Proceedings of the 89th Annual International Meeting of the Psychometric Society, Prague, Czech Republic (2024).

Disentangling Groupwise Heterogeneity in the Mediating Relationships of Yearly Renewable Education Programs

Hanna Kim [*†] and Jee-Seon Kim [‡]

†Psychology Department, University of California, Santa Cruz, Santa Cruz, 95064, CA, United States
‡Department of Educational Psychology, University of Wisconsin-Madison, Madison, 53706, WI, United States
*Corresponding author. Email: hanna_kim@ucsc.edu

Abstract

Although recent advances in causal mediation analysis have clarified the necessary conditions for identifying direct and indirect effects using intuitive closed-form estimation, these effects are often assumed to remain constant between subpopulations. In the context of yearly renewable education programs, this assumption implies that the effects of initial program attendance - both its long-term effects and the mediated effects through subsequent program attendance - are deemed homogeneous, despite the presence of distinct subgroups who may benefit differently from such programs.

We introduce a set of regression-based causal mediation estimators that account for such groupwise heterogeneity not only in the main effects of the exposure and mediator on the mediator and the outcome, but also in the exposure-mediator interaction within the outcome model. A real data analysis illustrates the importance of examining groupwise heterogeneity in causal mediation analysis, highlighting markedly different effects of the early and regular Head Start programs on English receptive vocabulary skills among Dual Language Learners and primarily English-speaking children. We conclude by discussing how extensions of causal mediation analysis methods can advance the study of heterogeneity in developmental processes and outline directions for future research.

Keywords: causal mediation analysis, effect heterogeneity, treatment-mediator interaction, regression-based estimation

1. Introduction

Over the past few decades, various causal mediation analysis methods have illuminated the conditions necessary for drawing causal conclusions about psychological processes (Imai et al., 2010; Nguyen et al., 2021; Pearl, 2001; VanderWeele, 2015). A particularly promising context for applying these causal mediation analysis methods is the evaluation of yearly renewable education programs. Rather than treating participation in the initial and subsequent rounds of such programs as separate events, conceptualizing them as the exposure and mediator within a mediation analysis framework allows researchers to examine how subsequent program attendance transmits the initial program's effect to a later outcome (H. Kim & Kim, 2024). In this framework, the average treatment effect (ATE) of the initial program attendance can be decomposed into the natural indirect effect (NIE) transmitted through subsequent program attendance and the natural direct effect (NDE) of the initial program attendance.

Among the various methods developed to estimate the NDE and NIE, closed-form point estimators based on mediation and outcome regression models (Li, Mathur, et al., 2023; Valeri &

VanderWeele, 2013) are particularly useful. They allow NDE and NIE to be expressed as contrasts of potential outcomes within the causal mediation framework while preserving the intuitive appeal of interpreting mediation effects as combinations of regression coefficients. However, the validity of these regression-based parametric estimators depends on the correct specification of both the mediator and outcome models. In addition to accounting for all variables that may confound the relationships between the exposure, mediator, and outcome, the models must be specified in functional forms that closely resemble the underlying data generating processes.

One common source of model misspecification arises when group differences in responses to the exposure – either in terms of the mediator or the outcome – are not modeled. In the context of yearly renewable education programs, this implies that the long-term effects of initial program attendance, as well as the mediated effects through subsequent participation, are incorrectly assumed to be homogeneous across subgroups.

A notable example of a renewable education program is the national Head Start program, which is administered over two consecutive years as Early and Regular Head Start, serving children from ages three to four to promote school readiness (Puma et al., 2010). Although Head Start is broadly accessible to low-income children and children with disabilities, studies have documented varying improvements in school readiness measures across subgroups of children from diverse backgrounds (Lee et al., 2021; Puma et al., 2010; Xie et al., 2020). These findings highlight the diversity of the population served by Head Start and underscore the need for a more nuanced understanding of the potentially heterogeneous effects of Early and Regular Head Start attendance.

Among the various measures of school readiness, we focus on English receptive vocabulary skills – children's ability to comprehend spoken English words – as our outcome of interest, given its foundational role in supporting communication and learning in the American school system. Extensive research shows that Dual Language Learners acquire English skills in ways that qualitatively differ from those of monolingual children from English-speaking households during early childhood (e.g., Castro, 2014; Choi et al., 2023). This motivates our focus on the home language environment as a key grouping variable for investigating heterogeneity in the mediating effect of Regular Head Start attendance on the impact of Early Head Start participation. Specifically, we examine whether the development of English receptive vocabulary skills through Early and Regular Head Start differs between primarily English-speaking children and Dual Language Learners.

To address groupwise heterogeneity in the mechanisms through which a renewable education program exerts its effect, we extend the parametric closed-form estimator for the NDE and NIE by incorporating a predefined grouping variable as a moderator in both the mediator and outcome models. This includes not only the main effects of the exposure and mediator, but also their interaction term in the outcome model. This extension allows researchers to answer the following research questions:

- What is the effect of initial program attendance within each predefined group?
- How is this effect decomposed into long-term effects of initial program attendance and mediated effects through subsequent program attendance within each group?
- How does the long-term direct effect of initial program attendance differ between groups?
- How does the mediated effect of initial program attendance through subsequent program attendance differ between groupos?

