



Fitting the Fractional Polynomial Model to Non-Gaussian Longitudinal Data

Ji Hoon Ryoo^{1*}, Jeffrey D. Long², Greg W. Welch³, Arthur Reynolds⁴ and Susan M. Swearer⁵

¹ Educational Leadership, Foundations, and Policy, University of Virginia, Charlottesville, VA, United States, ² Department of Psychiatry, University of Iowa, Iowa City, IA, United States, ³ Buffett Early Childhood Institute, University of Nebraska, Lincoln, NE, United States, ⁴ Institute of Child Development, University of Minnesota, Minneapolis, MN, United States, ⁵ Department of Educational Psychology, University of Nebraska, Lincoln, NE, United States

As in cross sectional studies, longitudinal studies involve non-Gaussian data such as binomial, Poisson, gamma, and inverse-Gaussian distributions, and multivariate exponential families. A number of statistical tools have thus been developed to deal with non-Gaussian longitudinal data, including analytic techniques to estimate parameters in both fixed and random effects models. However, as yet growth modeling with non-Gaussian data is somewhat limited when considering the transformed expectation of the response via a linear predictor as a functional form of explanatory variables. In this study, we introduce a fractional polynomial model (FPM) that can be applied to model non-linear growth with non-Gaussian longitudinal data and demonstrate its use by fitting two empirical binary and count data models. The results clearly show the efficiency and flexibility of the FPM for such applications.

Keywords: fractional polynomial, generalized additive model, Non-Gaussian longitudinal data, Chicago longitudinal study, reading of the mind

OPEN ACCESS

Edited by:

Holmes Finch,
Ball State University, United States

Reviewed by:

Christoph Scheepers,
University of Glasgow,
United Kingdom
Jaehwa Choi,
George Washington University,
United States

*Correspondence:

Ji Hoon Ryoo
jr3gv@virginia.edu

Specialty section:

This article was submitted to
Quantitative Psychology and
Measurement,
a section of the journal
Frontiers in Psychology

Received: 15 February 2017

Accepted: 07 August 2017

Published: 22 August 2017

Citation:

Ryoo JH, Long JD, Welch GW,
Reynolds A and Swearer SM (2017)
Fitting the Fractional Polynomial Model
to Non-Gaussian Longitudinal Data.
Front. Psychol. 8:1431.
doi: 10.3389/fpsyg.2017.01431

INTRODUCTION

Just as in cross sectional studies, longitudinal studies frequently utilize non-Gaussian data that involve binomial, Poisson, Gamma, inverse-Gaussian distributions, and multivariate exponential families. Fitzmaurice and Molenberghs (2008) categorized models for non-Gaussian longitudinal data into three types based on the way they account for the correlation among the repeated measures and interpret the regression parameters: (1) marginal or population-average models, (2) generalized linear mixed models (GLMM), and (3) conditional and transition models. In the research reported here, we opted to use a GLMM as this emphasizes individual differences and takes into account random effects.

Given the prevalence of non-Gaussian longitudinal data, it is not surprising that a number of statistical tools have been developed for estimating parameters in both fixed and random effects models. Examples include Stiratelli et al. (1984) and Schall (1991), who proposed a penalized quasi-likelihood estimation; Pinheiro and Bates' (1995) Laplace approximation; and Breslow and Clayton (1993), who combined marginal quasi-likelihood estimation with a penalized quasi-likelihood estimation. Other approaches that have been proposed include Anderson and Aitkin's (1985) adaptive Gaussian quadrature; McCulloch's (1997) Monte Carlo EM algorithm; Kuk and Cheng's (1997) Monte Carlo Newton-Raphson algorithm; and Zeger and Karim's (1991) Monte Carlo

integration via Gibbs sampling within the framework of the GLMM. However, in spite of the many estimating methods that have been suggested, growth modeling with non-Gaussian data remains somewhat limited when modeling the transformed expectation of the response when applying a linear predictor as a functional form for the explanatory variables.

As summarized by Fox (2016), the generalized linear model (GLM) consists of: (1) a random component specifying the conditional distribution of a response variable, Y_i ; (2) a linear predictor as a linear function of multiple regressors, $\eta_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$; and (3) a smooth and invertible linearizing link function, $g(\cdot)$, that transforms the expected value of the response variable to a linear predictor such as $g(E(Y_i)) = \eta_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$. In the applications of GLM reported in the literature, the random component and the link function are generally well-organized, for example in the form of exponential families, as in Nelder and Wedderburn (1972). However, the functional form for the linear predictor has been largely restricted to a linear combination of explanatory variables, interactions, and polynomial regressors. Although the link function resolves the issue related to conditional distributions in the regression, the trends exhibited in the associations between the latent variable of a dependent variable and the explanatory variables within the linear predictor would be more various and complex. For example, the association could be an asymmetric curve that could not be resolved simply by interactions or polynomial regressors. In this paper, we introduce a more flexible functional form for the linear predictor, which is known as a fractional polynomial regressor (Royston and Altman, 1994; Long and Ryoo, 2010).

In the social and behavioral sciences, modeling individual differences in a change process is vital when compensating for a small sample and/or reducing sampling bias. Random effects models can be utilized to model these individual differences. A GLMM can be defined by adding random effects representing within-subject variations in the linear predictor (Laird and Ware, 1982):

$$\eta_i = g(\mu_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip} + b_0 + b_1 Z_{i1} + \dots + b_q Z_{iq} = X_i \beta + Z_i b \quad (1)$$

where μ_i is the expected value of the response variable, X_i is a known design matrix linking β and η_i and Z_i is a known design matrix linking b and η_i . In practice, Z_i is often constrained within the random intercept model, since non-Gaussian data provide relatively little information about individual heterogeneity beyond variability in the random intercept (Long et al., 2009). In the work reported here, Z_i is extended to random slope variability to account for non-Gaussian data, including both count data and binary data.

In longitudinal data, the time variable is treated as a covariate due to its key role in modeling change over time. For that reason, the first step in model selection involves a visual inspection of mean change over time (Pinheiro and Bates, 2000; Molenberghs and Verbeke, 2005). Though the result of the visual inspection is subjective and not necessarily representative, it does establish

a starting point for the analyses. In this step, a generalized additive model (GAM; Hastie and Tibshirani, 1986; Berhane and Tibshirani, 1998) can offer useful support for the results obtained by a visual inspection of the data. As a non-parametric analytic tool, GAM is well-known and performs better than the local regression model (LOWESS; Cleveland, 1979; Cleveland and Devlin, 1988) when modeling the relationship between a response variable and each predictor.

