the analyses with factor scores saved using the regression method.

Self-efficacy

Self-efficacy relating to their perceived competence in psychology quizzes was assessed using a single item 'How well do you think you can perform in the quiz this week?' This is suitable and consistent with other measures of self-efficacy which assess task-specific behaviour (Bandura, 2006; Hoeppner, Kelly, Urbanoski, & Slaymaker, 2011) particularly in education research (Gardner, Cummings, Dunham, & Pierce, 1998). Participants responded to the item on a 6-point scale from 1 (*Extremely well*) to 6 (*Extremely poorly*). Responses were reversed such that a higher score indicated higher self-efficacy.

Academic achievement

For this study, academic achievement was assessed with *quizzes* that were standard assessment practices over the semester. During the academic term, students were invited to participate in fortnightly surveys which were timed with their regular formal quiz assessments which were distributed across the academic semester and contributed 20% of their overall mark. Assessment protocols were such that the unit convenor would only count the 5 best (out of 7) quiz scores towards the final course grade. Quizzes were undertaken online and preceded by a short questionnaire. As there was no obligation for students to participate, students were provided an option to proceed directly to their quiz without completing the survey. Students were provided with immediate feedback of their quiz marks in accordance with educational practice at the University. Quiz grade scores reflected the score for each quiz out of 10.

Statistical analysis

To facilitate direct effect size comparisons and to improve model estimation, the self-concept, self-efficacy, and quiz score variables were rescaled onto similar scales by *T*-scoring each of the variables respective to their grand mean and standardizing to a mean of 50 and a standard deviation of 10. Before examining the *reciprocal effects model* between our factors of interest, we first used multi-level regression models to examine the trajectories of change in the Competence and Affect Self-Concept factors, self-efficacy, and quiz score over the term to assess the extent to which factors exhibited systematic growth over the semester. Linear, quadratic, and cubic metrics were explored and compared with a saturated means model in which time is treated as a discrete function. Then, both CLPM and RI-CLPM were implemented in a structural equation framework in MPlus v7 and were estimated with a full-information maximum-likelihood (FIML) estimator. Therefore, participants with any missing or non-response observations are retained and the FIML estimation adjusts the likelihood function so that each case contributes information on the variables that are observed. In the CLPM and RI-CLPM analyses, both the self-concept factors and self-efficacy were estimated concurrently with achievement. Comparisons between models were made by comparing the change in a range of goodness-of-fit indices (GFI). Specific cut-offs for different GFI are well described (Marsh, Hau, Balla, & Grayson, 1998; Marsh *et al.*, 2009). For each model, chi-square and change in chi-square values (Kline, 1998), Akaike (AIC) and Bayesian information criteria (BIC), the comparative fit index (CFI; optimal values > .90 (Bentler, 1990)), the Tucker–Lewis index (TLI; optimal values > .90 (Hu & Bentler, 1999)), root-mean-square error of approximation (RMSEA; optimal values < .06 (Hu & Bentler, 1999)), and their corresponding 95% confidence intervals are reported.

Results

Baseline descriptive statistics (mean and standard deviation) of the variables are reported in the diagonal cells in Table 1. These means and standard deviations reflect the untransformed values before T-standardizing variables and centring to each variable's grand mean. In addition, inter-variable correlations are reported below the diagonal cells. These correlations reflect the average correlation between variables over the study period and are drawn from a bootstrap of 1,000 samples with standard error (reported in the parentheses) that are adjusted for the non-independence of repeated observations. Overall, there are small to moderate correlations between the variables over time.

Exploring stability and variation in self-concept, self-efficacy, and achievement within an academic semester

A series of multi-level regression models examined the trajectories of change in the Competence and Affect Self-Concept factors, self-efficacy and quiz score over the semester to assess the extent to which factors exhibited systematic growth over time.

