Journal of School Psychology 103 (2024) 101283

Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

Third variables in longitudinal research: Application of longitudinal mediation and moderation in school psychology

Jacqueline M. Caemmerer [a,*], Briana Hennessy [a], Christopher R. Niileksela [b]

[a] University of Connecticut, Neag School of Education, 249 Glenbrook Road, Storrs, CT 06269, USA
[b] University of Kansas, JRP Hall Room 640, 1122 West Campus Road, Lawrence, KS 66045, USA

ARTICLE INFO

Editor: Craig A. Albers
Action Editor: Danny Hajovsky

Keywords:
Longitudinal
Moderation
Mediation
Indirect effect
Interaction
Moderated mediation
Mediated moderation

ABSTRACT

Third variable models, such as mediation and moderation, can identify contextual factors that help explain the relation between two variables. Although used less frequently in school psychology research, longitudinal mediation, longitudinal moderation, and the integration of these two approaches can be used to describe the developmental changes in children's psychological and behavioral processes throughout the school years and beyond. This article provides conceptual descriptions of longitudinal mediation, longitudinal moderation, longitudinal moderated mediation, and longitudinal mediated moderation and demonstrates the use of these methods with a large sample of elementary students. Extensions of these methods and applied examples from the literature are also discussed. The M*plus* syntax from our illustrated examples are provided for those interested in reproducing the analyses.

1. Introduction

Many fundamental research questions in school psychology and related fields focus on the relations among different variables. Although two variables may be related, researchers are often interested in finding out why that relation exists and whether other variables may explain or influence that relation. For example, one hypothesis may be that individual differences in one variable cause individual differences in another variable. However, a different hypothesis may suggest that individual differences in a third variable may explain *why* or *under which circumstances* two variables are related to each other. Understanding how other variables affect the relations among the primary variables of interest is an important part of developing a strong understanding of how and why constructs are related to each other. Likewise, having a strong sense of moderation and mediation is essential for developing a comprehensive understanding of psychological and educational constructs and outcomes.

Moderation and mediation are often referred to as third variable models. *Mediators* are intervening variables that either partially or completely explain the relation between an independent and dependent variable. Mediators explain how or why two variables are related. In regression, the terms mediation and indirect effects are used synonymously. In contrast, *moderators* are variables that explain for whom a relation between two variables applies; in other words, moderation examines the generalizability of the relation between two variables (Fairchild & McQuillin, 2010). In moderation, the magnitude of the effect of the independent variable on the dependent variable is conditional, or depends, on the level of the moderator. In regression, the terms moderation and interactions are

* Corresponding author.
E-mail addresses: jacqueline.caemmerer@uconn.edu (J.M. Caemmerer), briana.hennessy@uconn.edu (B. Hennessy), chrisn@ku.edu (C.R. Niileksela).

https://doi.org/10.1016/j.jsp.2024.101283
Received 29 July 2022; Received in revised form 12 July 2023; Accepted 4 January 2024
Available online 20 January 2024

used synonymously.

Both mediation and moderation can identify contextual factors that influence the relations between variables. The study of mediators and moderators are an indicator of a maturing profession (Judd et al., 1995) that advances both theory and clinical practice (Holmbeck, 1997; Rose et al., 2004). In 2010, Fairchild and McQuillin published a review article in the *Journal of School Psychology* (*JSP*) describing mediation and moderation and how the approaches were applied at that time in the field of school psychology. Since their publication, additional methodological advances became more widely available that are worthy of exploration. Such advances include longitudinal mediation, longitudinal moderation, and the integration of these two concepts through longitudinal moderated mediation and mediated moderation.

Longitudinal research involves repeated measurement of constructs and examines the "time-ordered study of processes" (Baltes & Nesselroade, 1979, p.2). Repeated measurement requires a minimum of two measurement occasions. In longitudinal mediation and moderation, the predictor, mediator or moderator, and outcome are often measured at different time points (Fairchild & McQuillin, 2010). Longitudinal moderation, mediation, and integrated approaches characterize the development of psychological and behavioral processes in school children and can offer theoretical insights to inform intervention and instruction. Although these longitudinal and integrated approaches have been discussed by methodologists for quite some time (e.g., Baron & Kenny, 1986) their use has not been as widespread in the field of school psychology. For example, we searched within the *JSP,* our intended audience, for articles related to longitudinal moderation, longitudinal mediation, mediated moderation, or moderated mediation to demonstrate the limited use of these approaches in the field of school psychology. Our search, which was open to all publication years, revealed approximately 20 articles that examined longitudinal mediation or longitudinal moderation and approximately five articles that tested either mediated moderation or moderated mediation. This finding supports the need for descriptions and demonstrations of how longitudinal mediation, longitudinal moderation, longitudinal moderated mediation, and mediated moderation can be applied to school psychology research.

Our article seeks to extend Fairchild and McQuillin's (2010) article by describing these longitudinal methodological advances and demonstrating their use with a large sample of elementary students. The purpose of our article is to describe and provide applied examples from the literature (a) of longitudinal mediation, (b) of longitudinal moderation, (c) of integrated approaches, and (d) to illustrate the use of longitudinal moderation, longitudinal mediation, longitudinal moderated mediation, and longitudinal mediated moderation to further elaborate on these advances. This article is a conceptual overview of these topics. References are provided throughout for readers interested in technical aspects of these methods.

1.1. Conceptual overview of mediation

Prior to an in-depth discussion of longitudinal mediation and moderation we present a general overview of both. A mediator is an intermediary or intervening variable (Z) between a predictor or independent variable (X) and outcome or dependent variable (Y) that may explain how or why those two variables are related. Fig. 1 shows a traditional example of mediation using a single measurement occasion. Mediation is often examined in intervention research to identify the mechanisms through which interventions exert their effects. In other types of research mediation explains how a predictor influences change in an outcome; mediation is also popular in prevention research (Preacher et al., 2007).

Mediators can be continuous or categorical variables. For example, researchers may be interested in testing whether high school grade point average (GPA; i.e., X) predicts a person's income as an adult (i.e., Y). It is possible that years of postsecondary education (i.e., Z) mediates the relation between high school GPA and income, where those who have higher GPAs tend to obtain post-secondary degrees, which in turn results in a higher income. Using the same example, as a categorical mediator the type of post-secondary degree obtained may mediate the relation between high school GPA and income, where people with stronger high school GPAs obtain medical

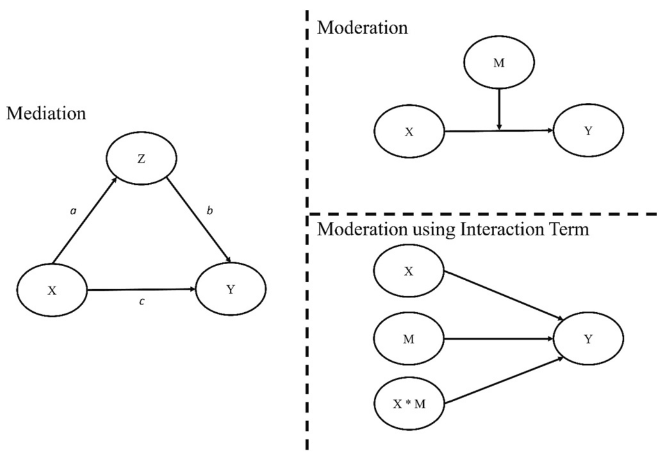

Fig. 1. Examples of Mediation and Moderation Measured at Single Time Points.
Note. X = predictor; Y = outcome; Z = mediator; M = moderator.

degrees, which in turn results in a higher income.

Mediation allows researchers to decompose the effects between the predictor and outcome into (a) a direct effect between the predictor and outcome and (b) an indirect effect from the predictor through the mediator to the outcome (Fairchild & McQuillin, 2010). The direct effect between the predictor and outcome is labeled *c* (see Fig. 1). The indirect effect is the composition of the paths from the predictor to the mediator (labeled *a)* and from the mediator to the outcome (labeled *b*; Baron & Kenny, 1986). Full or complete mediation is present when the direct effect (*c*) between the predictor and outcome is no longer statistically significant after the mediator is added to the model. Partial mediation is present when the direct effect is reduced in magnitude but remains statistically significant (Selig & Preacher, 2009).

There are a few approaches to test mediation, and one such approach evaluates the causal steps of mediation. Baron and Kenny's (1986) approach, also known as the causal steps approach, includes the following four components: (a) the predictor has a statistically significant relation with the outcome, (b) the predictor has a statistically significant relation with the mediator, (c) the mediator has a statistically significant relation with the outcome, and (d) the direct effect between the predictor and outcome is reduced when the mediator is included in the model. The extent to which (d) is true indicates the extent of the mediation (Baron & Kenny, 1986). Although the first component (i.e., the statistically significant relation between the predictor and outcome) was traditionally described as a prerequisite of mediation (Baron & Kenny, 1986), some experts have argued the indirect/mediation effect can be relevant without a significant relation between the predictor and outcome (MacKinnon et al., 2002; Shrout & Bolger, 2002; Wu & Zumbo, 2008). Additionally, the first prerequisite may inadvertently exclude mediation effects in which the direct and indirect effects have opposite signs and cancel each other out (MacKinnon et al., 2000). Finally, there are also concerns about statistical power with this four-step approach because it requires multiple statistical tests.

Baron and Kenny (1986) recommended the Sobel method to test the statistical significance of the indirect effect. The Sobel test is also known as the product-of-coefficients approach (Preacher & Hayes, 2008). Separate regressions estimate the paths from the predictor to the mediator (*a*) and from the mediator to the outcome (*b*). The indirect effect is calculated as the product of *ab,* with a standard error based on the standard errors of *a* and *b* (Sobel, 1982). Sobel's test assumes a normal distribution, which is problematic because the distribution of *ab* is only normal in large sample sizes (Preacher & Hayes, 2008; Shrout & Bolger, 2002). Those interested in testing this approach can try Kristopher Preacher and Geoffrey Leonardelli's interactive calculation tool (http://quantpsy.org/sobel/sobel.htm).

Bootstrapping, another method to test indirect effects, is currently the preferred approach because it overcomes the limitations of the other two methods. Bootstrapping is less dependent on sample size and is a nonparametric procedure that does not require any assumptions about the sampling distribution (Preacher & Hayes, 2008). Many random samples from the data are drawn and replaced to compute the indirect/mediated effect. These resampling methods are repeated a prespecified number of times, often at least 1000 times (Preacher et al., 2007). Estimates of the indirect effect are calculated for each of the bootstrapped samples and a sampling distribution of the indirect effect is created (Preacher et al., 2007). Percentile-based confidence intervals are generated from the frequency distribution. If zero is outside the confidence interval the mediated effect is statistically significant. Computation of the confidence intervals can be time intensive even with the use of statistical programs. Slightly different confidence intervals will be formed each time the procedure is used with the data due to the different resamples used, but this variation decreases as the number of resampling draws increases (Preacher et al., 2007).

An indirect effect is the product of two effects ($a*b$), which complicates the calculation of the effect size of the indirect effect. Just as there are multiple approaches to test the statistical significance of an indirect effect, there are multiple approaches to determine the effect size of the indirect effect. Preacher and Kelley (2011) recommended selecting an effect size unaffected by sample size and scaled on a meaningful metric and reporting confidence intervals around the effect size. Preacher and Kelley also recommended researchers use the ratio of the obtained indirect effect to the maximum possible indirect effect (k^2) as an effect size measure because k^2 meets these criteria. Researchers are encouraged to report multiple effect size measures in a single study because there may be instances where unstandardized indirect effects, completely standardized indirect effects, or proportion metrics are useful (Preacher & Kelley, 2011). Each effect size measure has limitations. For example, these effect size measures only apply to simple mediation models with a single mediator (Preacher & Kelley, 2011).

Mediation models often test a single mediator, but a mediation model can also test multiple mediators simultaneously. Such models are referred to as *multiple mediator* or *complex mediation* models (Fairchild & McQuillin, 2010; VanderWeele, 2016). Multiple mediators are a likely scenario in the field of psychology because psychological phenomena are often explained by multiple causes (in the earlier example, there are likely several variables that mediate the relation between high school GPA and later income), thus complete mediation by a single mediator is rare (Baron & Kenny, 1986). Multiple mediators can be measured at the same time or cross-sectionally, or the mediators may be measured in a temporal sequence or longitudinally (Fairchild & McQuillin, 2010). In models with multiple mediators an individual indirect effect is calculated for each mediator, therefore several indirect effects can be examined. These individual paths are sometimes referred to as specific indirect pathways. Specific indirect pathways may or may not be of interest depending on the research hypotheses being tested, which depends on whether there is a rationale to examine a particular indirect effect at a specific time interval (Cole & Maxwell, 2003; Little, 2013). Additionally, complex mediation models estimate a total indirect effect, which is the sum of all the individual specific indirect effects. Often, the total indirect effect is of great interest because it provides a broader estimate of the mediation effect across the full model and there may not be strong hypotheses about individual specific indirect effects. The total indirect effect of multiple temporally sequenced mediators is relevant in longitudinal models, which are described in a later section.

1.2. Conceptual overview of moderation

Moderation is examined when the strength or direction of the relation between a predictor or independent variable (X) on an outcome or dependent variable (Y) is influenced by one or more moderators (M). In other words, the relationship between a predictor and outcome variable depends on the value of a third variable, or moderator. In essence, moderators divide the independent variable (X) into subgroups to determine zones of "maximal effectiveness" for a dependent variable (Y; Baron & Kenny, 1986). This focus on subgroups means moderation can address questions about the generalizability of research findings because moderation shows if there are differences in the outcome for subgroups of individuals in the sample. Moderation can also be a potential explanation for a weak or non-significant finding when there is in fact a true effect that does not generalize to all subgroups. This can occur because either there is an effect for only a small portion or subgroup of the sample, or the effect may be positive for some individuals and negative for others, thereby canceling out the overall effect (Wu & Zumbo, 2008). The right side of Fig. 1 shows traditional examples of moderation using a single measurement occasion. In the top panel, the moderator M influences the size of the path between X and Y. Moderation is typically estimated in regression using the example in the bottom panel, where the main effects of X and M and the interaction between X and M (X*M) directly predict Y.