To model the time transformation and articulate change over time, a conventional polynomial model (CPM) is typically utilized to identify the function with the best fit (Fitzmaurice et al., 2004, chap. 12). However, a CPM seldom offers the optimum solution and neither does it provide useful descriptive information, since the CPM is always symmetric around the local maxima and local minima and also requires an additional parameter for each fluctuation in the pattern of change. There are many alternatives to modeling asymmetric trends that provide more parsimonious models with the same quality of information on model fit (Sterba, 2014). Here, we focus on the use of fractional polynomial models (FPMs; Royston and Altman, 1994). Compared with CPMs, FPMs have received relatively little attention in the context of non-Gaussian longitudinal data in the social and behavioral sciences, even though FPMs are more flexible than CPMs and provide broader classes for model selection. The flexibility and parsimoniousness of FPMs in model selection for Gaussian longitudinal data have been discussed in Long and Ryoo (2010). This paper investigates model selection in the linear predictor for non-Gaussian longitudinal data, with an extension of the time transformation and random effects onto FPMs. This is also referred to as the generalized fractional polynomial mixed model (GFPMM).

METHODS

Two real world datasets were analyzed to demonstrate a unified framework for fitting GFPMM. The model selection procedure for GFPMM suggested by Ryoo (2011) in LMM was used to demonstrate the process involved in building the model by fitting the time transformation, static predictor, interactions, and random effects in Equation (1).

Data Sources

To demonstrate the parsimoniousness and flexibility in fitting non-Gaussian data using the GFPMM, two longitudinal datasets including binary and count responses were used. The first dataset consisted of a binary response over four time points, while the second dataset consisted of a count response over three time points. In the second dataset, two static predictors were also considered. In this exploratory data analysis, those two variables were tested to determine whether they were significant variables for the best fitting model.

Chicago Longitudinal Study

As part of the Chicago Longitudinal Study (CLS, <http://www.cehd.umn.edu/icd/cls/>), data were collected from 1,539 Chicago public school students deemed to be at risk of educational and social difficulties due to economic disadvantage. Based on

parents and teachers' surveys at ages 3, 8, 12, and 17, a risk index was computed by combining risk factors such as mother's age, mother's education level, single parenting status, number of children in the household, eligibility for free school lunches, mother's employment, and the poverty index in the local school area, a total of four data points for each child. The scale of the ordinal risk index from 0 to 8 was then further dichotomized as 0 for low risk (<4) and 1 for high risk (4 or higher). In this demonstration, the change in the proportion of high risk students was examined by applying GFPMM.

Based on the dichotomous classification risk index for each of the 1,539 students included in the dataset, 1,122 (72.9%), 1,004 (65.2%), 983 (63.9%), and 823 (53.5%) were deemed to be at-risk at ages 3, 8, 12, and 17 years, respectively. Plotting the mean proportion changes (Figure 1), the numbers of at-risk students clearly decrease over time in a non-linear manner, with the proportion at age 17 dropping significantly below that at age 12. To further investigate the trend, the non-linear relationship between the binary response variable and the continuous time variable was modeled by applying GFPMM and GAM. To visualize the individual differences in terms of the change of at-risk status, we also plotted separate curves for 24 individuals randomly selected from the 1,539 participants (Figure 2). Within these small randomly selected datasets, we observed nine difference patterns in terms of their at-risk status. To articulate both the overall change and individual change patterns, we utilized GAM analysis as a diagnostic tool to capture these patterns.

Reading of the Mind in the Eyes Test-Revised

The Reading of the Mind in the Eyes Test-Revised (RMET-R; Baron-Cohen et al., 2001) is a 28-item measure designed

to assess children's ability to make social judgments based on stimuli such as the eye region of another person's face. This test involves 28 photographs of the facial eye region, each accompanied by four words describing various mental states. In the present study, 212 participants were asked to select the word that best described the mental state exhibited in the accompanying image, with a point awarded for each correct answer. The points were then summed to compute an overall score representing the participant's response at each time point, calculated as 28 minus the total score. Lower scores thus indicate higher "Theory of Mind." Data were collected at three different time points spaced at ~6 month intervals. Two demographic variables, gender (1 for girl and 2 for boy) and first language (1 for English and 2 for Others), are included in this demonstration.

The means for the RMET-R test results are plotted in Figure 3. At first glance, the sharp drop between Time 2 and Time 3 would seem to indicate that a non-linear model would be more appropriate for modeling the change than a linear model. However, as the data presented in Figure 4 suggests, this may not in fact be the case; the wide variations in the individual curves for 20 randomly selected participants may indicate non-linearity and decreases over time. To articulate both the overall and individual change patterns, we therefore opted to utilize GAM analysis as a diagnostic tool in order to capture both patterns.

Generalized Additive Model (GAM)

To aid in the visual inspection of the data, GAM incorporating non-parametric regression and smoothing was utilized. GAM was initially used to model the change at the group level by taking individual changes into account. Unlike GLM, the linear predictor for GAM uses smooth functions rather than linear

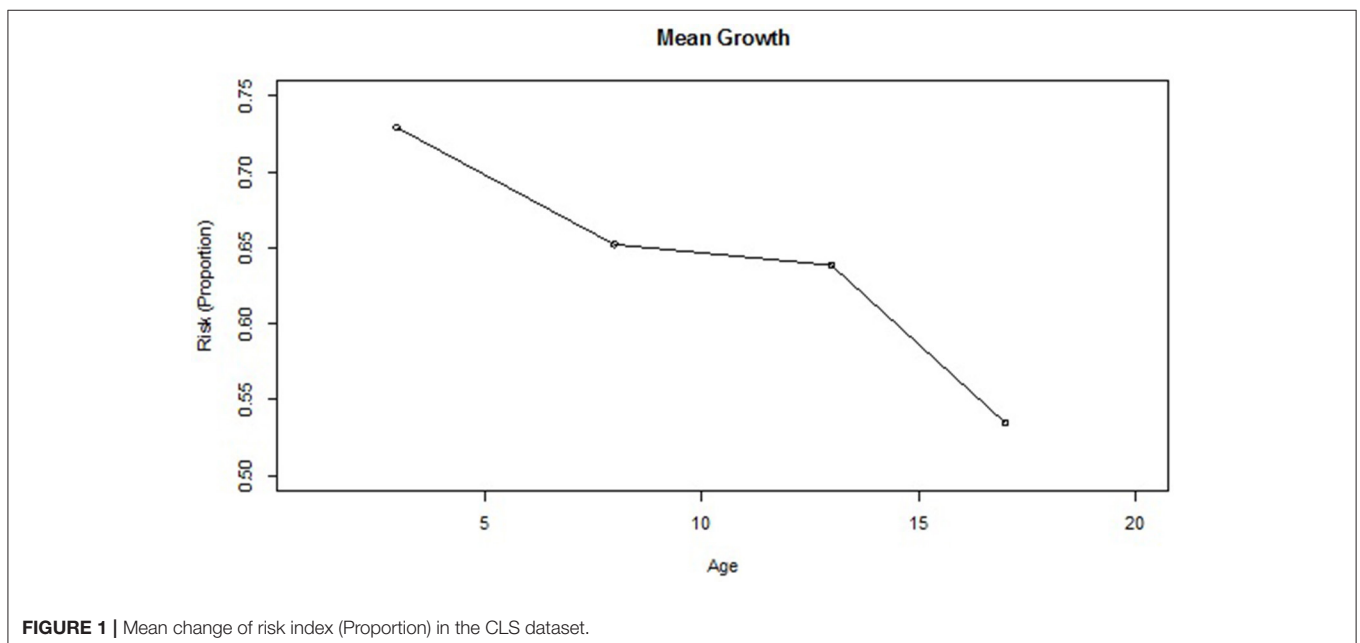


FIGURE 1 | Mean change of risk index (Proportion) in the CLS dataset.