Table 1. Descriptive statistics of the baseline untransformed mean and standard deviations (diagonal cells) and bootstrapped correlations and standard error between the variables averaged over the study period

	Quiz score *M(SD)/r(SE)*	Self-concept: Affect *M(SD)/r(SE)*	Self-concept: Competence *M(SD)/r(SE)*	Self-efficacy *M(SD)/r(SE)*
Quiz score	8.44 (2.41)			
Self-concept: Affect	0.12 (0.02)[a], *	0.00 (0.95)		
Self-concept: Competence	0.27 (0.02)[a], *	0.58 (0.02)[a,*]	0.00 (0.94)	
Self-efficacy	0.22 (0.02)[a], *	0.26 (0.02)[a,*]	0.47 (0.02)[a,*]	2.97 (7.8)

Notes. [a]Correlations and standard error (in parentheses) reflect bootstrapped correlations of 1,000 samples with standard error weighted to account for repeated observations within individuals.
*$p < .001$.

Table 2. Comparison of trajectories in self-concept, self-efficacy, and quiz score

	Affect Coeff. (*SE*)[a]	Competence Coeff. (*SE*)[a]	Self-efficacy Coeff. (*SE*)[a]	Quiz Coeff. (*SE*)[a]
Fixed effects				
Intercept	49.52 (0.60)***	49.27 (0.57)***	50.58 (0.63)***	52.27 (0.46)***
Time	−0.04 (0.09)	−0.03 (0.08)	−0.25 (0.13)*	−0.10 (0.08)
Random effects				
Intercept	87.61 (9.22)	78.70 (8.47)	58.05 (10.47)	36.66 (5.48)
Time	0.87 (0.20)	0.34 (0.15)	0.92 (0.39)	0.05 (0.18)
ϵ (residual)	15.99 (0.94)	17.87 (1.03)	49.42 (2.74)	46.73 (1.89)

Note. [a]Robust MLE standard errors were adjusted for repeated observation.
*$p < .05$; ***$p < .001$.

Results are displayed in Table 2. As quiz assessment was based on different weekly topics, there was no reasonable expectation that scores would exhibit any systematic form of change. Indeed, non-significant slopes for time indicate that quiz scores were stable over time. Changes in the fixed effects for time for the self-concept factors (Affect and Competence) were not reported. A very small linear decline was noted in self-efficacy and the small magnitude of this effect perhaps should not be over-emphasized. However, this decline was examined with additional non-linear models which incorporated quadratic and cubic functions of time. Results indicated better fit for a model in which self-efficacy was modelled as a function of cubic (curvilinear) time (BIC = 7,631; AIC = 7,597) and quadratic (BIC = 7,692; AIC = 7,665) time, in comparison with a linear (BIC = 7,698; AIC = 7,676) model. Finally, a saturated means model was estimated in which time was expressed as a discrete function. Overall model fit was far superior for this saturated model (BIC = 7,623; AIC = 7,573) and suggests no particular systematic trend in changes in self-efficacy over the term. It was concluded that findings indicate no substantive or systematic growth in the either the self-referent attitudes or achievement scores and so no need to estimate a general growth function, such as through the estimation of latent slope factors to capture developmental-related changes.

It is important to note that the proportion of the random effects for time indicate little variance around the time slope in comparison with the variance reported at the intercept and residuals. The proportion of the variance reported at the intercept and the residual at each measurement occasion (ϵ), indicate that there is substantial variation at each measurement occasion and between individuals. Overall, the proportional size of random effects for self-efficacy and quiz was similar suggesting that differences reflect between-person differences, discrete moments of experience at each measurement occasion, and not any systematic developmental process. Most of the variance in the two self-concept factors 'Affect' and 'Competence' were reflected at the intercept and reflect substantial variation between individuals, which, taken with the fixed effects, suggests they are far more stable constructs.