Like mediators, moderators can be continuous or categorical. The data type determines how moderation is tested. Continuous moderators are examined through interaction terms. Continuing with the previous example, there may be a relation between high school GPA and income. Age may be a continuous moderator that affects the relation between high school GPA and income because people who are older may have been in their jobs longer, resulting in a higher income than those who are younger, even though their high school GPAs may be the same. In regression, the main effects of the predictor and moderator variables (X and M) on the outcome are examined first, and then the effect of the interaction or cross-product between the predictor and moderator variables (X*M) on the outcome are examined second. Variables in the cross-product are typically mean-centered, where the sample mean is subtracted from all scores on both variables. Mean-centering does not affect the strength of the relation between the two variables, but it is used to enhance interpretation and to reduce the potential of multicollinearity (some experts disagree about the latter reason; see McClelland et al., 2017). If the interaction term has a statistically significant effect on the outcome, then it suggests moderation is present. The significant interaction means there are different slopes for different levels of the continuous moderator or for different groups of the categorical moderator ("different slopes for different folks"; Keith, 2019, p.178). The inclusion of the main effects of the predictor and moderator variables in the regression is critical even if the effects are not statistically significant. Their inclusion avoids the artificial inflation of the interaction by partialling them out of the interaction term (Aiken & West, 1991; Fairchild & McQuillin, 2010). Although linear interactions are often examined in moderation analyses, non-linear functional forms can also be tested, including quadratic or exponential forms (Selig et al., 2012).

Although less common, researchers can test interactions with latent, unobserved variables. Latent interactions are estimated without measurement error, which can result in more accurate estimates. Modern software, such as M*plus*, can test latent interactions and does not require centering of the observed indicators. This is demonstrated in our analyses below. Unfortunately, M*plus* does not report most fit indices, standardized coefficients, or effect sizes for interactions with latent variables, but researchers can compute this information themselves using the methods described in Maslowsky et al. (2015).

Categorical moderators can be tested with regression approaches if binary variables are coded dichotomously or categorical variables with multiple categories are coded as dummy variables. Continuing with the previous example, gender may be a binary categorical moderator. Gender may moderate the relationship between high school GPA and income because women have historically earned less money than men, so the relationship between high school GPA and income may be different for men and women. With categorical variables, moderation can also be tested through a multi-group model using structural equation modeling (SEM). In this approach model parameters (e.g., factor loadings, intercepts, structural paths) can be compared across groups to determine if they are statistically significantly different from each other. This approach is referred to as factorial or measurement invariance, which is described further in the next section. In SEM and confirmatory factor analysis (CFA), the examination of measurement invariance across groups is a form of moderation. These multigroup comparisons determine whether constructs are comparable across groups, which must be established before other between group differences are tested (Little, 1997). Measurement invariance addresses validity issues related to the internal structure of a measure and tests the extent of measurement equivalence across diverse groups (see Pendergast et al., 2017, for a detailed example). For instance, if the size of the factor loadings or intercepts for a depression scale differs across boys and girls, the assumption of measurement invariance would not be supported. The relation between the latent variable of depression (X) and its observed indicators (Y), which may be items or subscales, are moderated by gender (M). Although measurement invariance is often tested for two groups, it is possible to test measurement invariance for several groups

1.3. Longitudinal mediation and moderation

The general concepts for mediation and moderation can be extended longitudinally. Fig. 2 shows how mediation may be applied in a longitudinal model, and Fig. 3 and Fig. 4 show how moderation may be applied in a longitudinal model. Fig. 3 shows a traditional moderation diagram. Fig. 4 shows how moderation is displayed using a multi-group SEM model. The general concepts are the same as the single measurement occasion models described above, extended to examine multiple time points and how mediators and moderators may affect longitudinal relationships among variables. Although there are other configurations of paths possible (e.g., Z may be regressed on X during the same timepoint, then both predict Y at the subsequent timepoint), decisions about how longitudinal paths are included in the model should be based on theoretical and conceptual understanding of variables and how they relate to each other.

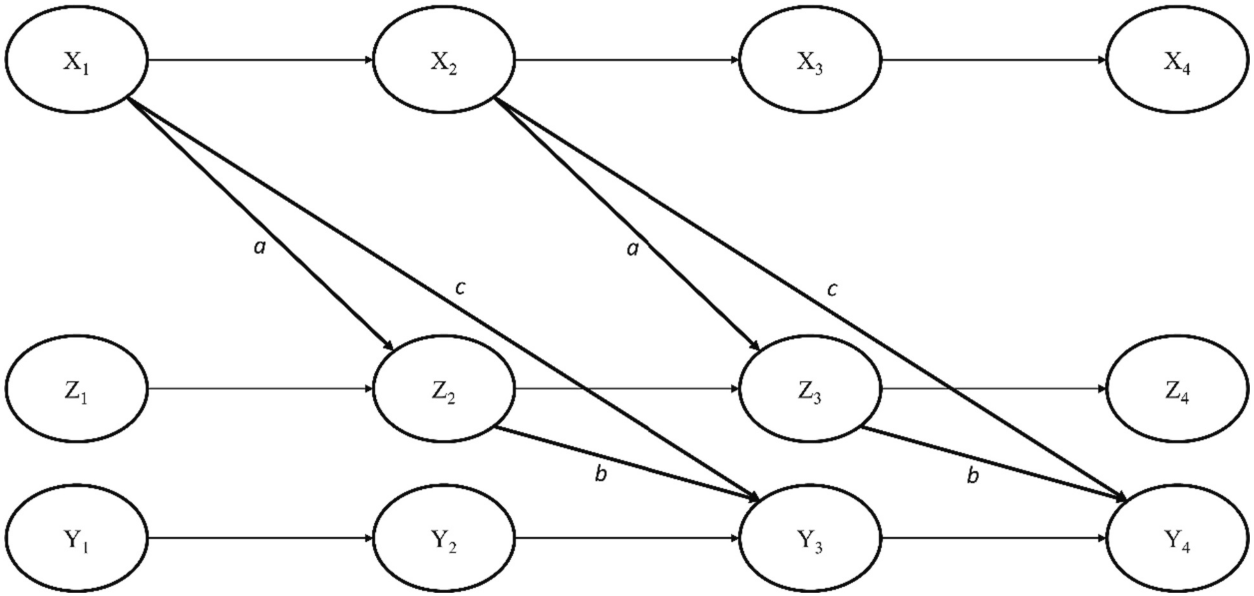

Fig. 2. Example of Longitudinal Mediation.

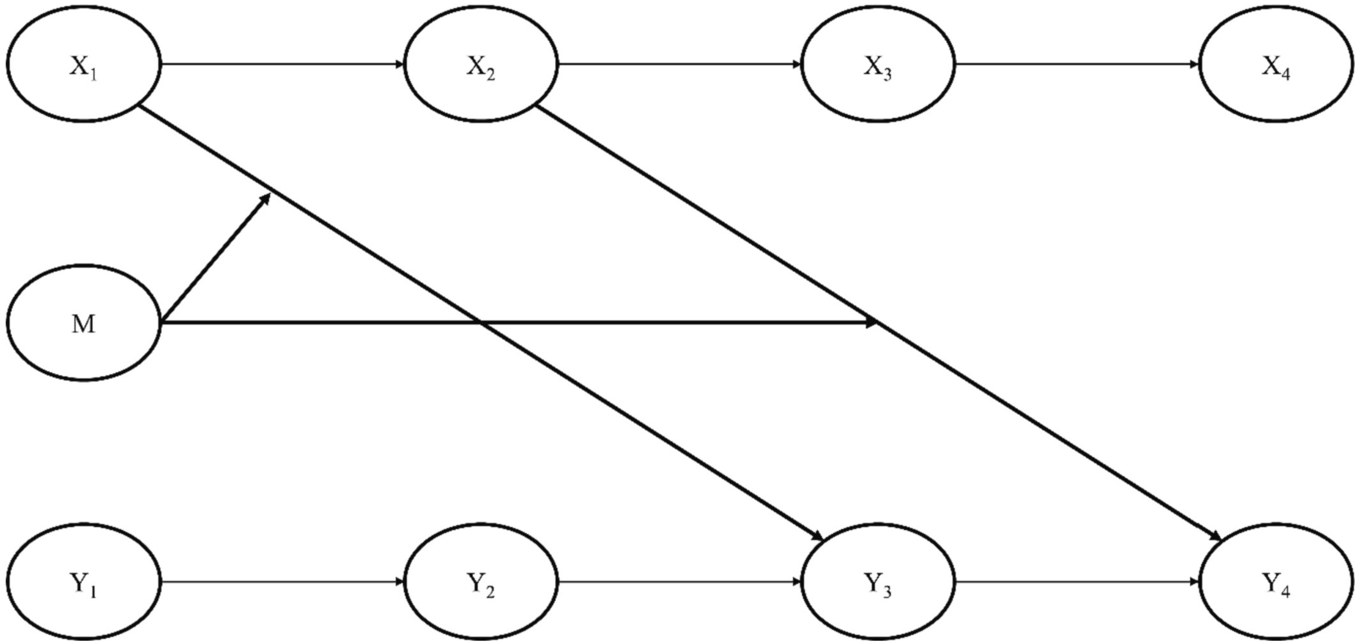

Fig. 3. Example of Longitudinal Moderation.

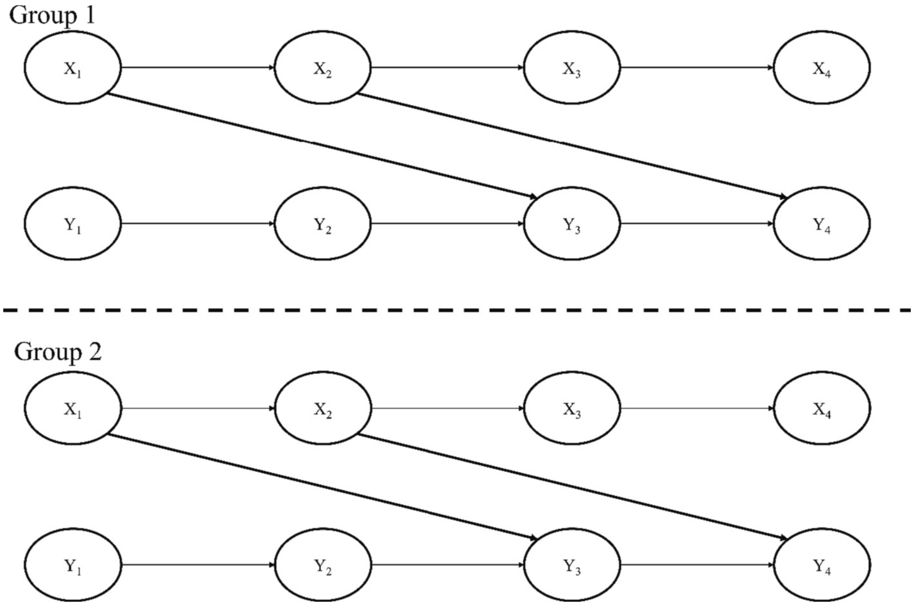

Fig. 4. Example of Longitudinal Moderation Using Multigroup SEM.

1.3.1. Model considerations: measurement invariance and inclusion of covariates

The multiple measurement occasions in longitudinal data introduce important issues researchers should consider before they can test their substantiative mediation and moderation research questions. One such consideration is the prerequisite of measurement invariance. Longitudinal analysis assumes the repeated measurements of each construct are measured similarly over time, but this assumption must be examined through tests of factorial invariance (Meredith, 1993; Vandenberg & Lance, 2000; Widaman et al., 2010), which provide necessary evidence of longitudinal construct equivalence. Equivalence of the constructs across time must be tested before the predictive relations between the constructs can be examined (Little, 2013; Widaman et al., 2010).

Factorial invariance tests require multiple indicators of each latent or unobserved variable. Invariance tests are completed in a series of steps using a multi-group confirmatory factor analysis (CFA; Little, 1997; Meredith, 1993; Widaman et al., 2010). Increasingly stringent constraints are placed on the measurement model parameters; the measurement model includes the factor loadings from the latent variables to the measured or observed variables and the observed variable intercepts. Weak or metric invariance, or the assumption of equal factor loadings, is required if the relations among latent variables are to be compared over time. Strong or intercept invariance, or the assumption of equal intercepts, is required for mean comparisons of scores over time (Keith, 2019). With a categorical variable strong invariance is known as threshold invariance and may require alternative approaches (Svetina et al., 2020). If invariance is supported across the time points in the study, the measurement constraints will be incorporated into the structural model; the structural model includes the paths and covariances between the latent variables. For longitudinal mediation and moderation using SEM, the presence of weak invariance is required because these approaches are primarily designed to examine relations among variables rather than mean differences across groups.[1] A full review of factorial invariance is outside the scope of this article, but interested readers are referred to the cited invariance references for a more in-depth discussion.

Another consideration in longitudinal research is how to include the influence of covariate variables when there are three or more time points. Covariates are also referred to as control, confounding, or instrumental variables. Covariates are included in models to account for mean-level differences of the covariate and to estimate the relations more accurately between the predictor and outcome variables (Little, 2013). Theory should guide the selection of covariates. In longitudinal models, researchers should consider covariates that explain the context of the developmental processes under study. The inclusion of too many covariates, which may not be supported by theory, may lead to overcontol, which is when covariates are related to the constructs of interest due to random chance alone. This random variation is removed from the constructs of interest and can reduce, or weaken, the effects of interest (Little, 2013).