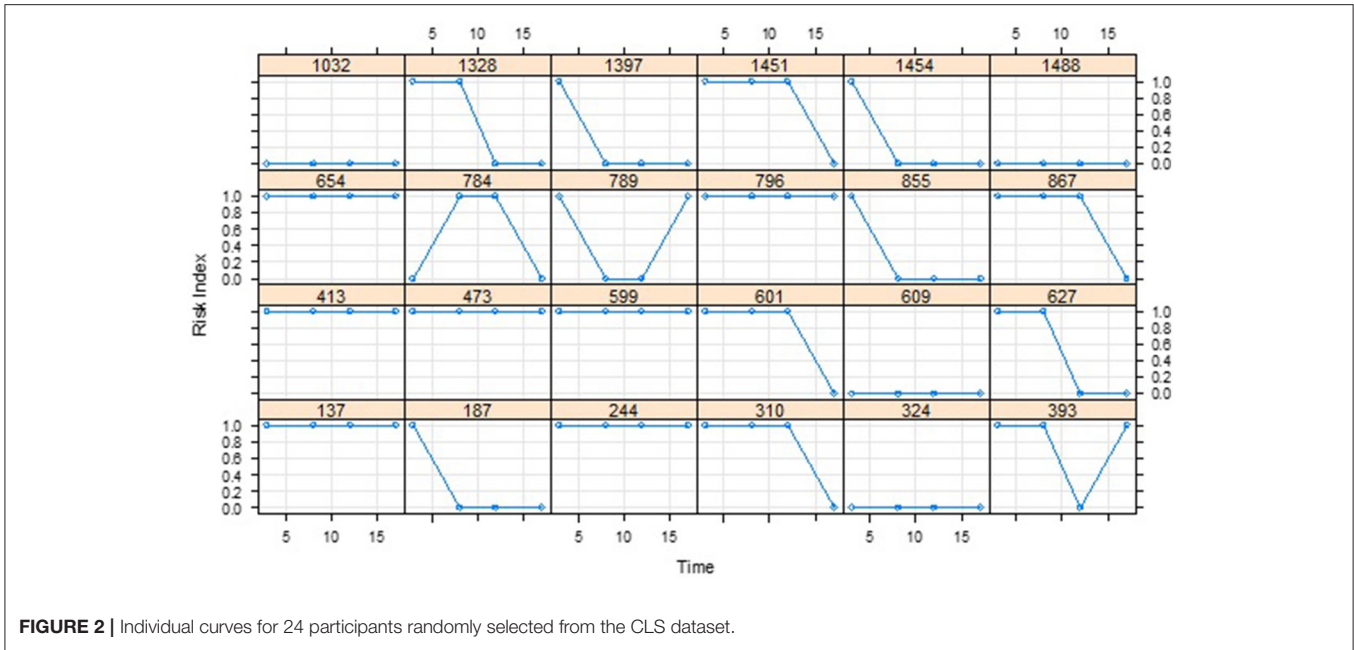


FIGURE 2 | Individual curves for 24 participants randomly selected from the CLS dataset.

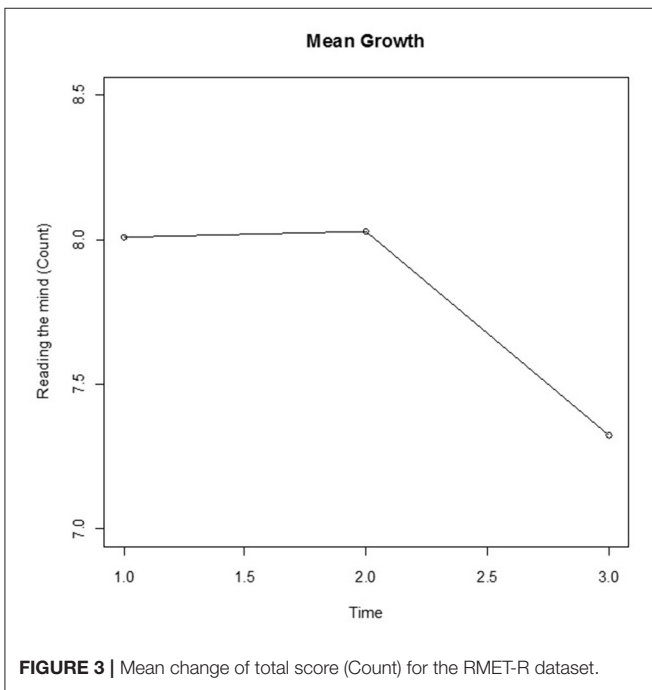


FIGURE 3 | Mean change of total score (Count) for the RMET-R dataset.

means of examining possible non-linear change between the response variable and each predictor variable.

Generalized Fractional Polynomial Mixed Model (GFPM)

The primary statistical model utilized in this study was GFPM, which can be written in terms of the linear predictor, η_{ij} , using the following the individual level form:

$$\begin{aligned} \eta_{ij} &= \beta_0 + \sum_{b=1}^p \beta_b f_b(t_{ij}) + \sum_{e=1}^l \gamma_e X_{ej} + b_{i0} + \sum_{g=1}^q b_{ig} f_g(t_{ij}) \\ &= X_i \beta + Z_i b \end{aligned} \tag{3}$$

where p and q indicate the orders of the FP terms for the time variable for the fixed effects and the random effects, respectively, and l indicates the number of static predictors. The indices i and j denote the i th time and j th person, respectively. An FP function, f , is defined as:

$$f_h(y_{ij}) = \begin{cases} y_{ij}^{(m_h)}, & \text{if } m_h \neq m_{h-1}, \\ f_{h-1}(y_{ij}) \cdot \log(y_{ij}), & \text{if } m_h = m_{h-1}, \end{cases} \tag{4}$$

where $m_1 \leq m_2 \leq \dots \leq m_p$. The parentheses on the exponent signify Box and Tidwell's (1962) transformation:

$$y_{ij}^{(m_h)} = \begin{cases} y_{ij}^{m_h}, & \text{if } m_h \neq 0, \\ \log(y_{ij}), & \text{if } m_h = 0, \end{cases} \tag{5}$$

with the constraint $y_{ij} > 0$, so that all transformations are defined. The power terms (exponents), m_h , are taken from the set of values suggested by Royston and Altman (1994) for general curve fitting. The set is:

$$M = \{-2, -1, -0.5, 0, 0.5, 1, 2, \max(3, m)\} \tag{6}$$

combination of predictors as follows:

$$\eta = s_0 + \sum_{i=1}^p s_i(X_i) \tag{2}$$

where $s_1(\cdot), \dots, s_p(\cdot)$ are smooth functions.