Comparing the reciprocal effects model with CLPM and RI-CLPM frameworks

Model parameters for the CLPM and RI-CLPM models are reported in Table 3. Model fit indicators suggest the RI-CLPM is a better fitting model than the CLPM. A chi-square difference test revealed that the RI-CLPM was a better fitting model than the CLPM, $\chi^2_{\text{diff}} = 646.08$ ($\text{df}_{\text{diff}} = 10$), $p < .001$. Comparison of the model estimates between the two frameworks indicates substantial variation in model parameters. Autocorrelations across all measures were substantial and statistically significant in the CLPM, from $b = .42$ ($SE = .03$) for quiz score to $b = .83$ ($SE = .01$) for the Affect Self-Concept factor. However, in the RI-CLPM the substantive magnitude of these autoregressive parameters was attenuated substantially from .83 to .20 for Affect, from .79 to .15 for Competence, and from .61 to .21 for self-efficacy. The autoregression for Quiz was fully attenuated (no longer statistically different from zero) in the RI-CLPM. This indicates that most of the stability reported in the CLPM is between-person variance.

In terms of the main research focus, examining the cross-lagged parameters also indicated substantial differences between the CLPM and RI-CLPM. In the CLPM, Competence was the only significant driver of quiz score. Only a small effect for self-efficacy was reported. However, in the RI-CLPM, Competence was no longer significantly associated with quiz score, and instead, self-efficacy reported a substantial and statistically significant relationship with future change in quiz score. In terms of the cross-lagged effect of quiz score on future self-concept and self-efficacy, results were more comparable between models. Quiz performance was unrelated to the subsequent Affect factor in either the CLPM or RI-CLPM, but reported significantly substantive effects on Competence and self-efficacy. However, it is important to note that the effect of Quiz on Competence was lower in the RI-CLPM model whilst the effect on self-efficacy was higher.

The models reported in Table 3 were estimated in a way in which variables at one measurement occasion were regressed on variables at a preceding measurement occasion. This is standard within the self-concept *reciprocal effects model* framework (Arens *et al.*, 2017; Marsh, 1990; Marsh & Yeung, 1997; Pinxten, De Fraine, Van Damme, & D'Haenens, 2013; Pinxten *et al.*, 2014; Seaton *et al.*, 2014, 2015; Valentine & DuBois, 2005). However, a more accurate model, particularly with the current design, is one in which quiz score is regressed on self-concept and self-efficacy within each measurement occasion as the self-referent items preceded the quiz. Quiz score then predicts the self-referent factors at the subsequent measurement occasion (see Figure 3). Both CLPM and RI-CLPM were estimated with this design, and model estimates are reported in Table 4.

Re-estimating the model with this design did not substantially change model estimates. As with the earlier analyses, a range of model fit indicators suggest the RI-CLPM is a better

Table 3. Estimates from CLPM and RI-CLPM analyses of self-concept, self-efficacy, and achievement

	Cross-lagged panel model				Random intercept cross-lagged panel model			
	Quiz Coeff. (*SE*)	Affect Coeff. (*SE*)	Competence Coeff. (*SE*)	Efficacy Coeff. (*SE*)	Quiz Coeff. (*SE*)	Affect Coeff. (*SE*)	Competence Coeff. (*SE*)	Efficacy Coeff. (*SE*)
Quiz	0.42 (0.03)**	0.05 (0.02)	0.16 (0.02)**	0.14 (0.03)**	−0.01 (0.03)	0.02 (0.03)	0.09 (0.03)**	0.18 (0.05)**
Affect	−0.02 (0.03)	0.83 (0.01)**			0.05 (0.06)	0.20 (0.05)**		
Competence	0.13 (0.03)**		0.79 (0.02)**		−0.13 (0.07)		0.15 (0.05)*	
Efficacy	0.06 (0.03)*			0.61 (0.03)**	0.12 (0.04)**			0.21 (0.06)**
Model fit								
AIC	36,830				36,203			
BIC	37,106				36,519			
RMSEA (95% CI)	.08 (0.08; 0.09)**				.04 (0.03; .04)			
CFI	.79				.96			
χ^2 (df)	1,154.36 (362)**				508.28 (352)**			

Notes. AIC = Akaike information criteria; BIC = Bayesian information criteria; CFI = comparative factor index; RMSEA = root-mean-square error of approximation.
*p < .01; **p < .001.