Some covariates are referred to as time-invariant because the covariate itself does not change throughout the period studied, such

[1] The large number of constraints involved in latent variable longitudinal models can result in statistically significant differences in factor loadings. Little (1997) distinction between a "modeling" rationale and a "statistical" rationale for acceptance of a claim (e.g, invariance between groups) explains that a model may be selected for good overall fit (modeling rationale) or because differences between the chosen model and a baseline model are statistically significant (statistical rationale). Using the modeling rationale to support measurement invariance is useful because it only requires that the invariant model fit well as a whole, even if the $\Delta\chi^2$ is statistically significant (a likely outcome with a large sample size and many parameter restrictions). The statistical rationale is important for substantive claims. These conclusions are based on theoretically important differences and require a statistical test for each conclusion.

as an individual's ethnicity. Other covariates are referred to as time-variant because the covariate does change throughout the study, such as the amount of intervention time a youth receives (Little, 2013). With time-variant covariates, measurements from multiple time points are typically included in the model, whereas time-invariant covariates are only included once. Covariates can be included in longitudinal models in several ways (Little, 2013). One approach includes paths from the covariates to all the variables at each time point. This approach fully controls for the influence of the covariates on all the predictors and outcomes. If the covariates and predictors are all measured at the first time point, another approach is to model all Time 1 variables together as exogeneous (independent) variables and include paths from the covariates to Time 2 and Time 3 variables. Another approach is to only include paths from the covariates to the first time point. The influence of the covariates on later time points is accounted for indirectly because variables at the first time point predict variables at the second time point and so on. This assumes the influence of the covariate on later time points lessens with the passage of time (Little, 2013). This assumption can be tested by examining whether paths from the covariates to variables measured at later time points are statistically significant. Little recommended pruning nonsignificant covariate effects from the model to avoid overcontrol of the nonsignificant parameters "soaking up variance that may in fact be random perturbations" (p. 195), but not all methodologists would agree with such pruning. Theory and time precedence should be considered when determining how and when covariates will be included (see Little, 2013, pp. 15–17, 194–198, for further discussion of covariates in longitudinal models).

1.3.2. Temporal precedence and sequencing

Although both longitudinal moderation and mediation can be examined, currently longitudinal mediation appears to be more commonly used in studies. In fact, in Fairchild and McQuillin's (2010) article, only longitudinal mediation was briefly described, whereas a discussion of longitudinal moderation was absent. Therefore, many of the references in this section focus solely on longitudinal mediation because it is more commonly used, but many of the same principles can be applied to longitudinal moderation.

Although mediation has frequently been examined with cross-sectional data, some experts have argued longitudinal data are better suited for mediation tests. This assertion is based on the logic that the passage of time is required for the effects of mediation to occur (Cole & Maxwell, 2003; Preacher, 2015; Selig & Preacher, 2009). In cross-sectional models the mediator is often measured concurrently with the predictor and outcome. This timing assumes the effects in the mediation model occur immediately (Cole & Maxwell, 2003). This assumption is fallacious because mediation models often examine developmental processes that logically require the passage of time to unfold (Maxwell et al., 2011; Selig & Preacher, 2009). In longitudinal mediation models at least two time points are needed, but often three time points are examined. Three time points allow the predictor, mediator, and outcome to be temporally sequenced at different time points (Fairchild & McQuillin, 2010; Maxwell et al., 2011). The inclusion of time and temporal sequencing strengthens the causal inferences examined in longitudinal mediation (Cole & Maxwell, 2003; MacKinnon et al., 2002; Preacher, 2015). Beyond this logical reasoning, cross-sectional estimates of mediation have been found to be biased. Cross-sectional estimates may either under- or overestimate the longitudinal mediation effect (Maxwell et al., 2011). Additionally, cross-sectional studies may find a mediation effect when longitudinal mediation is not present, particularly in the case of complete mediation. Alternatively, cross-sectional data may not support a significant mediation effect when longitudinal mediation does in fact exist (Jose, 2016; Maxwell et al., 2011). In mediation, the predictor is expected to causally precede the mediator, but in moderation the moderator is itself a predictor variable and does not necessarily require temporal precedence (Baron & Kenny, 1986). Therefore, a cross-sectional examination of moderation is more defensible.

Temporal sequencing must be carefully considered in longitudinal mediation and moderation models to account for the dynamic developmental context of effects. Mediation and moderation effects may depend on the developmental period studied. The timing of measurement in a study and the timing of the effects of the construct may differ (Cole & Maxwell, 2003; Selig et al., 2012). In mediation, researchers should carefully consider the time interval required for the predictor to have an expected effect on the mediator and the mediator to have an expected effect on the outcome (Cole & Maxwell, 2003). The magnitude of effects may vary across time intervals and should not be assumed equal, although this can be specifically tested using equality constraints in a path model. Furthermore, the length of the time lag between measurements can influence the magnitude of effects in both longitudinal mediation and moderation models (Selig et al., 2012; Selig & Preacher, 2009). The examination of effects at multiple time lags can improve the understanding of the relations (Preacher, 2015). In addition to the advantages of temporal sequencing, there are other advantages to longitudinal models. The inclusion of previous measurements of a variable control for an individual's earlier standing on that construct through autoregressive paths, which allows for more accurate estimates of paths in the mediation and moderation models (Selig & Preacher, 2009). Briefer intervals for autoregressive paths may result in larger estimates, particularly if the developmental process being examined occurs at a steady rate (Little, 2013).

1.3.3. Models to test longitudinal mediation and moderation

Structural equation modeling (SEM) is well suited to test longitudinal mediation and moderation models given their complexity. The latent variables in SEM also have the advantage of being estimated without measurement error, which can result in more accurate estimates of the paths in the model. Longitudinal mediation and moderation can be examined using several different models, including cross-lagged panel models, multigroup models, latent growth curve models, latent change score models, and multilevel models. In longitudinal mediation models the cross-lagged panel model is used most often and the latent change score model is used least often (Preacher, 2015). In longitudinal moderation models with categorical variables, multigroup models are commonly used to examine whether the longitudinal relations differ based on group membership. Although less common, it is also possible to test longitudinal moderation effects of continuous variables through the inclusion of interaction terms. All of the models account for prior measurements of variables, but in different ways.

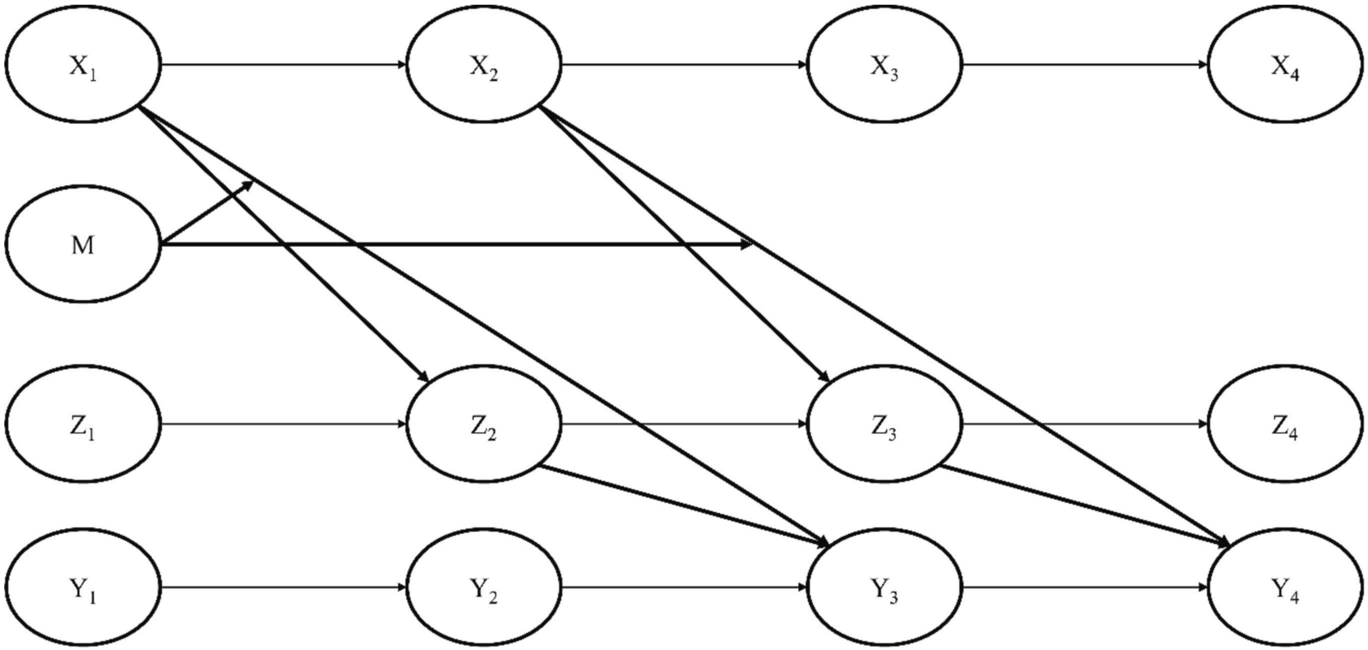

Fig. 5. Example of Longitudinal Moderated Mediation.
Note. X = predictor; Y = outcome; Z = mediator; M = moderator.

Cross-lagged panel models examine how one construct influences another construct at a later occasion. These models assume the association between constructs change as a function of time between measurements (Selig et al., 2012). In cross-lagged panel models autoregressive or stability paths link latent variables representing the same construct across time. These paths from prior measurements account for the individual's prior standing on each construct (Little, 2013). In a mediation model cross-lagged paths link the predictor, mediator, and outcome latent variables across the time-lagged measurement occasions (Preacher, 2015). For instance, researchers tested cross-lagged paths from children's achievement in one grade to their social skills in the next grade, and vice versa, from kindergarten to eighth grade (Caemmerer & Keith, 2015). In this longitudinal cross-lagged panel model the prior measurements of social skills and achievement acted as multiple temporally sequenced mediators. Longitudinal moderation is also possible in this example as the magnitude of the relations between children's social skills and achievement may differ for boys and girls. Using a multigroup SEM model, a cross-group equality constraint can be added to the cross-lagged paths to determine if they are statistically significantly different for boys and girls. If they are, this suggests that the cross-lagged relation between children's social skills and achievement is moderated by gender.

Two other longitudinal models are commonly used in research; a brief mention of them are worth including because they can also be used in longitudinal mediation and moderation studies. Latent growth models examine intraindividual change across time and interindividual variability in that intraindividual change (Little, 2013; Preacher et al., 2007). The time lag between repeated measures of the same variable is less important in latent growth models than cross-lagged panel models, but the time lag between repeated measures of different variables can have an important influence on the paths (Preacher, 2015). Latent growth models estimate individuals' intercepts (initial level of the construct) and slopes (growth in the construct across time). Intercepts and slopes typically have freely estimated means and variances, which represent average levels and individual differences in these two variables, respectively. Latent growth models are quite flexible, and both linear and non-linear growth trajectories can be tested. Preacher (2015) explained several options for the inclusion of a mediator in a latent growth model. For example, if one were interested in the relationship between externalizing behavior in kindergarten and school dropout in high school, it may be possible to determine if this relationship is mediated by reading achievement. Students who have stronger academic skills may be less likely to drop out, even though they display significant behavioral concerns in early grades. Here, school dropout would be regressed on externalizing behavior in kindergarten, and the intercept (initial level) and slope (growth) in reading achievement would be included as mediators between externalizing behavior and school dropout. Moderation can also be tested in latent growth models. Latent interactions between the intercept and slope factors may be of substantiative interest (Maslowsky et al., 2015). For example, does growth (slope) in children's reading skills depend on the starting point (intercept) of their externalizing skills?

Latent change score models, also referred to as latent difference score models (McArdle, 2001), specifically estimate the amount of change between adjacent measurements of the same variable that is interpreted as the rate of change between each interval (Preacher, 2015). These models include a latent intercept and latent linear slope (both with freely estimated means and variances), and the non-linear change over time is estimated by regressing the latent change score variable at one time point on the score at the previous time point. If this regression coefficient is negative, it suggests a deceleration in change over time. If this regression coefficient is positive, it suggests an acceleration in change over time. Latent change score models focus on change in constructs across time and the relations between change in constructs across time, but not the relations between the constructs themselves over time, which is examined in cross-lagged panel models (Preacher, 2015). In contrast to the latent growth models, change in the latent change score models is examined at single measurement periods and change is not assumed to be constant across lags (Preacher, 2015). Multiple variables are often examined in latent change score models, specifically whether changes in one variable precede or influence changes in another variable. For instance, latent change score models have examined whether changes in language precede changes in reading comprehension, whether changes in reading comprehension precede changes in language, or whether there are reciprocal relations between changes in language and reading comprehension (e.g., Quinn et al., 2015; Reynolds & Turek, 2012). Mediation can be examined using latent change score models, although this has not been extensively examined in the literature and comprehensive descriptions of how these models can be estimated have only recently been described (Hilley & O'Rourke, 2022).

Multilevel models are also well suited for the examination of longitudinal mediation and moderation (see Preacher et al., 2016, for details on multilevel SEM for moderation within and across levels of analysis). In research with large samples, children are often nested within classrooms or schools, and the effects at the school or classroom level can be examined in multilevel modeling. The repeated measurements in longitudinal data can be treated as a nesting variable as time is nested within individuals. For instance, researchers used multilevel models to test the influence of adolescents' moral disengagement, perceived school climate, and the interaction between these variables on bullying perpetration across 18 months (Teng et al., 2020). At the between-person level was moral disengagement and perceived school climate, and at the within-person level was the longitudinal change of moral disengagement and perceived school climate across the three measurement occasions. Interactions between-person, within-person, and cross-level were examined.

1.4. Integrated approaches: mediated moderation and moderated mediation

Researchers frequently examine mediation and moderation separately, but the simultaneous examination of both can answer other complex research questions. Moderated mediation can examine if a mediation effect is the same across different groups, contexts, or another continuous variable. Likewise, mediated moderation can examine if a moderation effect is mediated by another external variable. Both integrated approaches share similarities, but their start points are different as they either begin with significant moderation or significant mediation. Thus, the two integrated approaches answer different research questions (Muller et al., 2005).