Applying back-fitting and local scoring algorithms and generalized cross validation (GCV; Wahba, 1990), GAM estimates the smooth functions and consequently provides a

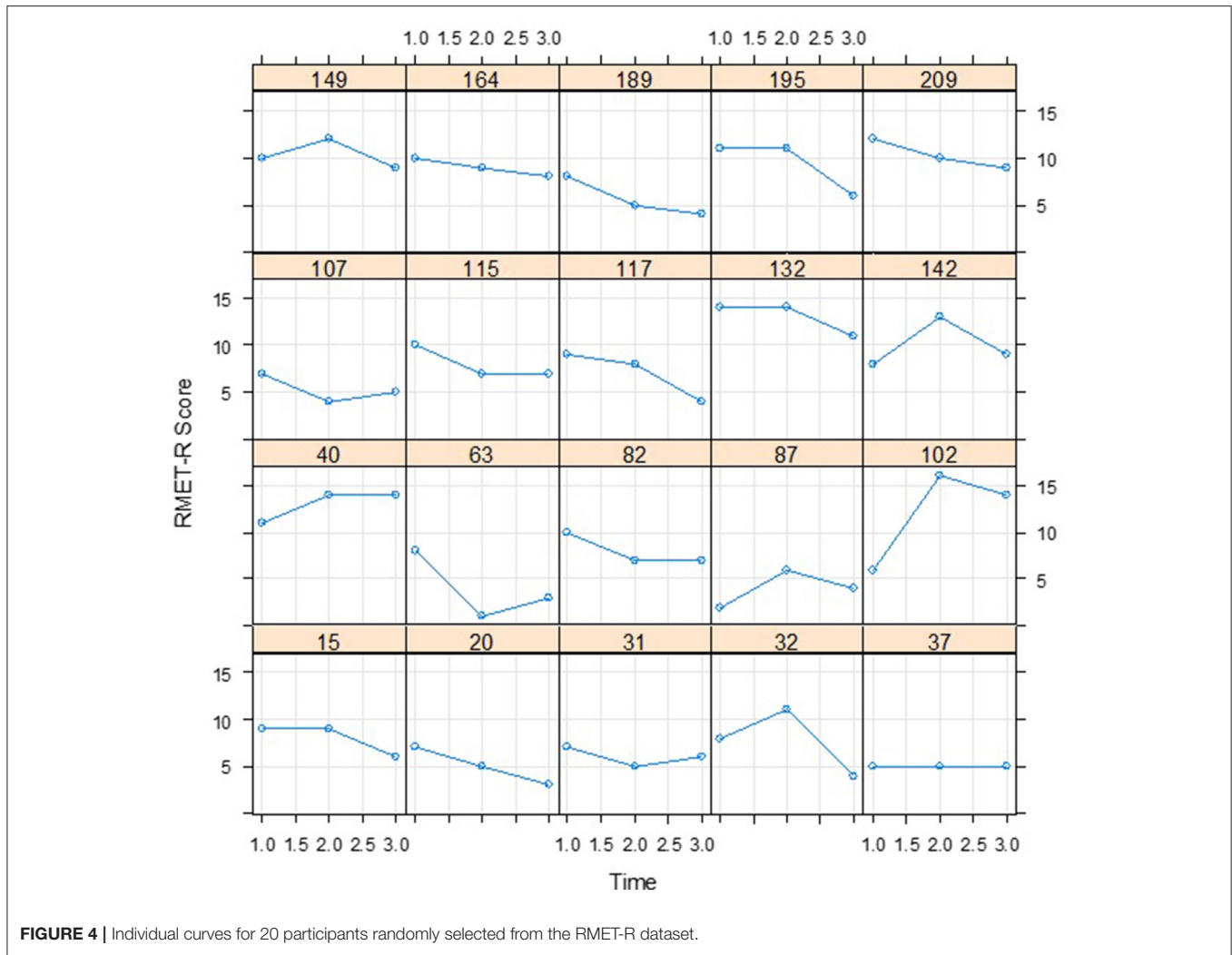


FIGURE 4 | Individual curves for 20 participants randomly selected from the RMET-R dataset.

where m is the order of the FP. The values in M offer a wide variety of curve shapes and constitute transformations that are familiar to applied researchers. **Figure 5** presents the shapes of the first order FPs, demonstrating this approach's flexibility in terms of fitting a fixed model.

In the equation for the linear predictor, η_{ij} , the first three terms refer to the fixed effects and the last two to the random effects. Applied researchers may also be interested in the interactions between time variables and additional explanatory variables. This term can be added by multiplying the 2nd and 3rd terms in the equation for the linear predictor, η_{ij} . If no between-person variation is expected, the random effects terms can be dropped.

Maximum Likelihood Estimation for GFPM

Within the format of a GLMM, GFPM can also be estimated using one of three common estimation methods: penalized quasi-likelihood (PQL; Schall, 1991; Breslow and Clayton, 1993), integral approximations such as the Laplace approximation (Pinheiro and Bates, 1995), or Gauss-Hermite quadrature

(McCulloch, 1997). In the current study, we used the Laplace approximation as we consider it to obtain more accurate results than PQL and to avoid the additional computational burden imposed by Gauss-Hermite quadrature.

The GFPM can be written in matrix format as follows:

$$\eta = X \cdot \beta + Z \cdot b \text{ and } g(\mu) = g[E(y|b)] = \eta \quad (7)$$

where $b \sim N(0, \Psi)$. The mean and variance conditioned random effects can also be written as $E(y|b) = \mu$ and $V(y|b) = \varphi \cdot v(\mu)$, respectively, where φ is a dispersion parameter and $v(\mu)$ is a variance function that depends on the distribution family. Based on the multivariate normality assumption for $v(\mu)$, we can estimate the fixed effects β , along with the dispersion parameter φ , by maximizing the marginal distribution:

$$p(y|\beta, \varphi) = \int_b p(y|\beta, \varphi, b)p(b|\Psi)db \quad (8)$$

We can now re-estimate the variance function, $v(\mu)$, using a pseudo-variable, y^* , that is predicted by the linear predictor.