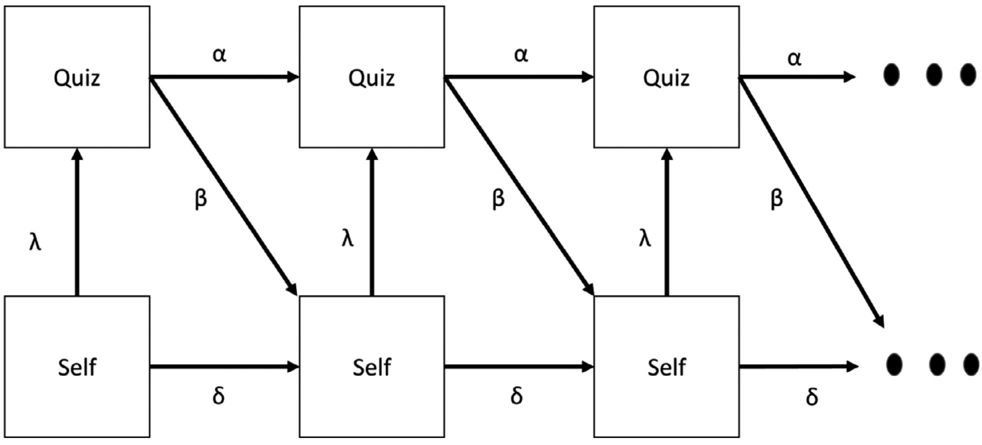

Figure 3. A re-estimation of the pathways between quiz score and self-referent factors within measurement occasion for both CLPM and RI-CLPM (RI-CLPM latent factors not drawn for clarity). *Note*. Quiz reflects achievement scores; Self reflects the self-referent variables (self-efficacy) and two self-concept factors (Competence and Affect) modelled simultaneously.

fitting model than the CLPM. A chi-square difference test revealed that the RI-CLPM was a better fitting model than the CLPM, $\chi^2_{\mathrm{diff}} = 633.64$ ($\mathrm{df}_{\mathrm{diff}} = 10$), $p < .001$. Autoregressive parameters across all measures in both the CLPM and RI-CLPM were comparable with the earlier analyses as they were estimated in a comparable way. Substantial effects were reported in the CLPM, from $b = .40$ ($SE = .03$) for quiz score to $b = .83$ ($SE = .01$) for the Affect Self-Concept factor, and in the RI-CLPM analysis, the substantive magnitude of these autoregressive parameters was again strongly attenuated from .83 to .20 for Affect, from .79 to .15 for Competence and from .61 to .21 for self-efficacy. The autoregression for Quiz was similarly fully attenuated in the RI-CLPM. The results of the cross-lagged parameters conformed with the initial CLPM and RI-CLPM and again reported substantial differences between the CLPM and RI-CLPM. In the CLPM, Competence and self-efficacy were drivers of quiz score, but in the RI-CLPM, only self-efficacy reported an association with quiz score. In terms of the cross-lagged effect of quiz score on future self-concept and self-efficacy, results were more comparable between models. Quiz performance was unrelated to subsequent Affect in either the CLPM or RI-CLPM, but reported significantly substantive effects on Competence and self-efficacy. Again, it is important to note that the effect of Quiz on Competence was strongly attenuated in the RI-CLPM model.

Discussion

The primary aim of this study was to compare the analysis of a *reciprocal effects model* of self-concept, self-efficacy, and achievement implemented within both CLPM and RI-CLPM frameworks. This comparison allows examination of the extent to which partialling between- and within-person variance in self-concept, self-efficacy, and academic achievement, in an RI-CLPM framework, accounts for the cross-lagged associations frequently reported in CLPM analyses. Results from this study showed that the RI-CLPM framework was a better model to fit to the data than the traditional CLPM analysis. The findings from the CLPM analyses in this paper are in line with prior evidence for a reciprocal relationship between self-concepts and achievement (Arens *et al.*, 2017; Guay