Fig. 5 shows how moderated mediation appears in a longitudinal study. Here, the direct effect of X on Y is mediated by Z, but that

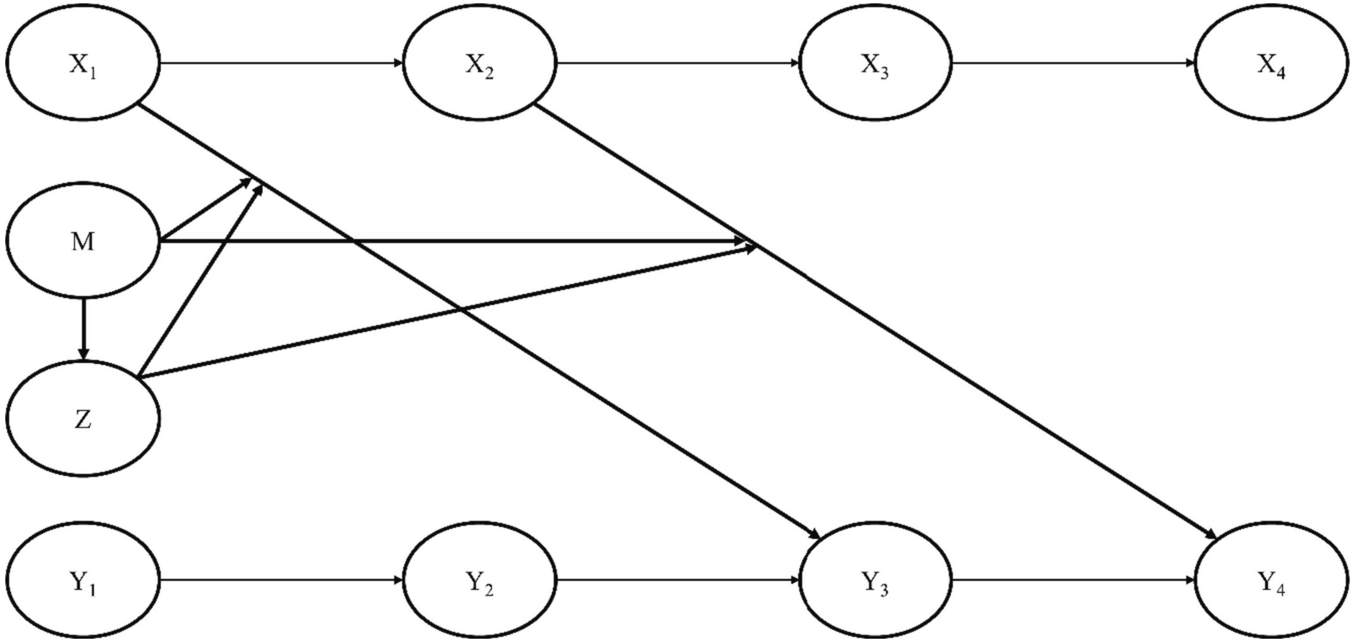

Fig. 6. Example of Longitudinal Mediated Moderation.
Note. X = predictor; Y = outcome; Z = mediator; M = moderator.

mediation effect is moderated by M. Moderated mediation, also referred to as conditional indirect effects, tests whether the influence of the mediator depends on the level of a moderator (Muller et al., 2005); the indirect effect is conditional on the value of a moderator. Although Fig. 5 only includes an effect of the moderator on the direct effect from the predictor (X) and outcome (Y), the moderator may alter the path between the predictor (X) and mediator (Z), or the mediator (Z) and outcome (Y; see Preacher et al., 2007, for five possible models). As an example, externalizing behavior in earlier grades may predict reading achievement in later grades, although this effect may be mediated by parenting style. Additionally, the mediation effect may be moderated by gender because boys tend to have higher levels of externalizing behavior and parents may use different parenting styles with boys and girls, so the relationships may not generalize across gender. For instance, using cross-sectional data, researchers found that mother's encouragement to diet mediated the influence of a child's body-mass index on the child's body dissatisfaction. There was significant moderated mediation as the mother's own body mass index moderated the effect of the mediator (Karazsia et al., 2014).

If a researcher finds a significant moderation effect, they may be curious to understand the mechanisms or intervening variables (i.e., mediators) underlying the moderation effect. Mediated moderation can explain how the moderation effect occurs; moderation occurs through a mediating variable (Muller et al., 2005). Mediated moderation tests whether the strength of a moderator, or interaction term, can be explained by a mediator (Karazsia et al., 2014; Little et al., 2007). Mediated moderation implies the overall moderation effect is reduced once the mediation effect is included (Muller et al., 2005). Fig. 6 shows how mediated moderation would appear in a longitudinal study. Essentially, the relationship between X and Y is moderated by M, but the moderation effect is also mediated by Z. As an example, externalizing behavior in earlier grades may predict reading achievement in later grades, but this effect may be moderated by gender because boys tend to have higher levels of externalizing behavior than girls. However, this moderation effect may be mediated by parenting style, where the moderation between externalizing behavior and reading achievement is mediated by parenting style. These types of questions are quite complex and require careful thought and grounding in theory and previous research. Mediated moderation is less common in the literature but has been used in some research. For instance, using longitudinal data, researchers found a significant interaction between parental divorce and child gender on the child's math performance. Children's approaches to learning were tested as a potential mediator of the moderation effect. Mediated moderation was not supported, however, because the magnitude of the interaction did not change due to the mediator (Anthony et al., 2014).

2. Application of longitudinal mediation, moderation, moderated mediation, and mediated moderation methods

2.1. Sample

Using a large national sample we demonstrate the longitudinal mediation, moderation, moderated mediation, and mediated moderation approaches described above. The sample was the public version of the Early Childhood Longitudinal Study: Kindergarten 2011 (ECLS-K:2011) Cohort, which is conducted by the National Center for Education Sciences. For more information, see https://nces.ed.gov/ecls/kindergarten2011.asp. Our samples included 7536 (Model 5), 8677 (Model 3), 16,939 (Model 2 and Model 4), and 16,967 children (Model 1 and baseline model) who were followed from kindergarten to fifth grade. Children who repeated kindergarten were excluded from our analyses. The sample was 48.9% female, and race and ethnicity included 46.6% white, 13.1% Black/African American, 25.4% Hispanic, 8.6% Asian, 0.6%, Native Hawaiian/Pacific Islander, 0.9% Native American/Alaskan Native, and 4.5% identified with multiple categories. Median household income was between $50,000 and $60,000 depending on the wave.

2.2. Models

We tested five substantiative longitudinal models for demonstration purposes. In our models we analyzed four to six occasions of measures separated by a year lag. We used data collected during the spring of kindergarten, first, second, third, fourth, and fifth grades, but fewer time points were used for some measures. Structural equation modeling was used; the latent variable approach had the advantage of accounting for measurement error in our models and allowed for tests of measurement invariance across grade levels and groups. There were four main variables of interest, including approaches to learning, attentional focus, academic achievement, sex, and income. These variables are described in more detail below.

The longitudinal predictive relation of children's approaches to learning on their later academic achievement served as the underlying model to which we introduced mediation, moderation, moderated mediation, and mediated moderation variables. First, we examined if children's attentional focus mediated the relation between their earlier approaches to learning on their later academic achievement (i.e., longitudinal mediation). Second, we examined if children's sex moderated the relation between children's approaches to learning on their later academic achievement (i.e., longitudinal moderation with a time-invariant categorical variable). Third, we examined if time-varying household income moderated the relation between children's approaches to learning on their later academic achievement (i.e., longitudinal moderation with a time-variant continuous variable). Fourth, we examined if children's sex moderated the mediated relation of their approaches to learning and attentional focus on their later academic achievement (i.e., moderated mediation). Fifth, we examined if children's attentional focus mediated the moderated relation of families' household income and children's approaches to learning on their later academic achievement (i.e., mediated moderation).

2.3. Measures

Academic achievement was measured using direct child assessments administered during the spring of each year. A latent variable for achievement was measured using three tests, including reading, mathematics, and science. These tests were developed specifically for the ECLS-K:2011 study. Test content was informed by national and state performance standards, state and commercial assessments, and curriculum experts. Reading tests assessed children's basic skills, vocabulary, and reading comprehension. Mathematics tests assessed children's number properties and operations, measurement, geometry, data analysis and probability, and algebra. Science tests assessed children's scientific inquiry, life science, physical science, and Earth and space science (Tourangeau et al., 2019). Scores for each test are interval level and derived through item response theory. Theta scores were estimated for each student and those scores were converted to scaled scores representing the number of items each student would have answered correctly had they been administered the entire test (Tourangeau et al., 2019). Achievement from second to fifth grade was included in our models and Omega values were 0.88 across all grade levels.

Approaches to learning is a teacher rating scale created specifically for the ECLS-K study that measures student learning behaviors. The scale included seven items about being organized, being eager to learn, working independently, easily adapting to changes in routine, persisting on tasks, paying attention, and following classroom rules. For the present study, the attention item was excluded because attention was a separate construct of interest, so only six items were used. All items were rated on a four-point Likert frequency scale (1 = *Never*, 2 = *Sometimes*, 3 = *Often*, and 4 = *Almost Always*); a not applicable option was provided. Across waves of the ECLS-K:2011 dataset children's approaches to learning positively correlated with their interpersonal and self-control skills and negatively correlated with their externalizing and internalizing behaviors (Tourangeau et al., 2019). Approaches to learning from spring kindergarten through fifth grade was included in our models, and Omega values for approaches to learning ranged from 0.95 to 0.96 across grade levels.

Attentional focus is a teacher rating scale that measures how well students pay attention and can hold their concentration during class. The attentional focus subscale from the Temperament in Middle Childhood Questionnaire was used (Simonds & Rothbart, 2004). The ECLS-K:2011 study used six of the seven items from the subscale. Items were about being able to pay attention for long periods when listening to stories, being easily distracted, shifting to a different task without completing an earlier task, concentrating well, and working for long periods during class activities. Teachers were asked to consider the past 6 months and rate children's reactions to situations on a five-point Likert scale (1 = *Almost always untrue*, 2 = *Usually untrue*, 3 = *Sometimes true, sometimes untrue*, 4 = *Usually true,* and 5 = *Almost always true*); a not applicable option was provided (Tourangeau et al., 2019). Attentional focus from second to fifth grade was included in our models. A different measure of attentional focus was used in kindergarten and first grade, therefore earlier measurements of attentional focus were not included in our model. Omega values for Attentional focus ranged between 0.95 and 0.96 for both boys and girls.

Child sex and household income were reported by parents. Child sex was reported in kindergarten. Household income was reported in $5000 increments from $5000 or less to $75,000 (15 categories) and three higher categories ranging from $75,001 to $200,001 (Tourangeau et al., 2019). Due to the unequal intervals of the income variable, categories were transformed to the midpoint of the range and then *z*-scores were created. Household income was a time-varying variable. We analyzed measurements from the spring kindergarten, first, second, third, and fourth grade waves.

2.4. Data analysis

M*plus* 8.2 (Muthén & Muthén, 1998–2017) was used for modeling, whereas R was used for data preparation and summary of results (R Core Team, 2022). Full-information maximum likelihood estimation was used for all models to account for missing data due to nonresponse and attrition during the study. Maximum likelihood estimation is considered an appropriate method for estimating model parameters when data are assumed to be missing at random (MAR; Enders, 2010; Rubin, 1987). Missingness across waves ranged from 24.2% to 37.4% for achievement, 12.2% to 43.3% for approaches to learning, 30.7% to 43.4% for attentional focus, 25.6% to 43.8% for household income, and <1% in kindergarten for child sex.

2.4.1. Measurement invariance across time and groups

Prior to testing the five substantiative models, tests of factorial invariance were conducted. We specifically tested for weak (i.e., metric) invariance by conducting a series of factorial invariance tests to ensure achievement, approaches to learning, and attentional focus demonstrated both longitudinal invariance (i.e., invariance across grade levels) and cross-group invariance (i.e., invariance across gender groups). First, for longitudinal invariance a model was estimated without constraints over time, and then equality constraints were added to the corresponding factor loadings across each time point. Similarly, for gender invariance a model without constraints across gender groups was estimated and then cross-group equality constraints were added to the factor loadings. Metric invariance tested if the latent constructs were equivalently scaled and allowed for comparisons of paths across time and groups. We took a modeling rationale perspective by selecting a model based on overall good fit (Little, 1997) to determine if the assumption of metric invariance was supported. Strong or intercept invariance was not tested because latent means were not examined in this study (Keith, 2019).

2.4.2. Cross-lagged panel models

Longitudinal mediation, longitudinal moderation, longitudinal moderated mediation, and longitudinal mediated moderation were tested using cross-lagged panel models. Longitudinal panel models test predictive relations between latent constructs across time and examine individual differences between children as they change across time (Little, 2013). For each construct in this study (i.e., achievement, approaches to learning, and attentional focus), autoregressive paths were estimated across grade levels. Autoregressive paths linked each construct with itself across time and accounted for children's prior standings on each construct. Cross-lagged paths tested the influence from children's approaches to learning to their achievement at a later grade. It is important to note that reciprocal or bidirectional influences between two or more variables are often examined in cross-lagged panel models (Hamaker et al., 2015); however, our models examined only one direction of effects. Correlated disturbances were included across latent constructs within each grade level because directionality was not assumed (e.g., between approaches to learning and achievement, approaches to learning, attentional focus, and achievement, or approaches to learning, household income, and achievement). The residuals for the six approaches to learning items, six attentional focus items, and three achievement scales were correlated with the same item or scale at each time point to account for their relation with each other beyond the relation with the latent variable (Little, 2013). These correlated residuals across time account for the shared method variance typical in longitudinal research, which can cause inflated path coefficients (Cole & Maxwell, 2003).

2.4.3. Longitudinal baseline model

Six models were tested starting with a baseline model. The baseline model included autoregressive paths for achievement and approaches to learning across grade levels and cross-lagged paths from approaches to learning to achievement 2 years later. This model is shown in Fig. 7 and the M*plus* code is included in the Appendix. Correlated disturbances across latent constructs as well as correlated residuals for individual items and scales are not shown in the figures but were included as described above. Item loadings were fixed to equality over time, indicating longitudinal invariance. The baseline model provided estimates for the effect of approaches to learning on achievement as well as model fit information prior to the inclusion of mediators or moderators. This is Step 1 in Baron and Kenny's (1986) casual steps procedure for testing mediation. After the baseline model was estimated, five models of interest were estimated.

2.4.4. Longitudinal mediation

First, the longitudinal mediation model included attentional focus as the mediator between approaches to learning and achievement. Paths from approaches to learning in Grades 1, 2, 3, and 4 to attentional focus in the next year, as well as from attentional focus at Grades 2, 3, and 4 to achievement at the next year, were added (see Fig. 8). Tests for the indirect effect of approaches to learning on achievement via focus are available in M*plus* using the MODEL INDIRECT command. The estimate for the indirect effect is not normally distributed; the ANALYSIS: BOOTSTRAP = 1000 command requests bootstrapped confidence intervals to use in place of analytic standard errors (Preacher et al., 2007).