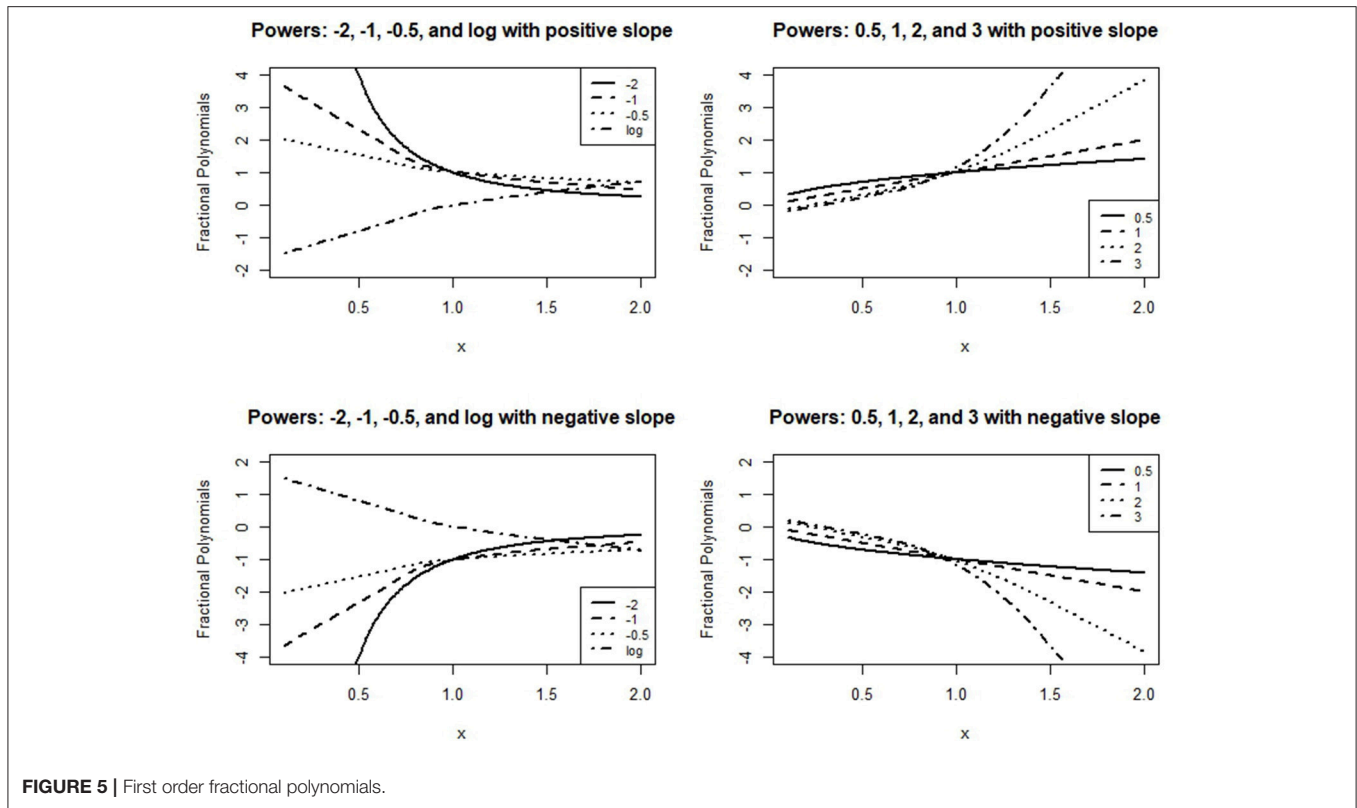


FIGURE 5 | First order fractional polynomials.

Estimation continues in this manner until convergence is achieved. The R `lme4` package (Bates et al., 2015) was used to estimate the parameters in GFPMM.

Model Comparison

After diagnosing the trends in the change patterns using mean changes and the GAM results, we applied the generalized conventional polynomial mixed model (GCPMM) to both datasets. We considered conventional up to quadratic polynomials for the CLS data and a linear model for the RMET-R data with up to random slope models due to their respective four and three time points. The best fit was obtained for the GFPMM with random slope models. We then compared the best fitting GCPMM and GFPMM to identify the model with the best fit. When comparing models, we utilized a generalized likelihood ratio test (GLRT; Cox and Hinkley, 1974; Wood, 2006) with a significance level of 0.05, and delta (Δ) methods (Burnham and Anderson, 2004) based on AICc (Hurvich and Tsai, 1989) and BIC (Schwarz, 1978), applying the criteria that models with $\Delta_i = \Delta AICc_i = AICc_i - \min(AICc) \leq 2$ have substantial support, models with $4 \leq \Delta_i \leq 7$ have considerably less support, and models with $\Delta_i > 10$ have essentially no support. This last delta criterion has also been applied to BIC in similar applications in the Bayesian literature (Raftery, 1996).

RESULTS

Chicago Longitudinal Study Dataset

As shown in Figure 1, the proportion of high risk students drops from age 3 to age 8, levels off to some extent from age 8 to

age 12, and then drops sharply from age 12 to age 17. This is somewhat inconsistent with the results obtained from the GAM analyses, which generates the linear model shown. The GAM curve follows a slightly different pattern from the mean change shown in Figure 1, indicating a decreasing pattern and faster drop for the oldest students. Interestingly, the GAM analysis yields an effective degree of freedom of one, meaning that the suggested value for the exponent of the time variable is linear. Based on both these findings, the quadratic polynomial model was considered to provide the best fit for the time transformation.

Analyses using the polynomial methods begin by implementing GCPMM. In the class of conventional polynomials with four time points, the suggested model would be considered to be either linear or quadratic polynomial; a cubic model would be a saturated model for fixed effects. For random effects, the models up to the random slope term were considered in a well-formulated model. The well-formulated model includes lower order terms when higher order terms are significant, regardless of the significance of the lower order terms (Morrell et al., 1997). Table 2 displays the results of the competing models utilizing GCPMM.

The results shown in Table 1 suggest that the best fit is obtained for the quadratic model with linear random effects, described in equation form as follows:

$$\log \left\{ \frac{\Pr(Y_{ij} = 1|b_{0i}, b_{1i})}{\Pr(Y_{ij} = 0|b_{0i}, b_{1i})} \right\} = \eta_{ij} = \beta_0 + \beta_1 \cdot time_{ij} + \beta_2 \cdot time_{ij}^2 + b_{0i} + b_{1i} \cdot time_{ij} \quad (9)$$

TABLE 1 | Results of model selection within generalized conventional polynomial mixed models (GCPMMs) for the CLS dataset.

Models	Number of parameters	BIC	AICC	GLRT	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
				Chisq (p -value)		
gcpm.10 ^a	3	6,179	6,159		52.46	69.03
gcpm.20 ^b	4	6,186	6,159	1.24 (0.27)	59.94	69.79
gcpm.11 ^c	5	6,126	6,093	68.67 (0.00)	0.00	3.13
gcpm.21 ^d	6	6,130	6,090	5.13 (0.02)	3.59	0.00

^aRandom intercept linear model;
^bRandom intercept quadratic model;
^cLinear model with random slope;
^dQuadratic model with random slope.