2. Decomposing the Total Effect into Groupwise Natural Direct and Indirect Effects

Within the causal mediation framework, the average effect of an exposure can be decomposed into natural direct and indirect effects (NDE and NIE) with respect to a mediator. To capture group-specific variation in responses to the exposure and mediator, we focus on groupwise NDE and NIE as the estimands of interest, rather than the marginal NDE and NIE that average across all groups. This approach allows us to investigate, for each group, the total effect of initial program attendance and understand how it unfolds through subsequent program attendance. Specifically, we compare across groups: (1) the initial program's effect that is not attributable to subsequent program participation (groupwise NDE), and (2) the initial program's effect that operates through changes in subsequent program participation following initial exposure (groupwise NIE).

Formally, let $M(a)_i$ denote the potential value of the mediator for subject i when their exposure is set to $A_i = a$, where $a \in \{0, 1\}$. When both the exposure and mediator are binary as in the case of participation in renewable education programs, we define the nested potential outcomes as $Y(A_i = a, M_i = M(a'))_i = Y(a, M(a'))_i$. The values a and a' coincide in naturally occurring cases, such as $Y(0, M(0))_i$ or $Y(1, M(1))_i$, where the mediator status of a subject follows their initial exposure. They differ, however, in counterfactual scenarios like $Y(0, M(1))_i$ or $Y(1, M(0))_i$, which allow us to conceptualize alternative mediation pathways.[1]

To allow for the possibility that the NDE and NIE vary across predefined subject groups, we define groupwise NDE and NIE as functions of group membership $G_i = g$, where $g \in \{1, ..., k\}$, in addition to the exposure level $A_i = a$ with $a \in \{0, 1\}$ and a vector of covariates $\mathbf{C}_i = \mathbf{c}$:

$$NDE(a|g, \mathbf{C}_i = \mathbf{c}) = E[Y(1, M(a))_i - Y(0, M(a))_i | G_i = g, \mathbf{C}_i = \mathbf{c}], \tag{1}$$

$$NIE(a|g, \mathbf{C}_i = \mathbf{c}) = E[Y(a, M(1))_i - Y(a, M(0))_i | G_i = g, \mathbf{C}_i = \mathbf{c}]. \tag{2}$$

In the presence of groupwise heterogeneity, the total effect of the exposure - also known as the average treatment effect (ATE) - may also vary across groups. For each group, the total effect can be decomposed into combinations of groupwise NDE and NIE, conditional on covariates $\mathbf{C}_i = \mathbf{c}$. Notably, there are two possible decompositions depending on the levels of hypothetical exposure (i.e., the values of a and a') used in defining the nested counterfactual outcomes. When an interaction exists between the exposure and mediator on the outcome, these decompositions yield different estimates of the direct and indirect effects.

$$\begin{aligned} TE|(g, \mathbf{C}_i = \mathbf{c}) &= E[Y(1, M(1))_i - Y(0, M(0))_i | G_i = g, \mathbf{C}_i = \mathbf{c}] \\ &= E[Y(1, M(1))_i - Y(0, M(1))_i | G_i = g, \mathbf{C}_i = \mathbf{c}] \\ &\quad + E[Y(0, M(1))_i - Y(0, M(0))_i | G_i = g, \mathbf{C}_i = \mathbf{c}] \\ &= NDE(1|g, \mathbf{C}_i = \mathbf{c}) + NIE(0|g, \mathbf{C}_i = \mathbf{c}). \end{aligned} \tag{3}$$

$$\begin{aligned} TE|(g, \mathbf{C}_i = \mathbf{c}) &= E[Y(1, M(1))_i - Y(0, M(0))_i | G_i = g, \mathbf{C}_i = \mathbf{c}] \\ &= E[Y(1, M(1))_i - Y(1, M(0))_i | G_i = g, \mathbf{C}_i = \mathbf{c}] \\ &\quad + E[Y(1, M(0))_i - Y(0, M(0))_i | G_i = g, \mathbf{C}_i = \mathbf{c}] \\ &= NIE(1|g, \mathbf{C}_i = \mathbf{c}) + NDE(0|g, \mathbf{C}_i = \mathbf{c}). \end{aligned} \tag{4}$$

3. Parametric Closed-Form Estimation of Groupwise Natural Direct and Indirect Effects

To consistently estimate $NDE(a|g, \mathbf{C}_i = \mathbf{c})$ and $NIE(a|g, \mathbf{C}_i = \mathbf{c})$ from observed data, despite their definition as functions of nested potential outcomes, the measured covariates $\mathbf{C}_i$ must include all relevant confounders and satisfy the following identification assumptions. For $a, a', a^* \in \{0, 1\}$, where $a \neq a^*$, the nonparametric identification assumptions are:

$$Y(a, M(a'))_i \perp A_i | \{G_i, \mathbf{C}_i\}, \tag{5}$$

$$Y(a, M(a'))_i \perp M_i | \{A_i, G_i, \mathbf{C}_i\}, \tag{6}$$

$$M(a)_i \perp A_i | \{G_i, \mathbf{C}_i\}, \tag{7}$$

$$Y(a, M(a'))_i \perp M(a^*) | \{G_i, \mathbf{C}_i\}. \tag{8}$$

In this framework, exposure levels are statistically independent of both potential outcomes and potential mediators within each group, conditional on observed confounders (Equations 5 and 7). Similarly, mediator levels and potential mediators are independent of the potential outcomes within each group, given observed confounders (Equations 6 and 8).