Table 4. Within-wave estimates from CLPM and RI-CLPM analyses of self-concept, self-efficacy, and achievement

	Cross-lagged panel model				Random intercept cross-lagged panel model			
	Quiz Coeff. (*SE*)	Affect Coeff. (*SE*)	Competence Coeff. (*SE*)	Efficacy Coeff. (*SE*)	Quiz Coeff. (*SE*)	Affect Coeff. (*SE*)	Competence Coeff. (*SE*)	Efficacy Coeff. (*SE*)
Quiz	0.40 (0.03)**	0.05 (0.02)	0.15 (0.02)**	0.13 (0.03)**	−0.01 (0.03)	0.01 (0.03)	0.10 (0.03)*	0.16 (0.05)**
Affect	−0.03 (0.02)	0.83 (0.01)**			−0.01 (0.06)	0.20 (0.05)**		
Competence	0.14 (0.03)**		0.80 (0.02)**		−0.08 (0.06)		0.15 (0.05)*	
Efficacy	0.08 (0.02)**			0.61 (0.03)**	0.13 (0.03)**			0.20 (0.06)**
Model fit								
AIC	36,821				36,207			
BIC	37,086				36,511			
RMSEA (95% CI)	.08 (0.07; 0.08)**				.04 (0.03; 0.04)			
CFI	.80				.96			
χ^2 (df)	1,151.78 (365)**				518.14 (355)**			

Notes. AIC = Akaike information criteria; BIC = Bayesian information criteria; CFI = comparative factor index; RMSEA = root-mean-square error of approximation.
*$p < .01$; **$p < .001$.

et al., 2003; Marsh, 1990; Marsh & Yeung, 1997; Pietsch *et al.*, 2003; Pinxten *et al.*, 2013; Seaton *et al.*, 2014, 2015; Valentine & DuBois, 2005). However, by decomposing construct variance into between- and within-person components within an RI-CLPM framework, the results do not support this relationship between self-concept and achievement, although a reciprocal relationship between self-esteem and achievement is reported. Indeed, the findings of the current study indicate that there appears to be a reciprocal relationship between self-efficacy and achievement. The relationship between self-concept and achievement would be best described by a *skill development model*. That is, self-concept, particularly the first-order Competence factor, is strongly driven by prior achievement, but future changes in achievement are not strongly related to prior self-concept. This suggests that between-person differences in self-concept are what drive many purported effects in prior CLPM analyses, including those CLPM analyses implemented in this study.

It is important to recognize that cross-lagged estimates from CLPM and RI-CLPM are not directly comparable. It must be emphasized that the major difference between the CLPM and RI-CLPM analyses is that with the inclusion of latent factors which capture the between-person variance in the RI-CLPM framework, the cross-lagged regressions now reflect whether changes from an individual's expected score on one variable are predicted from preceding deviations on a second variable, that is, a marker of within-person change. Simply, within the CLPM framework, the cross-lagged estimates comprise both between and within-person variation, whilst in the RI-CLPM, they reflect the average within-person change relative to individuals' estimated average level. Since self-concept and self-efficacy relate specifically to individual level attitudes of self, it is sensible to use a RI-CLPM framework to assess how individual level characteristics predict individual-level behaviour (i.e., academic achievement). Clearly, the CLPM framework fails to appropriately account for the within-person change in constructs over time.