The resulting GCPMM does not adequately describe the drops observed in the results for the quadratic model, however, since conventional polynomials are symmetric around vertexes. On the other hand, the shapes of fractional polynomials vary more widely than conventional polynomials, as can be seen in **Figure 5**. In addition, FPMs can be fit with a relatively small number of parameters; although the GCPMM with the best fit (Equation 9) includes six parameters, the GFPMM with the best fit (Equation 10) includes only five. The results of the model selection for the random intercept and intercept FPM are summarized in **Table 2**. The results suggest the FPM with power 3 provides the best fit in terms of the two delta methods, BIC and AICC; delta values for the GFPMM with power -1 are >10 . Thus, the model with the best fit for GFPMM combines a cubic polynomial term with random slope effects and can be written in equation form as follows:

$$\log \left\{ \frac{\Pr(Y_{ij} = 1|b_{0i}, b_{1i})}{\Pr(Y_{ij} = 0|b_{0i}, b_{1i})} \right\} = \eta_{ij} = \beta_0 + \beta_1 \cdot \text{time}_{ij}^3 + b_{0i} + b_{1i} \cdot \text{time}_{ij}^3 \quad (10)$$

When comparing the results of the best fitting models for GCPMM and GFPMM, GFPMM performs better than the best GCPMM, as indicated by the $\Delta(\text{BIC})$ and $\Delta(\text{AICC})$ values of 132.3 and 125.6, respectively, in **Table 3**. This result is also supported by **Figure 6**, where the GFPMM with the best fit shows a stable decrease between ages 3 and 12, at which point there is a sudden drop.

RMET-R

The RMET-R data consists of only three time points. **Figure 3** displays the mean changes, with a relatively stable value from Time 1 to Time 2 followed by a decrease from Time 2 to Time 3, which is similar to the results of the GAM analysis. However, the 20 randomly selected individual curves from among the 212 participants plotted in **Figure 4** fail to exhibit any coherent pattern but instead clearly show the between-subject variability.

TABLE 2 | Results of model selection in 1st order generalized fractional polynomial mixed models (GFPMMs) with random intercept and slope for the CLS dataset.

	Power	BIC	AICC	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
Random intercept	-2	6,149	6,129	151.15	164.60
GFPMM (df = 3)	-1	6,121	6,101	123.88	137.33
	-0.5	6,103	6,083	105.20	118.64
	0	6,084	6,064	86.43	99.87
	0.5	6,069	6,049	71.85	85.30
	1	6,062	6,042	64.46	77.91
	2	6,066	6,046	68.46	81.91
	3	6,079	6,059	81.92	95.36
Random slope	-2	6,278	6,244	280.02	280.02
GFPMM (df = 5)	-1	6,031	5,997	33.41	33.41
	-0.5	6,073	6,039	75.07	75.07
	0	6,098	6,065	100.61	100.61
	0.5	6,113	6,080	115.84	115.84
	1	6,126	6,093	128.73	128.73
	2	6,147	6,113	149.25	149.25
	3	5,998	5,964	0.00	0.00

TABLE 3 | Results of model comparison between the best fitted GCPMM and GFPMM for the CLS dataset.

Models	Number of parameters	BIC	AICC	GLRT	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
				Chisq (p -value)		
GFPMM ^a	5	5998	5964		0.0	0.0
GCPMM ^b	6	6130	6090	0.000 (1.000)	132.3	125.6

^a1st order fractional polynomial model with power 3 and random slope;
^bQuadratic polynomial model with random slope.

This random effect was thus taken into account for the between-subject variability. In the GCPMM class, the random effects can be either random intercept or random slope.

Both the GLRT and the delta method indicate that the random slope model performs significantly better than the random intercept model (**Table 4**). Next, variable selections were conducted via the GLRT and delta methods to select the variable that most affects the model. The result of both methods indicated that language is not significant but gender is. Interestingly, unlike the main effect of the gender variable, the interaction effect between the gender and time variables was not significant. The results for the gender variable are summarized in **Table 4**. The GCPMM with the best fit can be written as follows:

$$\log \{E(Y_{ij}|b_i)\} = \eta_{ij} = \beta_0 + \beta_1 \cdot \text{time}_{ij} + \beta_2 \cdot \text{gender}_i + b_{0i} + b_{1i} \cdot \text{time}_{ij} \quad (11)$$

Next, the class of model was extended to GFPMM. As can be seen in **Figure 3**, the mean remains roughly constant between Times 1 and 2 but then drops off sharply. The model with the best fit thus needs to describe this behavior over time. Among the random

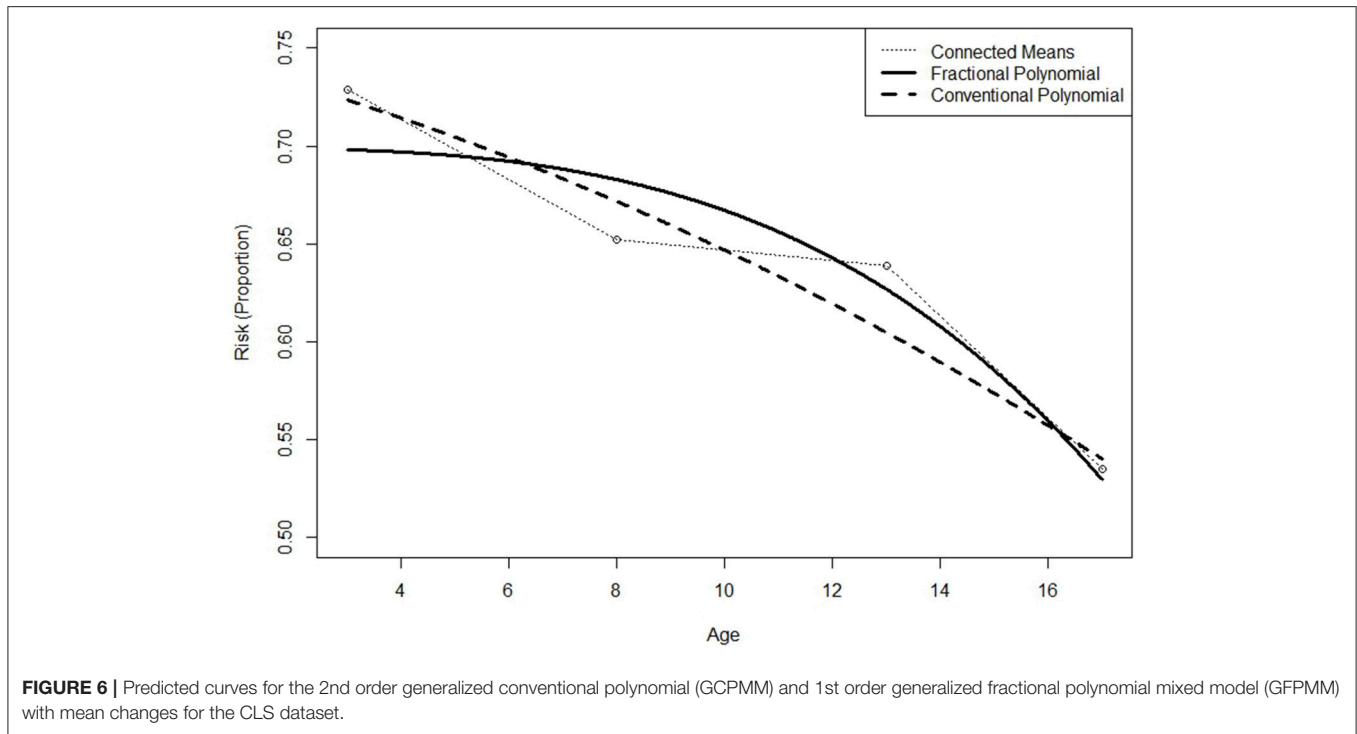


FIGURE 6 | Predicted curves for the 2nd order generalized conventional polynomial (GCPMM) and 1st order generalized fractional polynomial mixed model (GFPM) with mean changes for the CLS dataset.