Once groupwise NDE and NIE are deemed identifiable from observed data, researchers can apply a variety of estimation methods.[2] Closed-form estimators based on mediator and outcome regression models are appealing for their interpretability and ease of communication (Bollen & Pearl, 2013; Li, Yoshida, et al., 2023), but they require additional assumptions for valid inference. Specifically, the regression models should correctly represent the relationships among the exposure, mediator, outcome, and covariates. When there is a reason to believe that the effects of the exposure on the mediator and outcome, or the effects of the mediator on the outcome, vary across subject groups, we propose estimating groupwise NDE and NIE using mediator and outcome regression models that include groupwise interaction terms. Specifically, we incorporate interactions between group membership and the exposure, the mediator, and the exposure-mediator interaction related to the outcome.[3]

Formally, let the grouping variable $\mathbf{G_i}$ be represented by a vector of $k-1$ binary indicators (G_{2i}, ..., G_{ki}), where $G_{ji} = 1$ if subject i belongs to group j, and 0 otherwise, for $j = 2, \ldots, k$. The first group serves as the reference category and is represented by all zeros in $\mathbf{G_i}$. The full set of covariates is denoted by $\mathbf{C}_i$, with subsets $\mathbf{c_m}$ and $\mathbf{c_y}$ used in the mediator and outcome models, respectively. These subsets may or may not be identical.

A logistic regression model can be specified for a binary mediator, followed by a linear regression model for a continuous outcome. In this article, we focus on cases involving binary exposures and mediators followed by continuous outcomes, although similar derivation procedures may extend to cases involving other distributions of the exposure, mediator, or outcome.

$$logit\{P(M_i = 1 | A_i = a, \mathbf{G_i} = \mathbf{g}, \mathbf{C}_i = \mathbf{c_m})\} = \beta_0 + \beta_1 a + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m} + \boldsymbol{\beta}_4' a \times \mathbf{g}, \tag{9}$$

$$\begin{aligned} E(Y_i | A_i = a, M_i = m, \mathbf{G_i} = \mathbf{g}, \mathbf{C}_i = \mathbf{c_y}) = \theta_0 + \theta_1 a + \theta_2 m + \boldsymbol{\theta}_3'\mathbf{g} + \boldsymbol{\theta}_4'\mathbf{c_y} + \\ \theta_5 a \times m + \boldsymbol{\theta}_6' a \times \mathbf{g} + \boldsymbol{\theta}_7' m \times \mathbf{g} + \boldsymbol{\theta}_8' a \times m \times \mathbf{g}. \end{aligned} \tag{10}$$

These models allow for the possibility that group membership moderates the effect of (1) the exposure on the mediator ($\boldsymbol{\beta}_4'$), (2) the exposure on the outcome ($\boldsymbol{\theta}_6'$), (3) the mediator on the outcome ($\boldsymbol{\theta}_7'$), and (4) the interaction between the exposure and mediator on the outcome ($\boldsymbol{\theta}_8'$). Incorporating these groupwise interaction terms enables the estimation of groupwise heterogeneity in mediation effects that arise from differences in how groups respond to the exposure, the mediator, or their interaction. For simplicity, interaction terms between $\mathbf{C}_i$ and $\mathbf{G_i}$ are not shown, but they may be included as needed without substantially altering the estimator derivation. Regression parameters involving the grouping variable are expressed as vectors to generalize to settings where the grouping variable has more than two levels and is represented by multiple indicators.

Based on Equations 9 and 10, we derive the point estimators for groupwise NDE and NIE as follows.

$$NDE(a|g, \mathbf{C}_i = \mathbf{c}) = (\theta_1 + \boldsymbol{\theta}_6'\mathbf{g}) + (\theta_5 + \boldsymbol{\theta}_8'\mathbf{g})\frac{exp[\beta_0 + \beta_1 a + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m} + \boldsymbol{\beta}_4' a \times \mathbf{g}]}{1 + exp[\beta_0 + \beta_1 a + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m} + \boldsymbol{\beta}_4' a \times \mathbf{g}]}, \tag{11}$$

$$\begin{aligned} &NIE(a|g, \mathbf{C}_i = \mathbf{c}) \\ &= (\theta_2 + \theta_5 a + \boldsymbol{\theta}_7'\mathbf{g} + \boldsymbol{\theta}_8' a \times \mathbf{g})\{\frac{exp[\beta_0 + \beta_1 + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m} + \boldsymbol{\beta}_4'\mathbf{g}]}{1 + exp[\beta_0 + \beta_1 + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m} + \boldsymbol{\beta}_4'\mathbf{g}]} - \frac{exp[\beta_0 + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m}]}{1 + exp[\beta_0 + \boldsymbol{\beta}_2'\mathbf{g} + \boldsymbol{\beta}_3'\mathbf{c_m}]}\}. \end{aligned} \tag{12}$$

The groupwise NDE and NIE estimators depend on group membership, represented by the column vector $\mathbf{g}$, and reflect variability in the exposure and mediator effects across subject groups. Because the large-sample properties of these estimators have not been formally established and therefore normality cannot be assumed, standard errors are obtained using nonparametric bootstrapping.

Several characteristics of these estimators are worth noting. First, the groupwise NDE and NIE estimators represent a general form that incorporates three-way interaction terms $\boldsymbol{\theta}_8'$, capturing groupwise heterogeneity not only in the main effects of the exposure and mediator, but also in their interaction on the outcome. This formulation allows researchers to compare mediation effects across multiple groups simultaneously, without conducting separate mediation analyses for each group.

In practice, however, not all forms of groupwise heterogeneity may be present. Researchers may choose to exclude certain interaction terms to simplify the estimation models and improve parsimony. When there is no effect heterogeneity across subject groups - that is, when $\boldsymbol{\beta}_4' = \boldsymbol{\theta}_6' = \boldsymbol{\theta}_7' = \boldsymbol{\theta}_8' = \mathbf{0}'$ - the estimators reduce to the closed-form estimators for NDE and NIE proposed by Valeri and VanderWeele (2013).