2.4.5. Longitudinal moderation with a time-invariant categorical variable

Second, in the longitudinal moderation categorical variable model, children's sex was the moderator between approaches to learning and achievement (Fig. 9, top panel). This was modeled in M*plus* by assigning sex as a grouping variable. The effect from approaches to learning on achievement was allowed to vary across groups, whereas the remaining paths were constrained equal across groups (see Appendix). The MODEL TEST command jointly tested for a sex difference between paths from approaches to learning to achievement at all waves using a Wald χ^2 test (Muthén & Muthén, 1998–2017).

2.4.6. Longitudinal moderation with a time-variant continuous variable

Third, in the longitudinal moderation continuous variable model household income was tested as a moderator (Fig. 9, bottom panel). In M*plus*, an interaction between the observed household income variable and latent approaches to learning variable was defined at each wave using the XWITH option. Then achievement was regressed on household income, approaches to learning, and the

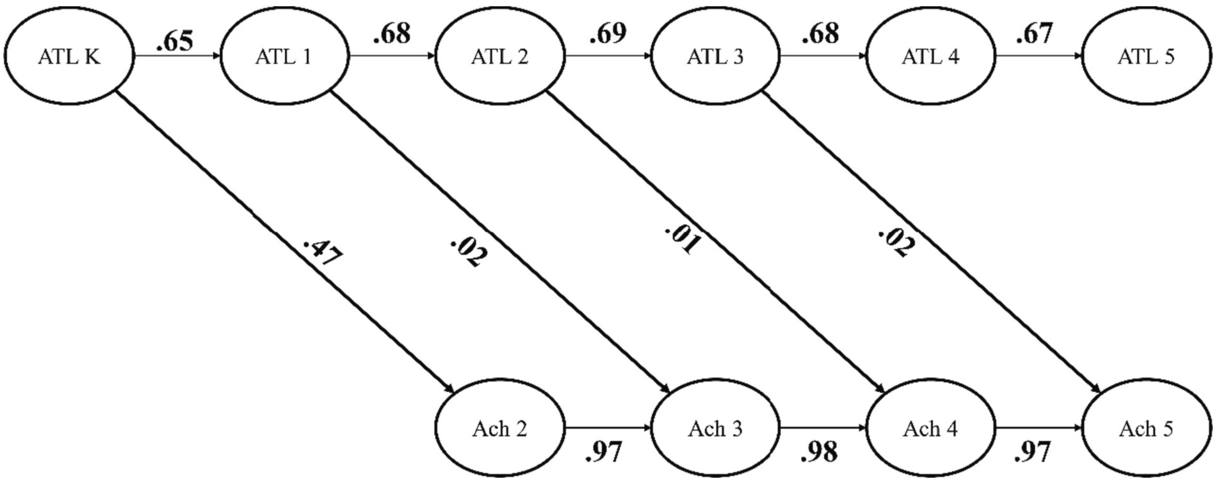

Fig. 7. Baseline Longitudinal Model 0 for the Effect of Approaches to Learning on Achievement.
Note. Standardized coefficients are reported. ATL = Approaches to Learning; Ach = Achievement. Correlated residuals between the same individual items and scales at each time point and correlated disturbances between latent variables measured at the same time point were included in Models 0–5 but are not shown in Figs. 7–11.

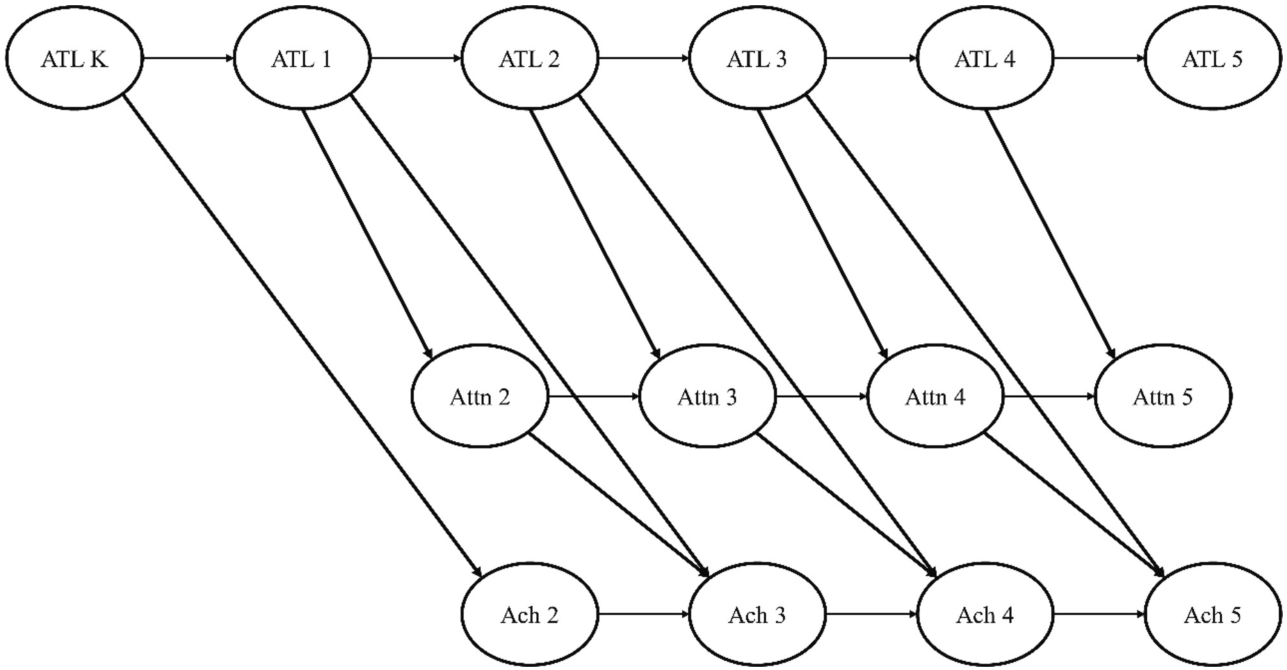

Fig. 8. Model 1 Longitudinal Mediation Model.
Note. ATL = Approaches to Learning; Ach = Achievement; Attn = Attentional Focus.

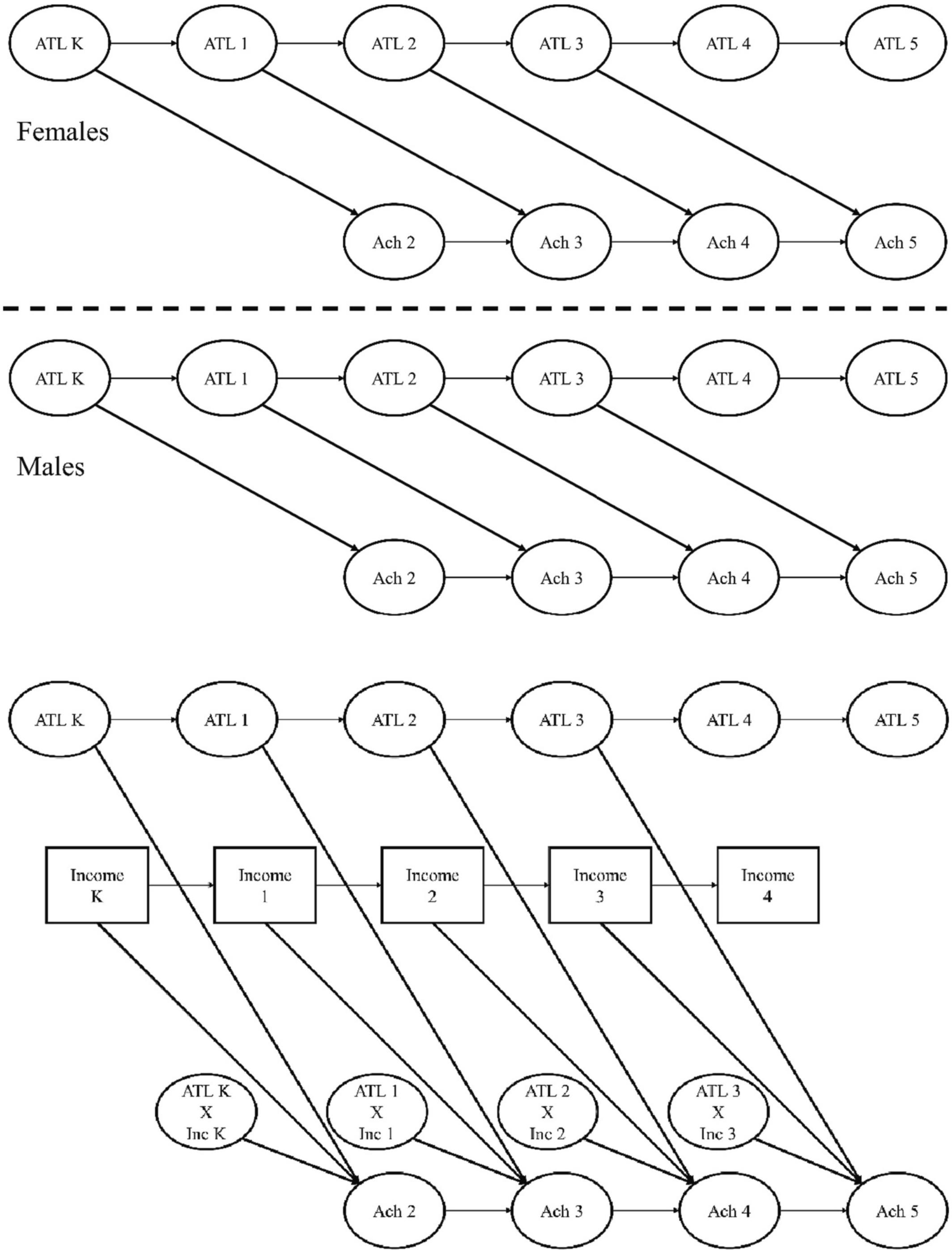

Fig. 9. Longitudinal Moderation: Model 2 Categorical Time-Invariant (Top) and Model 3 Continuous Time-Variant Variable (Bottom). *Note*. ATL = Approaches to Learning; Ach = Achievement.

interaction between income and approaches to learning. Because the interaction term included a latent variable a random effect was specified using TYPE = RANDOM in the ANALYSIS section. In addition, ALGORITHM = INTEGRATION was used to estimate robust standard errors with a numerical integration algorithm (Muthén & Muthén, 1998–2017).

2.4.7. Longitudinal moderated mediation

Fourth, in the moderated mediation model children's sex was tested as a moderator of the mediated relation between approaches to learning, attentional focus, and achievement (see Fig. 10). This was coded in M*plus* by defining the indirect effect via attentional focus

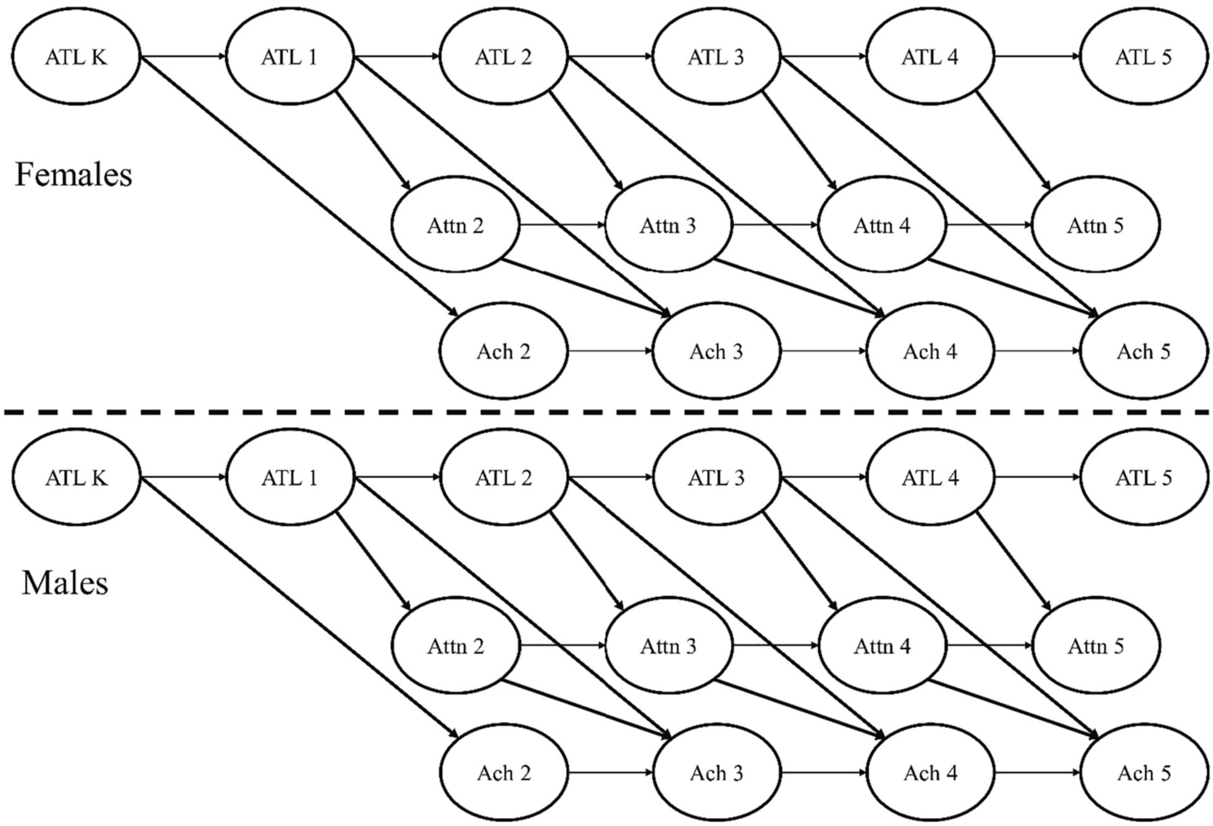

Fig. 10. Model 4 Longitudinal Moderated Mediation.
Note. ATL = Approaches to Learning; Ach = Achievement; Attn = Attentional Focus.

as in the longitudinal mediation model. Then, sex was used as a grouping variable as in the moderation with a categorical variable model. The MODEL CONSTRAINT option was used to define parameters representing the mediation effect as the product of the direct effects from approaches to learning to attentional focus and from attentional focus to achievement. Then MODEL TEST was used to jointly test whether there was any difference between girls and boys across waves.