TABLE 4 | Results of model selection within GCPMM for the “Reading the Mind” dataset ($N = 212$).

Models	Number of parameters	BIC	AICC	GLRT	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
				Chisq (p -value)		
gcpm.10 ^a	3	885.846	872.519		129.960	142.259
gcpm.11 ^b	5	755.887	733.706	142.870 (0.000)	0.000	3.446
gcpm.11g ^c	6	756.857	730.260	5.485 (0.019)	0.971	0.000
gcpm.11gi ^d	7	762.065	731.056	1.248 (0.264)	6.178	0.797

^aRandom intercept linear model.

^bLinear model with random slope.

^cLinear model with random slope and gender.

^dLinear model with random slope, gender, and interaction between gender and time.

intercept models tested, this would appear to be the GFPM with power 3 based on the results presented in **Table 5**. However, the GFPMs with powers 1, 2, and 3 all have substantial support based on the criteria suggested by Burnham and Anderson (2004) so instead of selecting a single model, these three GFPMs were all considered candidates for the best fit.

As discussed in the Introduction Section, the random effects were extended to the random slope. The results, summarized in **Table 5**, suggests that all the GFPMs except the GFPM with power 1 have substantial support. Regardless of the structure of the random effects, the GFPMs with powers 2 and 3 have substantial support. Since both these GFPMs performed better

TABLE 5 | Model selection in the 1st order generalized fractional polynomial mixed models (GFPMs) with random intercept and slope for the “Reading the Mind” dataset.

	Power	BIC	AICC	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
Random intercept GFPM (df = 3)	-2	889.069	875.742	153.200	162.054
	-1	888.242	874.914	152.373	161.226
	-0.5	887.703	874.375	151.833	160.687
	0	887.100	873.772	151.230	160.084
	0.5	886.466	873.139	150.597	159.451
	1	885.846	872.519	149.977	158.831
	2	884.803	871.475	148.933	157.787
	3	884.126	870.799	148.257	157.111
Random slope GFPM (df = 5)	-2	737.510	715.329	1.641	1.641
	-1	736.963	714.782	1.094	1.094
	-0.5	736.658	714.477	0.789	0.789
	0	736.365	714.185	0.496	0.497
	0.5	737.776	715.595	1.907	1.907
	1	755.887	733.706	20.018	20.018
	2	735.869	713.688	0.000	0.000
	3	736.076	713.895	0.207	0.207

in the further model comparison, they are jointly considered the best fitted models.

We further investigated the effect of gender by fitting all eight GFPMs with random slopes. In all the 1st order GFPMs, the gender is significant (all of $p \leq 0.032$); although there is no interaction effect with the time variable (all of $p \geq 0.110$) at

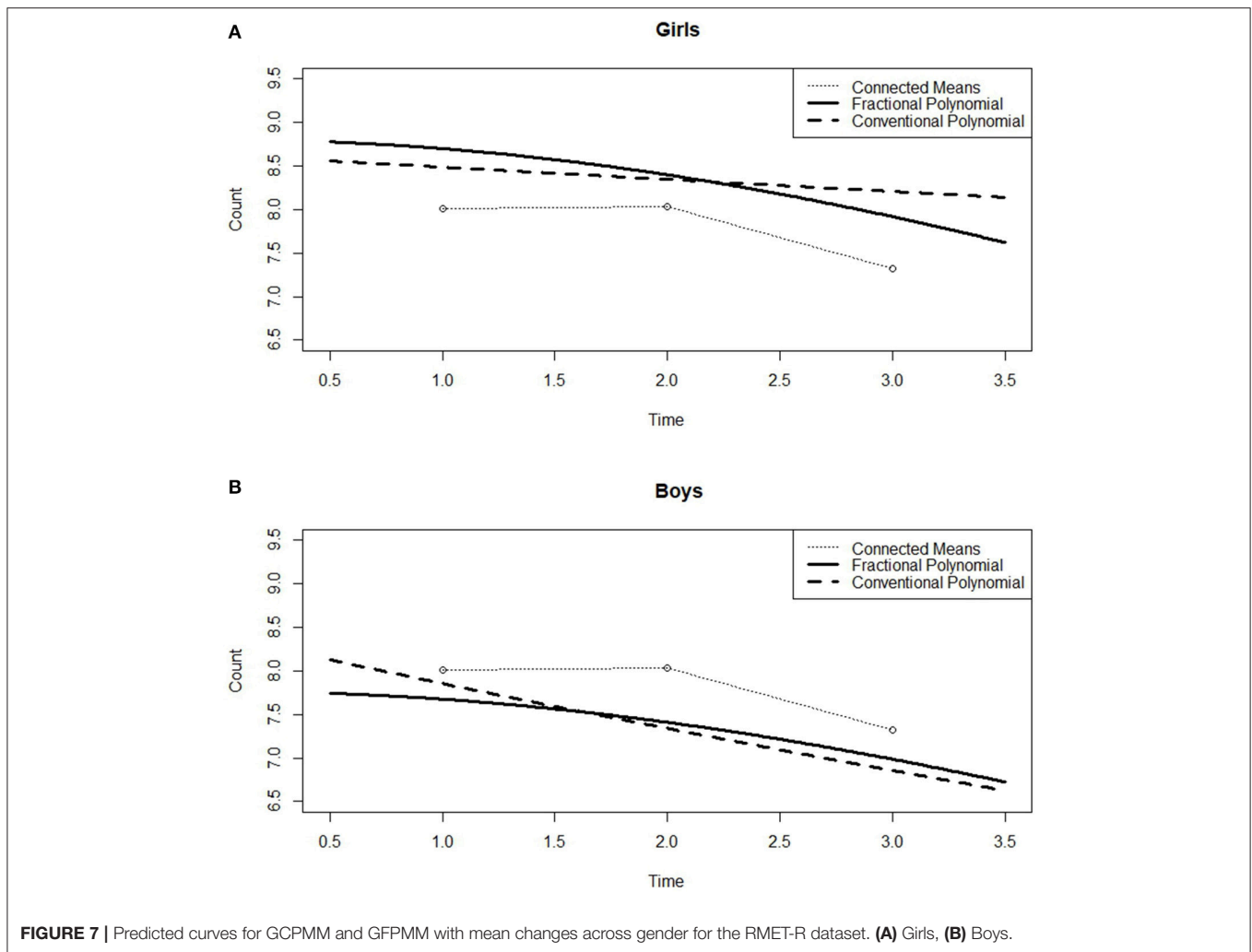


FIGURE 7 | Predicted curves for GCPMM and GFPMM with mean changes across gender for the RMET-R dataset. **(A)** Girls, **(B)** Boys.

a significance level of 0.05. As a result, the gender variable was added to the model as a main effect. This result can be confirmed by examining **Figure 7**. As the figure shows, the mean changes for boys and girls differ slightly. Comparing these mean changes with those predicted by the GFPMM with power 2 and the linear GCPMM, the mean for the girls is lower than that for the boys, which indicates that as a group, the girls have a higher “Theory of Mind” than the boys.