Second, covariates in the mediator and outcome models are included as main effects, thereby helping to satisfy the ignorability assumptions in 5 to 8 through statistical adjustment. Although the main effects of covariates in the outcome model ($\mathbf{C}_i = \mathbf{c_y}$) cancel out when deriving the point estimators, their inclusion reduces bias and improves precision of other partial regression coefficients, such as θ_5. The covariates used in the mediator and outcome models need not be identical; the researcher may specify different subsets of $\mathbf{C}_i$, such that $\mathbf{c_m} \neq \mathbf{c_y}$.

When estimating $NDE(a|g, \mathbf{C}_i = \mathbf{c})$ and $NIE(a|g, \mathbf{C}_i = \mathbf{c})$, the covariate values $\mathbf{c}$ refer specifically to those included in the mediator model. Selecting these values determines the individuals for whom the groupwise NDE and NIE are estimated. For example, to obtain groupwise NDE and NIE estimates for an average individual within group $G_i = g$, $\mathbf{c}$ can be set to $\bar{\mathbf{c}}_{\mathbf{m}|g}$, the mean vector of the covariates $\mathbf{c_m}$ in the mediator model (Equation 9).

4. An Empirical Analysis of the Effects of Early and Regular Head Start Participation by Children's Home Language Environment

4.1 Background and Data

To illustrate the methods discussed in the previous section, we examined the effects of Early and Regular Head Start attendance on children's English receptive vocabulary skills using data from the 3-year-old cohort of the Head Start Impact Study (HSIS).[4] HSIS is a nationally representative study designed to evaluate the effectiveness of Head Start as a national early childhood education program by following children and families eligible for enrollment in oversubscribed Head Start programs (Puma et al., 2010). The 3-year-old cohort includes 2,449 children (50% Female, 36% Black, 33% Hispanic), approximately 60% of whom attended Early Head Start.

As discussed in the introduction, we focused on children's home language environment by categorizing participants as either primarily English-speaking or Dual Language Learners, in order to examine groupwise heterogeneity in the effects of Early Head Start attendance on English receptive vocabulary skills, as mediated by Regular Head Start attendance. Building on prior research documenting differences in language development experiences between Dual Language Learners and English-monolingual children in Head Start settings (Garcia, 2018; Hammer et al., 2014; Piker & Rex, 2008), we investigated whether the NDE and NIE of Early Head Start attendance on children's English receptive vocabulary skills, through Regular Head Start attendance, differed across these groups.

Specifically, participation in Early Head Start and Regular Head Start was treated as the exposure and mediator, respectively, while children's English receptive vocabulary was measured using a shortened version of the Peabody Picture Vocabulary Test (PPVT) and treated as the outcome. Children's primary language at home, recorded as English and Spanish, was used as the grouping variable. Following previous Head Start studies, we included a range of potential confounders: children's demographic characteristics (gender, race/ethnicity, and special needs status), caregiving conditions (primary caregiver's age, co-residence with both biological parents, primary caregiver's level of depression, recent immigration of biological mother, mothers' marital status and education level, household risk level, and urbanicity), and children's receptive vocabulary skills measured at the baseline (Fall of the Early Head Start year) (Jenkins et al., 2018; H. Kim & Kim, 2024; Puma et al., 2010).

However, we were unable to empirically rule out two key sources of bias. First, the covariates included in our analysis did not account for post-exposure variables that, while independent of Early Head Start attendance, may have influenced both Regular Head Start attendance and children's PPVT scores. For example, access to alternative pre-K programs in the neighborhood could confound the mediator-outcome relationship.

Second, the cross-world ignorability assumption (Equation 8) is inherently untestable and may be violated in this context. Our analysis assumed that no consequence of participating in Early Head Start were causally related to both Regular Head Start attendance and later receptive vocabulary performance. However, this assumption may not hold, as caregivers' perception of their child's development following Early Head Start attendance, for instance, could influence their decision to continue Head Start enrollment, as well as the child's subsequent PPVT performance. The extent to which these potential sources of bias affect the results should be explored using additional analytic strategies, which falls beyond the scope of the current article.

Descriptive statistics for the variables used in our analyses are presented in Table 1. Multiple imputation was performed using the *mice* package (van Buuren & Groothuis-Oudshoorn, 2011) in R (v4.4.1; R Core Team, 2024) with predictive mean matching as the imputation algorithm. For simplicity, we used the first of 20 imputed datasets for our analysis, as the variable distributions were comparable across all imputed datasets.[5]

4.2 Analysis Results

The mediator and outcome models were fitted to the HSIS data following Equations 9 and 10. The resulting model coefficient estimates are presented in Tables 2 and 3, alongside estimates from models that assume no groupwise heterogeneity for comparison. Using the estimated model coefficients, we computed point estimates of the TE, NDE, and NIE separately for English-speaking children and Dual Language Learners by applying Equations 3, 11, and 12.[6] Standard errors and 95% confidence intervals for the point estimates were obtained via nonparametric bootstrapping with 10,000 replications. The estimated groupwise TE, NDE, and NIE estimates are presented in Table 4, followed by corresponding estimates under the assumption of constant mediation effects across groups in Table 5.[7]

The total effect of Early Head Start attendance and its decomposition into NDE and NIE differed substantially between English-speaking children and Dual Language Learners, assuming that all relevant confounders were appropriately adjusted for. Among English-speaking children ($G_i = 1$) with background characteristics similar to the average English-speaking Head Start-eligible child, Early Head Start attendance had a positive and statistically significant effect on receptive vocabulary development ($\hat{TE}|(1, \mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}|1}) = 5.991$, 95%$CI$ = [1.290, 10.733]).