2.4.8. Longitudinal mediated moderation

Fifth, in the mediated moderation model children's attentional focus was tested as a mediator of the moderated relation between families' household income and children's approaches to learning, which formed the interaction term, on children's later academic achievement (see Fig. 11).

Mediated moderation models pose significant challenges to estimation, especially in latent variable models. The default settings in M*plus* require TYPE = RANDOM when calculating latent interaction terms; however, the MODEL INDIRECT command is not available under TYPE = RANDOM. Researchers can explore alternatives to estimate mediated moderation models. Observed factor scores can be estimated for the latent variables and used in the model, which eliminates the need for TYPE = RANDOM. However, the model will no longer estimate the error terms for the latent variables; the scores are not considered error-free. Alternatively, the latent variables could be estimated without the MODEL INDIRECT command. The indirect pathways can be calculated by hand, but the bootstrapped standard errors will not be available. We illustrated the mediated moderation model with the estimated observed factor scores for the latent approaches to learning variables. These observed approaches to learning factor scores were used to create the observed interaction terms with the observed families' household income variables.

2.4.9. Model fit

Model fit was assessed using the χ^2, the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). The χ^2 was expected to be statistically significant because it is highly sensitive to sample size. CFI and TLI values between 0.95 and 0.97 suggest good fit, whereas values >0.97 suggest excellent fit. RMSEA values between 0.05 and 0.08 suggest good fit and values <0.05 suggest excellent fit. SRMR values between 0.05 and 0.10 suggest good fit, whereas values <0.05 suggest excellent fit (Schermelleh-Engel et al., 2003).

For nested model comparisons and factorial invariance tests, the $\Delta\chi^2$ was used, but is also sensitive to sample size and may be statistically significant with trivial sources of misfit. For factorial invariance, the ΔCFI ≤ -0.01 has been commonly used for

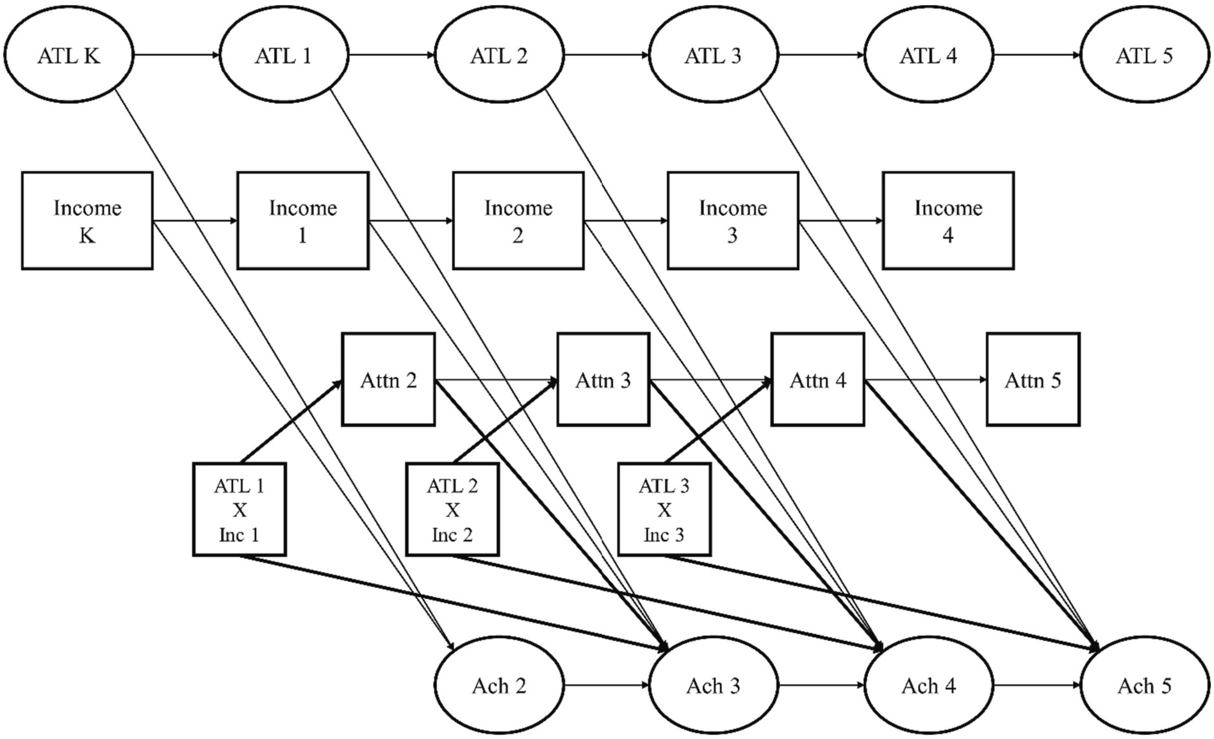

Fig. 11. Model 5 Longitudinal Mediated Moderation.
Note. ATL = Approaches to Learning; Ach = Achievement; Attn = Attentional Focus; Inc. = household income.

identifying noninvariance (Cheung & Rensvold, 2002; Little, 2013).

3. Results

3.1. Tests of measurement invariance

We fit models with loadings allowed to vary across time (longitudinal baseline), then fixed them equal (longitudinal invariance) and compared measures of model fit for each latent variable: achievement, approaches to learning, and attentional focus. Once we determined that longitudinal invariance held, we allowed loadings to vary by gender (group baseline) and fixed loadings equal across gender (group invariance). We examined the change in χ^2, CFI, TLI, AIC, and BIC to determine if assumptions of longitudinal and group invariance were violated.

The χ^2 change tested longitudinal and group invariance in achievement. Approaches to learning and attentional focus were all statistically significant ($p < .001$). Change in CFI and TLI between baseline and invariance models were always below 0.004, which is below the 0.01 threshold for invariance models. Thus, longitudinal and gender group invariance was supported for achievement, approaches to learning, and attentional focus using a modeling rationale (Little, 1997).

3.2. Longitudinal baseline: Model 0

Fig. 7 presents the baseline model for the effect of approaches to learning on achievement at Grades 2, 3, 4, and 5. Parameter estimates for this model help the researcher understand the magnitude of effects of the independent variables on the dependent variables. Model fit indices suggested good to excellent fit except for the SRMR (CFI = 0.966, TLI = 0.961, RMSEA = 0.031, SRMR = 0.121). With no other latent variables in the model, the unstandardized regression estimates were all significant at the $p < .05$ level. The unstandardized longitudinal autoregressive effects of approaches to learning at each grade level on approaches to learning at the next grade level ranged from 0.66 to 0.70 (β = 0.65–0.69). The unstandardized longitudinal autoregressive effects of achievement at each grade level on the next grade level ranged from 0.95 to 1.01 (β = 0.97–0.98). The unstandardized effect of approaches to learning on achievement 2 years later ranged from 0.20 to 0.48 (β = 0.01–0.47). The effect of approaches to learning on achievement at a given grade level was smaller than the effect of prior achievement.

Although the measurement model is not shown in Fig. 7, the six approaches to learning items and three achievement scales loaded highly on their respective latent variables at each grade level, with standardized loadings ranging from 0.66 to 0.85 and from 0.82 to 0.87, respectively. Most of the correlated residuals between the same approaches to learning items and achievement scales across grades were statistically significant ($p < .05$) except for the approaches to learning persists on tasks item, which did not significantly correlate with itself across time except for the fourth to fifth grade correlation. Non-significant correlated residuals were not pruned from any models.

3.3. Longitudinal mediation: Model 1

The longitudinal mediation model fit indices suggested good to excellent fit except for the SRMR (CFI = 0.966, TLI = 0.963, RMSEA = 0.026, SRMR = 0.133). The SRMR was only slightly worse than in the baseline model. The direct path from approaches to learning to attentional focus was statistically significant and negative at each time point (see Table 1; $\beta = -0.64$ to -0.45). The direct path from attentional focus to achievement was statistically significant and negative from second to third and fourth to fifth grades ($\beta = -0.02$) but was not statistically significant from third to fourth grade. The total effect of approaches to learning on achievement was positive and statistically significant from first to third ($c + ab$ $\beta = 0.02$) and third to fifth grades ($c + ab$ $\beta = 0.01$), but not from second to fourth grade ($\beta = 0.00$). The direct path from approaches to learning to achievement (c) was never statistically significant as the confidence interval always included zero. However, the indirect path via attentional focus was statistically significant from first to third (ab $\beta = 0.01$) and third to fifth (ab $\beta = 0.01$) grades. This indicates full mediation of the effect of approaches to learning on achievement in the first to third and third to fifth grade. Approaches to learning influenced achievement via the attentional focus variable (see Table 1).

Attentional focus was introduced to this model and because the measurement model is not shown, estimates for attentional focus are discussed here. The six attentional focus items loaded highly on their respective latent variables at each grade level; standardized loadings ranged from 0.82 to 0.93. Approximately half of the correlated residuals between attentional focus items and their measurements at later grades were statistically significant ($p < .05$); two of the attentional focus items did not correlate with their later measurements across time.

3.4. Longitudinal moderation with a time-invariant categorical variable: Model 2

The longitudinal moderation with a categorical variable model fit indices suggested good to excellent fit except for the SRMR (CFI = 0.959, TLI = 0.954, RMSEA = 0.033, SRMR = 0.125). The SRMR was only slightly worse than in the baseline model. The Wald χ^2 test was statistically significant (χ^2 (3) = 22.169, $p < .001$), indicating there was some difference in the effect of approaches to learning on achievement for boys and girls. The regression estimates indicated that the effect of approaches to learning on achievement was stronger for girls and statistically significant at all four grades, whereas the effect of approaches to learning on achievement was only statistically significant for boys from kindergarten to second and third to fifth grades (Table 2). The relationship between approaches to

Table 1
Model 1 Unstandardized Estimates of the Mediation of the Effect from Approaches to Learning on Achievement by Attentional Focus.

Outcome	Predictor		Estimate	SE[1]	95% CI LL	95% CI UL
Achievement at Grade 3						
	ACH Grade 2		0.952*	0.004	0.943	0.961
	Focus Grade 2		−0.211*	0.071	−0.350	−0.074
	ATL Grade 1					
		Direct Effect	0.153	0.123	−0.088	0.404
		Indirect Effect[2]	0.210*	–	0.072	0.349
		Total Effect	0.363*	0.099	0.178	0.568
Achievement at Grade 4						
	ACH Grade 3		0.948*	0.005	0.937	0.957
	Focus Grade 3		−0.115	0.063	−0.238	0.005
	ATL Grade 2					
		Direct Effect	−0.013	0.115	−0.217	0.233
		Indirect Effect[3]	0.081	–	−0.004	0.164
		Total Effect	0.069	0.098	−0.113	0.266
Achievement at Grade 5						
	ACH Grade 4		1.012*	0.005	1.003	1.021
	Focus Grade 4		−0.193*	0.074	−0.331	−0.044
	ATL Grade 3					
		Direct Effect	0.136	0.119	−0.085	0.389
		Indirect Effect[4]	0.138*	–	0.033	0.241
		Total Effect	0.274*	0.096	0.092	0.460

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit; ATL = Approaches to Learning; ACH = Achievement.
[1] Standard errors for indirect effects should not be interpreted.
[2] Specific indirect effect via Attentional Focus at Grade 2.
[3] Specific indirect effect via Attentional Focus at Grade 3.
[4] Specific indirect effect via Attentional Focus at Grade 4.
* $p < .05$.

Table 2
Model 2 Unstandardized Estimates of the Moderation of the Effect of Approaches to Learning on Achievement by Sex.

	Boys			Girls		
	Estimate	SE	p	Estimate	SE	p
Achievement at Grade 2						
ATL Grade K 10.489		0.287	< 0.001*	10.372	0.300	< 0.001*
Achievement at Grade 3						
ACH Grade 2	0.951	0.005	< 0.001*	0.951	0.005	< 0.001*
ATL Grade 1	0.198	0.122	0.106	0.707	0.132	< 0.001*
Achievement at Grade 4						
ACH Grade 3	0.942	0.004	< 0.001*	0.942	0.004	< 0.001*
ATL Grade 2	0.121	0.114	0.290	0.457	0.122	< 0.001*
Achievement at Grade 5						
ACH Grade 4	1.014	0.005	< 0.001*	1.014	0.005	< 0.001*
ATL Grade 3	0.347	0.118	0.003*	0.558	0.131	< 0.001*

Note. ATL = Approaches to Learning; ACH = Achievement.
* $p < .01$.

learning and achievement was moderated by gender.

3.5. Longitudinal moderation with a time-variant continuous variable: Model 3

Standalone fit indices are not available in M*plus* for the longitudinal moderation with a time-variant continuous variable model due to the latent approaches to learning variable in the interaction and use of the ALGORITHM = INTEGRATION option. The significance of household income as a moderator was tested using four interaction terms between time-varying household income and approaches to learning (see Table 3 for the time points involved). The regression estimates in Table 3 indicate household income in Grade 2 and Grade 3 was not a statistically significant predictor of achievement in Grade 4 or Grade 5, whereas the effects of income in kindergarten on Grade 2 achievement and income in Grade 1 on Grade 3 achievement were significant. The prior year's achievement, as well as the approaches to learning variable (except for Grade 2 approaches to learning on Grade 4 achievement), were both significant predictors with positive regression estimates. The interaction between household income and approaches to learning was significant at all four time points, but the regression estimate was negative. Taken together, this indicates that as approaches to learning increased, so did achievement; however, the impact was softened for students with a high household income. Conversely, the effect of approaches to learning on achievement was strengthened for students from a low household income. The relationship between approaches to learning and achievement was moderated by household income.

3.6. Longitudinal moderated mediation: Model 4

The longitudinal moderated mediation model fit indices suggested good to excellent fit except for the SRMR (CFI = 0.962, TLI =

Table 3
Model 3 Unstandardized Estimates of the Moderation of the Effect of Approaches to Learning on Achievement by Household Income.