Previous CLPM analyses have frequently unconstrained cross-lagged and autoregressive parameters. Consequently, inconsistent findings are often reported and describe how reciprocal effects may be apparent between some measurement points whilst a self-enhancement model is supported at other measurement points in the same model. As an example, Guay *et al.* (2003) conclude that 'results supported a *reciprocal effects model* for the first two waves of data collection and a self-enhancement model between the second and third waves' (p. 130). Such patterns of attenuation are reported in many studies of the *reciprocal effects model* (Marsh, 1990; Pinxten *et al.*, 2013, 2014). Consequently, perhaps more reflective of the multi-level or generalized estimating equations frameworks, in the above analyses, these parameters have been constrained as time-invariant. There are strengths and limitations of both methods. First, constrained models are easier to estimate, and secondly, they allow us to derive an average of the estimates over the study period. As there was no linear change in our constructs, it seems sensible to derive averages over the study period. Given there is no 'developmental' trajectory in those constructs estimated here, this is a sensible and defensible position. However, further examination of a model in which the coefficients are unconstrained might well focus on how the size of the structural relationships changes over time; this may be particularly informative to instructors who might determine which assessment of specific course content may have the strongest adverse impact on students' self-concept, or conversely how student self-concept impacts on achievement tasks related to specific course content.

It is important to emphasize some caveats to the findings of this research for generalization to other populations. First, the participants comprised a group of students enrolled in an introductory psychology unit. Not all students were enrolled in a psychology degree or major; some students were enrolled in nursing and physiotherapy degrees for whom introductory units are compulsory. Yet other students are enrolled in other degrees and undertook an introductory unit for other reasons. It is important to consider that enrolment in the unit may moderate these findings. Much of the previous literature on self-concept, self-efficacy, and achievement has focused on primary and secondary-aged pupils. Additionally, it is important to note that much of the prior research has focused on long-term changes in self-concept and achievement of between 6, 12, and 24 months. Whilst these issues may limit comparisons with prior studies, the CLPM implemented in the current study confirms a reciprocal effects relationship and suggests that these limitations have not adversely impacted on our findings.

In conclusion, RI-CLPM presents a sound and robust alternative to the CLPM approaches frequently utilized in much educational psychology research. Analyses of the reciprocal effects of self-concept, self-efficacy, and achievement within a CLPM framework may be informative, to an extent, but it is important to understand their limitations. Analyses within an RI-CLPM may provide additional and more nuanced understanding of the interrelationships between self-referent attitudes and achievement outcomes. It is important to recognize that no one single analytical framework can answer all possible questions in relation to longitudinal changes in constructs of interest, nor their inter relationships over time. But it is imperative that research fully appreciates the limitations of the analytical method they have utilized and make appropriate conclusions about the implications of their findings in the light of these limitations. At least, a more thorough and thoughtful examination of the relationship between self-concept, self-efficacy, and achievement is needed to assist professionals in a variety of fields to develop policies and programmes that will facilitate the enhancement of both self-attitudes and achievement.

References

Abu-Hilal, M. M., Abdelfattah, F. A., Alshumrani, S. A., Abduljabbar, A. S., & Marsh, H. W. (2013). Construct validity of self-concept in TIMSS's student background questionnaire: A test of separation and conflation of cognitive and affective dimensions of self-concept among Saudi eighth graders. *European Journal of Psychology of Education*, *28*, 1201–1220. https://doi.org/10.1007/s10212-012-0162-1

Arens, A. K., & Hasselhorn, M. (2015). Differentiation of competence and affect self-perceptions in elementary school students: Extending empirical evidence. *European Journal of Psychology of Education*, *30*, 405–419. https://doi.org/10.1007/s10212-015-0247-8

Arens, A. K., Marsh, H. W., Craven, R. G., Yeung, A. S., Randhawa, E., & Hasselhorn, M. (2016). Math self-concept in preschool children: Structure, achievement relations, and generalizability across gender. *Early Childhood Research Quarterly*, *36*, 391–403. https://doi.org/10.1016/j.ecresq.2015.12.024

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*, 621–634. https://doi.org/10.1037/edu0000163

Arens, A., Yeung, A., Craven, R., & Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology*, *103*, 970–981. https://doi.org/10.1037/a0025047

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. C. Urdan & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents*. Charlotte, NC: Information Age.