This total effect was decomposed into two sets of NDE and NIE estimates, depending on the hypothetically fixed level of Regular Head Start attendance probability - $M(0)$ versus $M(1)$ for NDE - and Early Head Start attendance status - $A = 1$ versus $A = 0$ for NIE. Among English-speaking children, those who attended Early Head Start were approximately 3.615 times more likely to attend Regular Head Start than their peers who did not, holding other characteristics constant ($exp(\hat{\beta}_1) = exp(1.285) = 3.615$, see Table 2). Given this difference in attendance probabilities, the positive estimate of $NDE(0|1, \mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}|1})$ suggests that Early Head Start attendance is beneficial through pathways other than its effect on increasing subsequent Regular Head Start attendance.

However, the estimated NDE of Early Head Start attendance when accounting for the increased likelihood of Regular Head Start attendance ($\hat{NDE}(1|1, \mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}|1}) = 4.365$, 95%$CI$ = [−0.950, 9.735]) was not statistically significant. Moreover, the mediated effect of Early Head Start through Regular Head Start attendance was not statistically significant, regardless of whether children had previously attended Early Head Start. In other words, attending Early Head Start programs did not lead to a statistically significant development of English receptive vocabulary skills through Regular Head Start attendance for average English-speaking Head Start-eligible children.

In contrast, Dual Language Learners ($G_i = 0$) with background characteristics comparable to the average Head Start-eligible Dual Language Learner did not experience a statistically significant overall improvement in English receptive vocabulary skills from attending Early Head Start. While the estimated $NDE(0|0, \mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}|0})$ was positive (3.394, 95%CI = [0.633, 6.167]), this effect was offset by a negative $NIE(1|0, \mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}|0})$ estimate (−1.275, 95%CI = [−2.451, −0.140]), which reflects the impact of Early Head Start attendance on receptive English vocabulary development via changes in Regular Head Start attendance. These findings suggest that Early Head Start attendance, after leading to an increased likelihood of attending Regular Head Start programs ($exp(\hat{\beta}_1 + \hat{\beta}_4) = exp(1.285 + 0.141) = 4.162$, see Table 2), did not help Dual Language Learners learn more English receptive vocabulary skills.

Our mediation analysis approach depends on the correct specification of both the mediator and outcome models. Here, we focus on the role of groupwise heterogeneity in model specification and demonstrate how failing to account for heterogeneity between home language environment groups can affect the estimated regression coefficients of the mediator and outcome models. As shown in Tables 2 and 3, omitting groupwise interaction terms and modeling only the main effects of the exposure, mediator, and grouping variable yields coefficient estimates that differ from those obtained when groupwise heterogeneity is explicitly modeled. Importantly, the coefficient estimates from models that assume no groupwise heterogeneity are not simple arithmetic means of group-specific coefficient estimates. Rather, they are weighted averages, where the weights depend on conditional variances, and thus may not correspond to any theoretically relevant parameter of interest (Elwert & Winship, 2010).

These differences in regression coefficient estimates feed directly into the mediation analysis results presented in Table 5, potentially leading to oversimplified or misleading conclusions when groupwise heterogeneity is ignored. For the average child eligible for Head Start across all home language environments, the total effect of Early Head Start attendance on English receptive vocabulary skills at the end of the pre-K year was positive and statistically significant ($\hat{TE}|(\mathbf{C_i} = \bar{\mathbf{c}}_{\mathbf{m}}) = 3.154$, 95%$CI$ = [0.873, 5.400]). When this total effect was decomposed, the benefits of Early Head Start appeared to operate primarily through a prolonged direct effect, rather than through increased participation in Regular Head Start. Specifically, both NDE estimates were positive and significant ($\hat{NDE}(0|\mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}}) = 4.424$, 95%$CI$ = [2.001, 6.878], $\hat{NDE}(1|\mathbf{C}_i = \bar{\mathbf{c}}_{\mathbf{m}}) = 3.023$, 95%$CI$ = [0.500, 5.534]). In contrast, the negative $NIE(1|\mathbf{C_i} = \bar{\mathbf{c}}_\mathbf{m})$ estimate suggests that, for an average Head Start-eligible child, the increased likelihood of attending Regular Head Start after attending Early Head Start may actually attenuate the overall benefit of Early Head Start on receptive vocabulary development, provided that all causal assumptions hold.

However, interpreting these findings as applying uniformly to all Head Start-eligible children regardless of their home language backgrounds risks underestimating the effectiveness of Early and Regular Head Start programs for English-speaking children and overestimating the program effects for Dual Language Learners. By allowing for the possibility that Early Head Start effects may be transmitted differently through Regular Head Start attendance by children's home language environment, we can better understand why English-speaking children may derive greater benefit from Early Head Start attendance, whereas Dual Language Learners may not, particularly compared to alternative childcare options available in their communities.

Overall, this empirical analysis underscores the importance of examining groupwise heterogeneity when assessing the plausibility of correct functional form assumptions underlying parametric estimation methods for causal mediation analysis. Neglecting such heterogeneity can obscure meaningful subgroup differences and lead to misleading conclusions about the effectiveness of yearly renewable education programs.