	Estimate	SE	p
Achievement at Grade 2			
ATL Grade K 9.328		0.311	< 0.001*
Income Grade K 3.613		0.142	< 0.001*
Income X ATL Grade K -1.036		0.287	< 0.001*
Achievement at Grade 3			
ACH Grade 2	0.930	0.006	< 0.001*
ATL Grade 1	0.442	0.126	< 0.001*
Income Grade 1	0.160	0.061	0.009*
Income X ATL Grade 1	−0.556	0.100	< 0.001*
Achievement at Grade 4			
ACH Grade 3	0.943	0.006	< 0.001*
ATL Grade 2	0.170	0.114	0.136
Income Grade 2	0.150	0.060	0.012
Income X ATL Grade 2	−0.336	0.099	< 0.001*
Achievement at Grade 5			
ACH Grade 4	1.002	0.007	< 0.001*
ATL Grade 3	0.507	0.118	< 0.001*
Income Grade 3	0.024	0.063	0.698
Income X ATL Grade 3	−0.448	0.107	< 0.001*

Note. ATL = Approaches to Learning; ACH = Achievement.
* $p < .01$.

0.959, RMSEA = 0.027, SRMR = 0.129). The Wald χ^2 test for the difference in the paths was not statistically significant (1.823 with DF = 3, $p > .05$). This indicates no significant difference by gender in the indirect effect of approaches to learning on achievement via attentional focus. Examining the estimates, the confidence intervals overlapped for the indirect effects for girls and boys (Table 4). The mediation effect was not moderated by sex.

3.7. *Longitudinal mediated moderation: Model 5*

The longitudinal moderated mediation model fit indices suggested good to excellent fit except for the SRMR (CFI = 0.949, TLI = 0.944, RMSEA = 0.047, SRMR = 0.194). The confidence intervals included zero for the indirect effect from the interaction terms to achievement (Table 5). This indicates that there was no significant difference in the moderation effect of income for children with different levels of attentional focus. The moderation effect was not mediated by attentional focus.

4. Discussion

4.1. *Interpretation of results*

4.1.1. *Mediation (Model 1)*

Recall that mediation decomposes the effect of the independent variable on the dependent variable into direct effects and indirect effects via the mediating variables. When the direct effect is not statistically significant, but the indirect effect is, the researcher can conclude that the effect is fully mediated by the third variable. This was the case for two of the three time points in Model 2; attentional focus *fully mediated* the effect of approaches to learning on achievement. When the indirect effect is not statistically significant (as occurred with approaches to learning at Grade 2 to achievement at Grade 4), there is no evidence of mediation. When both direct and indirect effects are statistically significant, there is evidence of *partial mediation* (not observed in our model).

4.1.2. *Moderation (Model 2 and Model 3)*

When the effect from the independent variable to the dependent variable differs between groups, there is evidence of moderation. In the longitudinal context, we use an omnibus test (Wald χ^2 test) to test if there are any differences and then examine the regression results to determine where those differences likely occur. In this case, the omnibus test was statistically significant, indicating a difference between girls and boys in the effect of approaches to learning on achievement. The regression estimates in Table 2 were used to determine where the differences were largest. If the omnibus test was not statistically significant, we would have concluded that gender *did not moderate* the relationship between approaches to learning and achievement.

When the moderator is a continuous variable, as family income is, it is modeled as an interaction with the independent variable. In the present study, the interaction was always statistically significant, indicating that the effect of approaches to learning on achievement was moderated by family income. However, the regression results in Table 3 must be interpreted carefully because the

Table 4
Model 4 Unstandardized Estimates of the Moderation by Sex of the Mediation of Approaches to Learning on Achievement by Attentional Focus.

	Boys	95% CI		Girls	95% CI	
	Estimate	LL	UL	Estimate	LL	UL
Achievement at Grade 3						
ACH Grade 2	0.949*	0.940	0.959	0.949*	0.940	0.959
Focus Grade 2	−0.241*	−0.421	−0.061	−0.173	−0.371	0.026
ATL Grade 1						
Direct Effect	−0.015	−0.312	0.282	0.548*	0.227	0.870
Indirect Effect	0.227*	0.057	0.397	0.162	−0.024	0.347
Total Effect	0.212	−0.024	0.449	0.710	0.455	0.965
Achievement at Grade 4						
ACH Grade 3	0.944*	0.935	0.952	0.944*	0.935	0.952
Focus Grade 3	−0.209*	−0.382	−0.035	−0.077	−0.259	0.105
ATL Grade 2						
Direct Effect	−0.167	−0.446	0.113	0.535	−0.018	0.584
Indirect Effect	0.139*	0.023	0.255	0.164	−0.075	0.185
Total Effect	−0.028	−0.258	0.203	0.544*	0.092	0.584
Achievement at Grade 5						
ACH Grade 4	1.014*	1.004	1.023	1.014*	1.004	1.023
Focus Grade 4	−0.164	−0.126	0.060	−0.232*	−0.436	−0.028
ATL grade 3						
Direct Effect	0.139	−0.158	0.437	0.236	−0.092	0.565
Indirect Effect	0.083	−0.040	0.206	0.160*	0.018	0.301
Total Effect	0.223	−0.018	0.463	0.396*	0.130	0.662

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit; ATL = Approaches to Learning; ACH = Achievement. Indirect Effect represents the specific indirect effect of ATL on achievement via Attentional Focus.
* $p < .01$.

Table 5
Model 5 Unstandardized Estimates of the Mediation by Attentional Focus on the Moderation of the Effect of Approaches to Learning on Achievement by Household Income.

Outcome	Predictor		Estimate	SE[1]	95% CI	
					LL	UL
Achievement at Grade 3						
	ACH Grade 2		0.927*	0.007	0.908	0.941
	Focus Grade 2		0.026	0.082	−0.146	0.178
	ATL Grade 1		0.310*	0.152	0.001	0.603
	Income Grade 1		0.207*	0.066	0.062	0.330
	ATL X Income Grade 1					
		Direct Effect	−0.597*	0.104	−0.795	−0.383
		Indirect Effect[2]	−0.001	–	−0.008	0.005
		Total Effect	−0.598*	0.104	−0.795	−0.385
Achievement at Grade 4						
	ACH Grade 3		0.943*	0.007	0.928	0.956
	Focus Grade 3		−0.002	0.076	−0.146	0.147
	ATL Grade 2		0.022	0.137	−0.270	0.269
	Income Grade 2		0.154*	0.063	0.034	0.279
	ATL X Income Grade 2					
		Direct Effect	−0.289*	0.102	−0.467	−0.081
		Indirect Effect[3]	0.000	–	−0.002	0.002
		Total Effect	−0.289*	0.102	−0.465	−0.081
Achievement at Grade 5						
	ACH Grade 4		1.000*	0.007	0.0986	1.014
	Focus Grade 4		−0.038	0.087	−0.217	0.124
	ATL Grade 3		0.295	0.148	−0.012	0.567
	Income Grade 3		0.060	0.068	−0.075	0.187
	ATL X Income Grade 3					
		Direct Effect	−0.439*	0.112	−0.646	−0.202
		Indirect Effect[4]	0.000	–	−0.001	0.006
		Total Effect	−0.439*	0.112	−0.646	−0.201

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit; ATL = Approaches to Learning; ACH = Achievement.
[1] Standard errors for indirect effects should not be interpreted.
[2] Specific indirect effect via Attentional Focus at Grade 2.
[3] Specific indirect effect via Attentional Focus at Grade 3.
[4] Specific indirect effect via Attentional Focus at Grade 4.

interaction term was negative. The effect of approaches to learning is positive (all time points); as approaches to learning increased, so did achievement. The effect of income was positive in kindergarten and first grade in that as income increased, so did achievement. The effect of income was not significant in second and third grade (income had no impact on the relationship between approaches to learning and achievement given the other variables in the model). The *interaction* was negative as when both approaches to learning and income increased, the effect of approaches to learning on achievement decreased.

4.1.3. Moderated mediation (Model 4)

Recall that moderated mediation (Model 4) occurs when the indirect effect (mediation) depends on the value of a moderator. In the present study, attentional focus mediated the relation between approaches to learning and achievement from Grades 1 to 3 and from Grades 3 to 5 (Table 1). In the fourth model, we tested whether the mediation relationship was moderated by gender (i.e., Was the indirect effect different for boys and girls?) We look to the confidence intervals for indirect effects (as with the mediation model) and we test the differences using an omnibus test (as with the moderation model). The omnibus test was not statistically significant, indicating no moderated mediation. Had the test been statistically significant, we would look to the indirect effects in Table 4 to determine the time points where boys and girls differed in the mediation of the relation between approaches to learning and achievement via attentional focus.

4.1.4. Mediated moderation (Model 5)

Mediated moderation describes moderated pathways that are mediated by a fourth variable. In the continuous moderation example (Model 3), family income moderated the relations between children's approaches to learning and achievement. This is modeled by including an interaction term between approaches to learning and income. The significant interaction indicated the relations between children's approaches to learning and achievement were different for children from families with different incomes. In the fifth model, we tested whether the moderation relation itself was mediated (i.e., Was the path from the interaction term to achievement mediated by children's attentional focus?). This model was tested for illustrative purposes and was not theoretically grounded.

The complexity of the mediated moderation model made its estimation and interpretation difficult. The indirect paths from the interaction to achievement were not statistically significant, indicating no mediated moderation (Table 5). Had the paths been significant we could have concluded the moderation of income on the relations between approaches to learning and achievement was partially or fully attributed to differences in attentional focus for children from families with different incomes.

4.2. Contextualization of results

Our findings are consistent with other longitudinal studies that have found children's approaches to learning predicted their later achievement with a small to moderate effect size (Anthony & Ogg, 2019, 2020; Barnard-Brak et al., 2016; Claessens et al., 2009; Duncan et al., 2007; Gullo & Impellizeri, 2022; Li-Grining et al., 2010; Sung & Wickrama, 2018). These studies also used the ECLS-K dataset, but more often used the older 1999 dataset, and most of these studies did not include children's science performance in their measures of achievement. Unsurprisingly, shorter time lags between the measurement of approaches to learning and achievement in other studies, such as those spaced 6–12 months apart, generally produced larger effects. This may explain the small effects observed in our study as the time lags were 2 years to allow for tests of longitudinal mediation.

In terms of our moderation results, there is some support for moderation of the approaches to learning and achievement relation by children's sex and SES. For moderation by sex (Model 2), Kuo et al. (2021) found the longitudinal relation between children's self-regulation, which overlaps to some extent with approaches to learning, and later achievement was stronger for high-achieving adolescent girls than high-achieving boys; we similarly found girls' approaches to learning had a stronger influence on their achievement. For moderation by SES (Model 3), Robinson (2013) utilized ECLS-K:1999 data and similarly found a negative interaction between socioeconomic status and approaches to learning. The influence of children's approaches to learning on their math gains from fall to spring kindergarten was stronger for children from lower socioeconomic backgrounds (Robinson, 2013). As for our mediation results (Model 1), the examination of attentional focus as a mediator of approaches to learning on achievement was not previously examined, but more broadly attention is often conceptualized as one component of executive functioning (Brock et al., 2009). Children's executive functioning, as measured by their cognitive flexibility and working memory, influenced their reading and math achievement trajectories from kindergarten to first grade and was mediated by their approaches to learning (ECLS-K:2011 study; Sung & Wickrama, 2018).

4.3. Limitations of results

Our findings must be contextualized within the limitations of our study. Regarding the fit of our models, three of the four standalone fit indices suggested good to excellent fit of the models to the data. The SRMR, however, consistently indicated model misfit. It is possible our model was too restrictive as the SRMR may be more sensitive to mis-specified factor covariances and latent structures (Hu & Bentler, 1999). The longitudinal models included up to 15 latent variables, with many restrictions on the covariance structure of the indicators; additional relations should not be added to the model without support from theory. In addition, the misfit did not grow significantly worse from baseline when testing mediation and moderation. The purpose of our analyses was to demonstrate the different longitudinal methods rather than a substantiative focus on the findings, but an interpretation of our findings should consider this SRMR caveat.

Additionally, our models focused on interindividual (i.e., between-person) change only and did not examine intraindividual (i.e., within-person). An alternative to the cross-lagged panel model, referred to as the random intercept cross-lagged panel model, has been proposed to account for these multilevel effects (Hamaker et al., 2015). Random intercept models were attempted in our analyses but did not converge and could not be estimated. However, our cross-lagged panel models differed from their typical use. Researchers often use cross-lagged panel models to examine the influences of two or more variables on each other over time (Hamaker et al., 2015). Our models did not examine these reciprocal or bidirectional relations over time, rather we focused on one direction of influence (i.e., from approaches to learning to achievement). These model issues are beyond the scope of our article, but readers are referred to Hamaker et al. (2015), Lüdtke and Robitzsch (2021), Mulder and Hamaker (2021), and Zyphur et al. (2020) for useful references and to Ozkok et al. (2022) for an applied example of the random intercept-cross-lagged panel model for longitudinal moderation.

4.4. Conclusion and extensions

The purpose of this article was to discuss the advantages of longitudinal mediation, moderation, and integrated approaches to advance the field of school psychology research. These methods can be used to examine multivariate change and children's development throughout the school years and can also contribute to theoretical understanding and inform intervention and instruction. Longitudinal mediation, moderation, moderated mediation, and mediated moderation can explain how, why, and for whom the relation between two variables exists. In this article, we provided a conceptual overview of these issues and then demonstrated how these longitudinal methods may be applied with a large national sample (i.e., the public version of the ECLS-K:2011).