Becht, A. I., Nelemans, S. A., Dijk, M. P. A., Branje, S. J. T., Van Lier, P. A. C., Denissen, J. J. A., & Meeus, W. H. J. (2017). Clear self, better relationships: Adolescents' self-concept clarity and relationship quality with parents and peers across 5 years. *Child Development*, *88*, 1823–1833. https://doi.org/10.1111/cdev.12921

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Burns, R. A., Crisp, D. A., & Burns, R. B. (2018). Competence and affect dimensions of self-concept among higher education students: A factorial validation study of an academic subject-specific self-concept. *European Journal of Psychology of Education*, *33*, 649–663. https://doi.org/10.1007/s10212-018-0369-x

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

Gardner, D. G., Cummings, L. L., Dunham, R. B., & Pierce, J. L. (1998). Single-item versus multiple-item measurement scales: An empirical comparison. *Educational and Psychological Measurement*, *58*, 898–915. https://doi.org/10.1177/0013164498058006003

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. https://doi.org/10.1037/0022-0663.95.1.124

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

Hoeppner, B. B., Kelly, J. F., Urbanoski, K. A., & Slaymaker, V. (2011). Comparative utility of a single-item versus multiple-item measure of self-efficacy in predicting relapse among young adults. *Journal of Substance Abuse Treatment*, *41*, 305–312. https://doi.org/10.1016/j.jsat.2011.04.005

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

Kline, R. B. (1998). *Principles and practice of structural equation modeling. Methodology in the social sciences*. New York, NY: Guilford Press.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic-achievement – A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, *82*, 646–656. https://doi.org/10.1037/0022-0663.82.4.646

Marsh, H. W., Craven, R., & Debus, R. (1999). Separation of competency and affect components of multiple dimensions of academic self-concept: A developmental perspective. *Merrill-Palmer Quarterly*, *45*, 567–601.

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*, 181–220. https://doi.org/10.1207/s15327906mbr3302_1

Marsh, H. W., Muthen, B., Asparouhov, T., Ludtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*, 439–476. https://doi.org/10.1080/10705510903008220

Marsh, H. W., & O'Neill, R. (1984). Self description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, *21*(2), 153–174. https://doi.org/10.1111/j.1745-3984.1984.tb00227.x

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology*, *89*, 41–54. https://doi.org/10.1037/0022-0663.89.1.41

Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, *95*, 589–603. https://doi.org/10.1037/0022-0663.95.3.589

Pinxten, M., De Fraine, B., Van Damme, J., & D'Haenens, E. (2013). Student achievement and academic self-concept among secondary students in Flanders: Gender and changes over time. *Irish Educational Studies*, *32*, 157–178. https://doi.org/10.1080/03323315.2012.749058

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, *84*, 152–174. https://doi.org/10.1111/bjep.12028

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258. https://doi.org/10.1037/0033-2909.88.2.245

Seaton, M., Marsh, H. W., Parker, P. D., Craven, R. G., & Yeung, A. S. (2015). The reciprocal effects model revisited: Extending its reach to gifted students attending academically selective schools. *Gifted Child Quarterly*, *59*, 143–156. https://doi.org/10.1177/0016986215583870

Seaton, M., Parker, P., Marsh, H. W., Craven, R. G., & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: Juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*, *34*, 49–72. https://doi.org/10.1080/01443410.2013.825232

Valentine, J. C., & DuBois, D. L. (2005). Effects of self-beliefs on academic achievement and vice versa. Separating the chicken from the egg. In H. Marsh, R. G. Craven & D. M. McInerney (Eds.), *International advances in self research: New frontiers for self research*. Greenwich, UK: Information Age.

Vangeel, L., Vandenbosch, L., & Eggermont, S. (2018). The multidimensional self-objectification process from adolescence to emerging adulthood. *Body Image*, *26*, 60–69. https://doi.org/10.1016/j.bodyim.2018.05.005

Yang, L., Arens, A. K., & Watkins, D. A. (2014). Testing the twofold multidimensionality of academic self-concept: A study with Chinese vocational students. *Educational Psychology*, *36*, 1651–1669. https://doi.org/10.1080/01443410.2014.995597

Received 4 September 2018; revised version received 19 November 2018