5. Discussion

This article demonstrated how to account for groupwise heterogeneity in the mediation effects of yearly renewable education programs by extending regression-based closed-form estimators of the NDE and NIE to allow the exposure, mediator, and their interaction to vary across predefined groups. By adopting this approach, researchers can detect and interpret heterogeneity in how intervention effects are transmitted through mediating pathways across subpopulations of interest. The accompanying empirical analysis illustrated how this extension can deepen our understanding of effect heterogeneity in both the mechanisms and overall effectiveness of yearly renewable education programs.

Before concluding, several avenues for future research merit attention. First, the asymptotic properties of the proposed groupwise NDE and NIE estimators should be systematically evaluated through simulation studies. While this article has focused primarily on demonstrating the importance of incorporating predefined groups as moderators in the mediator and outcome models under the assumption that nonparametric identification assumptions are met, the consequences of unmeasured confounding, particularly in interaction with group membership, warrants further investigation.

Second, the proposed groupwise NDE and NIE estimation approach should be further extended to accommodate features commonly encountered in empirical evaluations of education programs. For example, substantial clustering may arise when programs are administered at the classroom, school, or center level, necessitating the use of multilevel modeling or cluster-robust inference techniques. Additionally, when outcomes or confounders involve psychological constructs, measurement error becomes a critical concern. In such cases, it is important to incorporate methods that address measurement reliability and, where appropriate, to conduct sensitivity analyses to assess the robustness of the mediation findings.

Lastly, heterogeneity in mediation effects may not always align most clearly with predefined, manifest groups. Instead, there may be unobserved subpopulations differentiated by patterns of multiple background characteristics that exhibit distinct mediation effects. In such cases, a manifest grouping variable may serve only as a proxy for more complex latent structures. To better capture this context, causal mediation analysis could be augmented with classification techniques such as mixture modeling (C. Kim et al., 2018; Wang et al., 2021), enabling the identification of latent subgroups with distinct mediating mechanisms. Integrating these approaches may offer a more comprehensive understanding of effect heterogeneity in how education programs exert their influence.

Table 1. Combined Descriptive Statistics of the HSIS Data Before and After Imputing Missing Information

	Observed data set					Analysis data set[a]				
Variable	**N**	**Mean**	**SD**	**min**	**max**	**N**	**Mean**	**SD**	**min**	**max**
Early Head Start	2113	0.59	0.49	0.00	1.00	2449	0.56	0.50	0.00	1.00
Regular Head Start	2030	0.59	0.49	0.00	1.00	2449	0.59	0.49	0.00	1.00
Receptive vocabulary at baseline	2398	230.40	37.29	128.54	372.39	2449	230.13	37.15	128.54	372.39
Receptive vocabulary after Pre-K	2041	299.49	38.53	172.11	407.40	2449	299.74	38.39	172.11	407.40
Black	2449	0.36	0.48	0.00	1.00	2449	0.36	0.48	0.00	1.00
Hispanic	2449	0.33	0.47	0.00	1.00	2449	0.33	0.47	0.00	1.00
Female	2449	0.50	0.50	0.00	1.00	2449	0.50	0.50	0.00	1.00
Low academic skills at baseline	2449	0.24	0.43	0.00	1.00	2449	0.24	0.43	0.00	1.00
Primarily speaks English at home	2449	0.26	0.44	0.00	1.00	2449	0.26	0.44	0.00	1.00
Special educational needs	2449	0.12	0.32	0.00	1.00	2449	0.12	0.32	0.00	1.00
Caregiver age	2449	28.76	7.48	16.00	78.00	2449	28.76	7.48	16.00	78.00
Lives with both biological parents	2449	0.49	0.50	0.00	1.00	2449	0.49	0.50	0.00	1.00
Caregiver depression scale score	1991	0.77	0.98	0.00	3.00	2449	0.77	0.98	0.00	3.00
Mother recently immigrated	2449	0.16	0.37	0.00	1.00	2449	0.16	0.37	0.00	1.00
Mother less than High School	2449	0.35	0.48	0.00	1.00	2449	0.35	0.48	0.00	1.00
Mother High School equivalent	2449	0.35	0.48	0.00	1.00	2449	0.35	0.48	0.00	1.00
Mother never married	2445	0.42	0.49	0.00	1.00	2445	0.42	0.49	0.00	1.00
Mother currently married	2445	0.43	0.50	0.00	1.00	2445	0.43	0.50	0.00	1.00
Household Risk Index score	2449	0.32	0.61	0.00	2.00	2449	0.32	0.61	0.00	2.00
Urban	2449	0.83	0.38	0.00	1.00	2449	0.83	0.38	0.00	1.00

a The first out of the 20 imputed data sets.

Table 2. Mediation Regression Model Coefficients Excluding and Including Primary Language Group Heterogeneity