As described earlier, a variety of models can be used to answer questions related to longitudinal mediation, moderation, moderated mediation, and mediated moderation. Here we demonstrated these methods with cross-lagged panel models, but other models and extensions are possible and can answer different types of research questions. For researchers with clustered and non-independent data, multilevel modeling can be used to account for this nesting in longitudinal mediation and moderation models. If the clustering is ignored, standard errors may be underestimated and confidence intervals may be too narrow, which can lead to inflated Type I errors (Preacher, 2015). In an applied example, Benson (2020) used seven waves of data and combined multilevel modeling with longitudinal mediation and moderated mediation. The repeated measures were nested within mothers and their children's age was used as a Level 1 variable of time. Mothers' depressed mood, emotional well-being, and perceived support were tested as potential mediators of the relation between received support and marital quality. Baseline child behavior severity significantly moderated the link between mothers' received spousal support on perceived spousal support (mediator) on marital quality (Benson, 2020). For another useful applied example of multilevel longitudinal mediation, see Nie et al. (2020), and for a useful example of multilevel longitudinal moderation, see Teng et al. (2020). Preacher et al. (2016) provided technical information on moderation in multilevel SEM.

Latent growth models can also be used to test longitudinal mediation and moderation. Latent growth models address both intra-individual and interindividual change. Intraindividual change across time and interindividual variability in that intraindividual change are the focus (Little, 2013; Preacher et al., 2007). In an applied example, Sung and Wickrama (2018) analyzed four waves of the ECLS-K:2011 data from fall kindergarten through spring first grade and found children with higher initial levels and growth of executive functioning and approaches to learning demonstrated faster reading and math growth. Children's growth in approaches to learning mediated the influence of their growth in executive functioning on their growth in math and reading (Sung & Wickrama, 2018). Another possible extension is the use of categorical variables in mediation models. In our analyses, we examined a continuous mediator, which is frequently used in psychological research. Interested readers are directed to Preacher (2015) and VanderWeele (2016) for discussions of categorical variables in mediation models.

Finally, another extension concerns an approach to the interpretation of moderation effects. Often significant interactions are plotted to aid in their interpretation. An interaction with a categorical variable with a limited number of categories easily lends itself to be graphed by the continuous variable it interacts with. For interactions between two continuous variables researchers will often arbitrarily divide the continuous moderator into categories for illustration purposes. Often, individuals' scores will be divided into groups based on the mean (i.e., 1 *SD* below and 1 *SD* above the mean). Another lesser used approach is the Johnson-Neyman technique, which can also be applied to significant moderated mediation effects (Preacher et al., 2007). With the Johnson-Neyman technique, confidence bands are plotted around all values of the continuous moderator. Regions of significance for the moderator are determined in which the confidence band for the indirect effect does not contain zero (Preacher et al., 2007). For an illustration and M*plus* syntax, readers are directed to Clavel (2015) website.

In conclusion, third variable models can examine questions about how, why, and for whom the relation between two variables exists. The examination of these questions with longitudinal data improves the understanding of how these developmental processes unfold as children grow older. These questions are important to consider in school psychological research because the development of academic, social, emotional, and behavioral skills across childhood and adolescence is complex and can be influenced by many variables. Understanding how those variables are related and possibly influence each other is essential to building knowledge about development. Longitudinal mediation can help researchers and practitioners better understand which variables predict positive development in these areas and how those variables are related to each other. In cases where those third variables can be manipulated, intervention or prevention programs may be built around this knowledge. Similarly, longitudinal moderation can help clarify when and for whom specific relations among variables are important and the generalizability of the relation between two variables. If the success of specific intervention or prevention programs varies by group membership, this information can help researchers and practitioners consider why results vary for groups and whether those differences should result in changes to how programs or services are provided.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

The data file we used is shared here for ease of access. The ECLS-K:2011 data and related documentation are openly available at https://nces.ed.gov/ecls/kindergarten2011.asp.

Acknowledgements

This research was enhanced by training that Jacqueline Caemmerer and Briana Hennessy received during the American Educational Research Association Institute on Statistical Analysis: Development of Mathematics Competencies in Early Childhood. AERA receives funds for its "AERA Grants Program" from the National Science Foundation (NSF) under NSF award NSF-DRL #1749275. Opinions reflect those of the authors and do not necessarily reflect those AERA or NSF. The authors wish to thank Eunice Blemahdoo and Nivedhya Cheruvari for their literature search support and Natalie Charamut for her comments.

Appendix A. Appendix

M*plus* syntax for the five models examined in this study are provided. For laavan users we recommend the mplus2laavan converter package available at https://rdrr.io/cran/lavaan/man/mplus2lavaan.html. We are not laavan users but provide laavan syntax for the MODEL portion of the M*plus* syntax for the first three models to the best of our abilities. We also provided some resources for conducting Wald tests and latent variable interactions in laavan. Please be advised we did not extensively explore these models using laavan. We hope the laavan syntax will provide a useful starting point for more well-versed laavan users.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsp.2024.101283.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications, Inc.
Anthony, C. J., DiPerna, J. C., & Amato, P. R. (2014). Divorce, approaches to learning, and children's academic achievement: A longitudinal analysis of mediated and moderated effects. *Journal of School Psychology, 52*(3), 249–261. https://doi.org/10.1016/j.jsp.2014.03.003.
Anthony, C. J., & Ogg, J. (2019). Parent involvement, approaches to learning, and student achievement: Examining longitudinal mediation. *School Psychology, 34*(4), 376–385. https://doi.org/10.1037/spq0000282.
Anthony, C. J., & Ogg, J. (2020). Executive function, learning-related behaviors, and science growth from kindergarten to fourth grade. *Journal of Educational Psychology, 112*(8), 1563–1581. https://doi.org/10.1037/edu0000447.
Baltes, P. B., & Nesselroade, J. R. (1979). *Longitudinal research in the study of behavior and development*. Academic Press.
Barnard-Brak, L., Stevens, T., Xiao, F., & Chesnut, S. R. (2016). Approaches to learning and medicated ADHD: The potential impact on learning and assessment. *Learning and Individual Differences, 47*, 298–303. https://doi.org/10.1016/j.lindif.2015.12.029.
Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173.
Benson, P. R. (2020). Examining the links between received network support and marital quality among mothers of children with ASD: A longitudinal mediation analysis. *Journal of Autism and Developmental Disorders, 50*(3), 960–975. https://doi.org/10.1007/s10803-019-04330-4.
Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly, 24*(3), 337–349.
Caemmerer, J. M., & Keith, T. Z. (2015). Longitudinal reciprocal effects of social skills and achievement from kindergarten through eighth grade. *Journal of School Psychology, 53*(4), 265–281. https://doi.org/10.1016/j.jsp.2015.05.001.
Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. https://doi.org/10.1207/S15328007SEM0902.
Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth grade achievement: Evidence from the ECLS-K. *Economics of Education Review, 28*, 415–427. https://doi.org/10.1016/j.econedurev.2008.09.003.
Clavel, F. (2015, March 23). Advanced topics: Plotting better interactions using the Johnson-Neyman technique in Mplus. https://fredclavel.org/2015/03/23/advanced-topics-plotting-better-interactions-using-the-johnson-neyman-technique-in-mplus/.
Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112*(4), 558–577. https://doi.org/10.1037/0021-843X.112.4.558.
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428.
Enders, C. K. (2010). *Applied missing data analysis*. The Guilford Press.
Fairchild, A. J., & McQuillin, S. D. (2010). Evaluating mediation and moderation effects in school psychology: A presentation of methods and review of current practice. *Journal of School Psychology, 48*(1), 53–84. https://doi.org/10.1016/j.jsp.2009.09.001.
Gullo, D. F., & Impellizeri, W. E. (2022). Kindergarten teachers' ratings of children's behavioral and learning competencies: Predictive impact on children's third and fourth grade achievement trajectories. *Early Childhood Education Journal, 50*(2), 301–314. https://doi.org/10.1007/s10643-021-01157-6.
Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102–116. https://doi.org/10.1037/a0038889.
Hilley, C. D., & O'Rourke, H. P. (2022). Dynamic change meets mechanisms of change: Examining mediators in the latent change score framework. *International Journal of Behavioral Development, 46*(2), 125–141.
Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65*, 599–610. https://doi.org/10.1037/0022-006X.65.4.599.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
Jose, P. E. (2016). The merits of using longitudinal mediation. *Educational Psychologist, 51*, 331–341. https://doi.org/10.1080/00461520.2016.1207175.
Judd, C. M., McClelland, G. H., & Culhane, S. E. (1995). Data analysis: Continuing issues in the everyday analysis of psychological data. *Annual Review of Psychology, 46*, 433–465.
Karazsia, B. T., Berlin, K. S., Armstrong, B., Janicke, D. M., & Darling, K. E. (2014). Integrating mediation and moderation to advance theory development and testing. *Journal of Pediatric Psychology, 39*(2), 163–173. https://doi.org/10.1093/jpepsy/jst080.
Keith, T. Z. (2019). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling (3rd ed)*. Taylor & Francis.
Kuo, Y. L., Casillas, A., Allen, J., & Robbins, S. (2021). The moderating effects of psychosocial factors on achievement gains: A longitudinal study. *Journal of Educational Psychology, 113*(1), 138–156. https://doi.org/10.1037/edu0000471.
Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology, 46*(5), 1062–1077. https://doi.org/10.1037/a0020066.
Little, T. D. (1997). Mean and covariance structures (MACS) analysis of cross cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*(1), 53–76.
Little, T. D. (2013). *Longitudinal structural equation modeling. Methodology in the social sciences*. The Guilford Press.
Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., & Crandall, C. S. (2007). Structural equation modeling of mediation and moderation with contextual factors. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 207–230). Lawrence Erlbaum Associates, Inc.
Lüdtke, O., & Robitzsch, A. (2021). *A critique of the random intercept cross-lagged panel model* (pp. 1–25). https://doi.org/10.31234/osf.io/6f85c. PsyArXiv.
MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173–181.
MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83–104.
Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development, 39*(1), 87–96. https://doi.org/10.1177/0165025414552301.
Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research, 46*(5), 816–841. https://doi.org/10.1080/00273171.2011.606716.
McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analyses. In R. Cudeck, S. Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 342–380). Scientific Software International.

McClelland, G. H., Irwin, J. R., Disatnik, D., & Sivan, L. (2017). Multicollinearity is a red herring in the search for moderator variables: A guide to interpreting moderated multiple regression models and a critique of Iacobucci, Schneider, Popovich, and Bakamitsos (2016). *Behavior Research Methods, 49*(1), 394–402. https://doi.org/10.3758/s13428-016-0785-2.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. https://doi.org/10.1007/BF02294825.

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

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*(6), 852–863. https://doi.org/10.1037/0022-3514.89.6.852.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nie, Q., Yang, C., Teng, Z., Furlong, M. J., Pan, Y., Guo, C., & Zhang, D. (2020). Longitudinal association between school climate and depressive symptoms: The mediating role of psychological suzhi. *School Psychology, 35*(4), 267–276. https://doi.org/10.1037/spq0000374.

Ozkok, O., Vaulont, M. J., Zyphur, M. J., Zhang, Z., Preacher, K. J., Koval, P., & Zheng, Y. (2022). Interaction effects in cross-lagged panel models: SEM with latent interactions applied to work-family conflict, job satisfaction, and gender. *Organizational Research Methods, 25*(4), 673–715. https://doi.org/10.1177/10944281211043733.

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology, 60*, 65–82. https://doi.org/10.1016/j.jsp.2016.11.002.

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology, 66*, 825–852. https://doi.org/10.1146/annurev-psych-010814-015258.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879.

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods, 16*(2), 93–115. https://doi.org/10.1037/a0022658.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185–227. https://doi.org/10.1080/00273170701341316.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods, 21*(2), 189–205. https://doi.org/10.1037/met0000052.

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86*(1), 159–175.

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Reynolds, M. R., & Turek, J. J. (2012). A dynamic developmental link between verbal comprehension-knowledge (Gc) and reading comprehension: Verbal comprehension-knowledge drives positive change in reading comprehension. *Journal of School Psychology, 50*(6), 841–863.

Robinson, K. (2013). Early disparities in mathematics gains among poor and non-poor children: Examining the role of behavioral engagement in learning. *The Elementary School Journal, 114*(1), 22–47.

Rose, B. M., Holmbeck, G. N., Coakley, R. M., & Franks, E. A. (2004). Mediator and moderator effects in developmental and behavioral pediatric research. *Journal of Developmental & Behavioral Pediatrics, 25*, 58–67.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development, 6*(2–3), 144–164. https://doi.org/10.1080/15427600902911247.

Selig, J. P., Preacher, K. J., & Little, T. D. (2012). Modeling time-dependent association in longitudinal data: A lag as moderator approach. *Multivariate Behavioral Research, 47*(5), 697–716. https://doi.org/10.1080/00273171.2012.715557.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445. https://doi.org/10.1037/1082-989X.7.4.422.

Simonds, J., & Rothbart, M. K. (2004, October). The temperament in middle childhood questionnaire (TMCQ): A computerized self-report measure of temperament for ages 7–10. In *Poster session presented at the occasional temperament conference, Athens, GA*.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13*, 290–312.

Sung, J., & Wickrama, K. A. S. (2018). Longitudinal relationship between early academic achievement and executive function: Mediating role of approaches to learning. *Contemporary Educational Psychology, 54*, 171–183. https://doi.org/10.1016/j.cedpsych.2018.06.010.

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/sem tools packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 111–130. https://doi.org/10.1080/10705511.2019.1602776.

Teng, Z., Bear, G. G., Yang, C., Nie, Q., & Guo, C. (2020). Moral disengagement and bullying perpetration: A longitudinal study of the moderating effect of school climate. *School Psychology, 35*(1), 99–109. https://doi.org/10.1037/spq0000348.

Tourangeau, K., Nord, C., Le, T., Wallner-Allen, K., Vaden-Kiernan, N., Blaker, L., & Najarian, M. (2019). *Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011): User's manual for the ECLS-K:2011 kindergarten-fifth grade data file and electronic codebook, public version (NCES 2019–051)*. U.S. Department of Education National Center for Education Statistics.

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

VanderWeele, T. J. (2016). Mediation analysis: A practitioner's guide. *Annual Review of Public Health, 37*, 17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10–18.

Wu, A. D., & Zumbo, B. D. (2008). Understanding and using mediators and moderators. *Social Indicators Research, 87*(3), 367–392. https://doi.org/10.1007/s11205-007-9143-1.

Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., Hamaker, E. L., Shamsollahi, A., Pierides, D. C., Koval, P., & Diener, E. (2020). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods, 23*(4), 688–716. https://doi.org/10.1177/1094428119847280.