	Without groupwise heterogeneity			With groupwise heterogeneity		
Variable	$\hat{\beta}$	**SE**	p	$\hat{\beta}$	**SE**	p
Intercept	−0.343	0.429	0.423	−0.328	0.429	0.444
Early Head Start	1.321	0.089	<0.001	1.285	0.102	<0.001
Primarily speaks English at home	0.264	0.164	0.107	0.200	0.189	0.289
Receptive vocabulary at baseline	−0.001	0.001	0.327	−0.001	0.001	0.328
Black	−0.385	0.118	0.001	−0.382	0.118	0.001
Hispanic	−0.075	0.136	0.583	−0.075	0.136	0.580
Female	−0.005	0.089	0.953	−0.004	0.089	0.965
Low academic skills at baseline	0.336	0.109	0.002	0.335	0.109	0.002
Special educational needs	0.197	0.141	0.162	0.198	0.141	0.160
Urban	0.161	0.118	0.175	0.161	0.118	0.175
Caregiver age	−0.003	0.006	0.590	−0.003	0.006	0.595
Lives with both biological parents	−0.019	0.120	0.872	−0.015	0.120	0.903
Caregiver depression scale score	0.094	0.046	0.042	0.093	0.046	0.044
Mother recently immigrated	−0.217	0.159	0.171	−0.224	0.160	0.162
Mother less than High School	0.199	0.117	0.090	0.197	0.117	0.092
Mother High School equivalent	0.072	0.108	0.504	0.071	0.108	0.509
Mother never married	0.218	0.138	0.115	0.219	0.138	0.114
Mother currently married	0.150	0.157	0.340	0.147	0.157	0.350
Household Risk Index score	−0.063	0.084	0.453	−0.062	0.084	0.458
Early Head Start × Primarily speaks English at home	–	–	–	0.141	0.203	0.487

Acknowledgement

We are grateful to proceedings editor Irini Moustaki for her thoughtful and valuable feedback on

Table 3. Outcome Regression Model Coefficients Excluding and Including Primary Language Group Heterogeneity

Variable	Without groupwise heterogeneity			With groupwise heterogeneity		
	$\hat{\beta}$	SE	p	$\hat{\beta}$	SE	p
Intercept	237.442	5.619	<0.001	238.575	5.639	<0.001
Early Head Start	6.298	1.825	0.001	4.608	2.073	0.026
Regular Head Start	0.421	1.735	0.808	-1.123	2.025	0.579
Primarily speaks English at home	-21.832	2.095	<0.001	-26.904	3.076	<0.001
Receptive vocabulary at baseline	0.342	0.017	<0.001	0.342	0.017	<0.001
Black	-20.995	1.537	<0.001	-21.007	1.537	<0.001
Hispanic	-14.874	1.754	<0.001	-15.053	1.756	<0.001
Female	1.581	1.150	0.169	1.621	1.150	0.159
Low academic skills at baseline	-11.881	1.388	<0.001	-11.924	1.387	<0.001
Special educational needs	-2.644	1.772	0.136	-2.631	1.772	0.138
Urban	2.725	1.544	0.078	2.771	1.546	0.073
Caregiver age	0.216	0.078	0.006	0.220	0.078	0.005
Lives with both biological parents	-0.210	1.554	0.892	-0.080	1.555	0.959
Caregiver depression scale score	1.368	0.595	0.022	1.351	0.595	0.023
Mother recently immigrated	-11.460	2.016	<0.001	-11.433	2.019	<0.001
Mother less than High School	-9.677	1.516	<0.001	-9.723	1.517	<0.001
Mother High School equivalent	-4.485	1.411	0.001	-4.487	1.410	0.001
Mother never married	0.594	1.804	0.742	0.658	1.803	0.715
Mother currently married	0.473	2.046	0.817	0.422	2.046	0.836
Household Risk Index score	-2.695	1.088	0.013	-2.633	1.088	0.016
Early Head Start × Regular Head Start	-4.501	2.417	0.063	-3.033	2.796	0.278
Early Head Start × Primarily speaks English at home	-	-	-	7.156	4.382	0.103
Regular Head Start × Primarily speaks English at home	-	-	-	6.216	3.917	0.113
Early Head Start × Regular Head Start × Primarily speaks English at home	-	-	-	-6.433	5.614	0.252

Table 4. Groupwise Natural Direct and Indirect Effects of Early and Regular Head Start Attendance on English Receptive Vocabulary Skills[a]

Variable	English-Speaking Children				Dual Language Learners			
	Estimate	Bootstrap SE	95% CIL	95% CIU	Estimate	Bootstrap SE	95% CIL	95% CIU
$TE\|(g, \mathbf{C}_i = \bar{\mathbf{c}}_{m\|g})$	5.991	2.395	1.290	10.733	2.119	1.315	-0.430	4.716
$NDE(0\|g, \mathbf{C}_i = \bar{\mathbf{c}}_{m\|g})$	7.388	2.730	2.093	12.773	3.394	1.421	0.633	6.167
$NIE(1\|g, \mathbf{C}_i = \bar{\mathbf{c}}_{m\|g})$	-1.397	1.199	-3.863	0.908	-1.275	0.592	-2.451	-0.140
$NDE(1\|g, \mathbf{C}_i = \bar{\mathbf{c}}_{m\|g})$	4.365	2.706	-0.950	9.735	2.463	1.446	-0.390	5.266
$NIE(0\|g, \mathbf{C}_i = \bar{\mathbf{c}}_{m\|g})$	1.626	1.254	-0.735	4.186	-0.344	0.616	-1.544	0.869

a All covariates other than language group membership have been conditioned on the average values within each language group. Standard errors have been bootstrap across 10,000 bootstrap samples.

Table 5. Natural Direct and Indirect Effects of Early and Regular Head Start Attendance on English Receptive Vocabulary Skills

Estimand	Estimate	Bootstrap SE	95% CIL	95% CIU
$TE\|(\mathbf{C}_i = \bar{\mathbf{c}}_m)$	3.154	1.142	0.873	5.400
$NDE(0\|\mathbf{C}_i = \bar{\mathbf{c}}_m)$	4.424	1.246	2.001	6.878
$NIE(1\|\mathbf{C}_i = \bar{\mathbf{c}}_m)$	-1.271	0.527	-2.313	-0.242
$NDE(1\|\mathbf{C}_i = \bar{\mathbf{c}}_m)$	3.023	1.275	0.500	5.534
$NIE(0\|\mathbf{C}_i = \bar{\mathbf{c}}_m)$	0.131	0.559	-0.977	1.240

this manuscript. Any remaining shortcomings are our own.

Funding Statement Support for this research was provided by Graduate Education at the University of Wisconsin– Madison with funding from the Wisconsin Alumni Research Foundation.

Competing Interests None.

Notes

1 For example, $Y(0, M(1))_i$ represents the outcome that would occur if subject i did not participate in the initial program but experienced subsequent program participation at the level they would have if they had participated in the initial program.
2 For a review of estimation methods for causal mediation effects, we refer interested readers to Qin (2024).
3 We assume that the observed covariates satisfy the identification assumptions outlined above. Addressing potential violations of identification assumptions involves substantial methodological and substantive considerations that are beyond the scope of this article.
4 This study was granted an IRB exemption by the University of Wisconsin-Madison (8/12/2022) and restricted data access approval from the Inter-university Consortium for Political and Social Research (12/5/2022) for secondary analysis of the HSIS data.
5 Pooling results across imputed datasets while also bootstrapping standard errors for groupwise NDE and NIE point estimates is beyond the scope of this article. Interested readers may refer to Schomaker and Heumann (2018) for guidance on combining bootstrapping and multiple imputation.
6 As expected, applying equations 3 and 4 yielded identical TE estimates.
7 R code for computing the proposed closed-form estimates and bootstrap standard errors is available at https://github.com/mikannah/Groupwise-Heterogeneity-in-Mediation-Effects.

References

Bollen, K. A., & Pearl, J. (2013). Eight myths about causality and structural equation models. In *Handbook of causal analysis for social research* (pp. 301–328). Springer.

Castro, D. C. (2014). The development and early care and education of dual language learners: Examining the state of knowledge. *Early Childhood Research Quarterly*, *29*(4), 693–698.

Choi, J. Y., Van Pay, C. K., & Beecher, C. C. (2023). Preschool language exposure and use: A comparison study of dual-language learners and english monolingual children. *Infant and Child Development*, *32*(3), e2420.

Elwert, F., & Winship, C. (2010). Effect heterogeneity and bias in main-effects-only regression models. *Heuristics, probability and causality: A tribute to Judea Pearl*, 327–336.

Garcia, E. B. (2018). The classroom language context and english and spanish vocabulary development among dual language learners attending Head Start. *Early Childhood Research Quarterly*, *42*, 148–157.

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 715–733.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309.

Jenkins, J. M., Sabol, T. J., & Farkas, G. (2018). Double down or switch it up: Should low-income children stay in Head Start for 2 years or switch programs? *Evaluation Review*, *42*(3), 283–317.

Kim, C., Daniels, M., Li, Y., Milbury, K., & Cohen, L. (2018). A bayesian semiparametric latent variable approach to causal mediation. *Statistics in Medicine*, *37*(7), 1149–1161.

Kim, H., & Kim, J.-S. (2024). Assessing the effects of a yearly renewable education program through causal mediation analysis. *Quantitative Psychology: The 88th Annual Meeting of the Psychometric Society, College Park, MD, USA*, 155–164.

Lee, S. Y., Kim, R., Rodgers, J., & Subramanian, S. (2021). Treatment effect heterogeneity in the Head Start Impact Study: A systematic review of study characteristics and findings. *SSM-Population Health*, *16*, 100916.

Li, Y., Mathur, M. B., Solomon, D. H., Ridker, P. M., Glynn, R. J., & Yoshida, K. (2023). Effect measure modification by covariates in mediation: Extending regression-based causal mediation analysis. *Epidemiology*, *34*(5), 661–672.

Li, Y., Yoshida, K., Kaufman, J. S., & Mathur, M. B. (2023). A brief primer on conducting regression-based causal mediation analysis. *Psychological Trauma: Theory, Research, Practice, and Policy*, *15*(6), 930.

Nguyen, T. Q., Schmid, I., & Stuart, E. A. (2021). Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. *Psychological Methods*, *26*(2), 255.

Pearl, J. (2001). Direct and indirect effects. *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence, 2001*, 411–420.

Piker, R. A., & Rex, L. A. (2008). Influences of teacher–child social interactions on english language development in a Head Start classroom. *Early Childhood Education Journal*, *36*, 187–193.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., et al. (2010). Head Start Impact Study. Final Report. *Administration for Children & Families*.

Qin, X. (2024). An introduction to causal mediation analysis. *Asia Pacific Education Review*, 1–15.

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Schomaker, M., & Heumann, C. (2018). Bootstrap inference when using multiple imputation. *Statistics in medicine*, *37*(14), 2252–2266.

Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with sas and spss macros. *Psychological Methods*, *18*(2), 137–150.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

VanderWeele, T. J. (2015). *Explanation in causal inference: Methods for mediation and interaction*. Oxford University Press.

Wang, W., Xu, J., Schwartz, J., Baccarelli, A., & Liu, Z. (2021). Causal mediation analysis with latent subgroups. *Statistics in Medicine*, *40*(25), 5628–5641.

Xie, Y., Near, C., Xu, H., & Song, X. (2020). Heterogeneous treatment effects on children's cognitive/non-cognitive skills: A reevaluation of an influential early childhood intervention. *Social Science Research*, *86*, 102389.