After the gender variable was added into the 1st order GFPMMs, all the 1st order GFPMMs were compared and the results are summarized in **Table 6**. Although five of the GFPMMs have substantial support, the GFPMM with power 2 has the smallest BIC and AIC_C. The model equation can thus be written as follows:

$$\log \{E(Y_{ij}|b_i)\} = \eta_{ij} = \beta_0 + \beta_1 \cdot time_{ij}^2 + \beta_2 \cdot gender_i + b_{0i} + b_{1i} \cdot time_{ij}^2 \quad (12)$$

Within the various conditions tested for the FPMs, the models with the best fit are the GFPMM with power 2 in terms of BIC and the GFPMM with power 2 and the gender variable in terms

of AICc. Both GFPMMs have random slope as their random effect and both are substantially supported compared with the best fitted models shown in **Table 7**. However, GLRT indicates that the GFPMM with power 2 and the gender variable performs significantly better than the GFPMM with power 2 alone.

DISCUSSION

In this study, the efficiency and parsimoniousness of GFPMM have been investigated for non-Gaussian longitudinal data, specifically in the context of binary and count responses. In line with the flexibility of shapes in FPMs, the GFPMMs with the best fit performed better than GCPMMs for both types of data according to the GLRT and delta methods used to assess the results of two empirical studies. In addition, GFPMMs are more parsimonious than GCPMMs. As shown in the analyses above, GFPMMs can be utilized to achieve more parsimonious (i.e., requiring fewer parameters), better fitting models than GCPMMs.

The results of the GAM analyses provide fairly consistent models for the GFPMMs with the best fits, which shows that

TABLE 6 | Results of model selection in the 1st order generalized fractional polynomial mixed model (GFPMM) with random slope and gender for the “Reading the Mind” dataset.

Power	BIC	AICC	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
-2	739.359	712.761	2.052	2.051
-1	738.765	712.167	1.458	1.457
-0.5	738.423	711.825	1.116	1.115
0	738.081	711.483	0.774	0.773
0.5	739.555	712.957	2.248	2.247
1	756.857	730.260	19.550	19.550
2	737.307	710.710	0.000	0.000
3	737.381	710.783	0.074	0.073

TABLE 7 | Results of model comparison between the best fitted GCPMM and GFPMM in the “Reading the Mind” dataset.

Models	Number of parameters	BIC	AICC	GLRT	$\Delta(\text{BIC})$	$\Delta(\text{AICC})$
				Chisq (<i>p</i> -value)		
GFPMM ^a	5	735.869	713.688		0.000	2.979
GFPMM ^b	6	737.307	710.710	5.017 (0.025)	1.438	0.000
GCPMM ^c	6	756.857	730.260	0.000 (1.000)	20.988	19.550

^a1st order fractional polynomial model with power 2 and random slope;

^b1st order fractional polynomial model with power 2, random slope, and gender;

^cLinear model with random slope and gender.

the non-parametric modeling is well-incorporated with the parametric modeling conducted via standard model comparison tools, namely the GLRT and delta methods. Such compatibility is not achievable with the comparable GCPMMs due to the limited choice of shapes available. Thus, for the case of fitting non-linear trends in non-Gaussian longitudinal data, we recommend that applied researchers utilize GFPMM with GAM and using the GLRT and delta methods for their modeling. In this paper, we mainly focused on exploratory analysis. However, it is not limited to exploratory analysis but the procedure can also be used for any confirmatory analysis.

We have also demonstrated the use of GFPMM in the context of non-linear patterns of change. GFPMM offers a flexible and efficient means of modeling non-linear change in continuous, binary, and count data within the linear predictor. As our analyses of the CLS and RMET-R datasets showed, GFPMM can provide a more parsimonious solution than the traditional approach based on GCPMM.

The utility of this approach does require further investigation in the context of a multitude of different data situations where

non-linear change is present. In general, the main difficulty in fitting polynomial models lies in the interpretation of the results. That is, what does the model tell us about the substantive nature of the data? Answering this question lies at the heart of every investigation undertaken by applied researchers. In the context of the types of data illustrated above, a better model fit was achieved by using GFPMM rather than GCPMM. Of course, the need to estimate fewer parameters inevitably eases the interpretation of the model and in the cases illustrated above GFPMM provided a better model fit with fewer parameters.

Royston and Altman (1994) noted that a GFPMM with order >3 is rarely fitted due to its inherent complexity, which does not actually mean that their flexibility is limited but rather that GFPMMs up to order two provide a wide variety of shapes that are generally sufficient to model non-linear trends in longitudinal data. The shapes of a 2nd order GFPMM can be found in Long and Ryoo (2010). If there are many time points, then dynamic changes would be expected in practice. In the case, higher order GFPMMs can also be applicable. Such flexible and parsimonious properties in the GFPMM suggest applied researchers to consider it when asymmetric growth pattern is observed or expected. Although a number of time points will not be associated with using the GFPMM, three to six time points often favor the GFPMM than the GCPMM (Long and Ryoo, 2010; Ryoo et al., 2017). Future studies will examine the efficiency via a simulation study. By considering the random effects introduced to take in account the between-person variability, GFPMM incorporated with GAM provides an advanced statistical model that takes care of these individual differences. In the social and behavioral sciences, the articulation of individual difference is of paramount importance. Thus, fitting GFPMM to non-Gaussian longitudinal data will enable applied researchers in these areas to articulate individual differences in model selection.

ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

AUTHOR CONTRIBUTIONS

JR designed this study, performed the statistical analysis, and drafted the manuscript. JL and GW reviewed analysis, and drafted the manuscript; AR collected data and drafted the manuscript; SS collected data and drafted the manuscript. All authors read and approved the final manuscript.

REFERENCES

Anderson, D. A., and Aitkin, M. (1985). Variance components models with binary response: interviewer variability. *J. R. Stat. Soc. Ser. B* 47, 203–210.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Berhane, K., and Tibshirani, R. J. (1998). Generalized additive models for longitudinal data. *Can. J. Stat.* 26, 517–535. doi: 10.2307/3315715
- Box, G. E. P., and Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics* 4, 531–550. doi: 10.1080/00401706.1962.10490038
- Breslow, N. E., and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *J. Am. Stat. Assoc.* 88, 9–25. doi: 10.1080/01621459.1993.10594284
- Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Metho. Res.* 33, 261–304. doi: 10.1177/0049124104268644
- Cleveland, W. G. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74, 829–836. doi: 10.1080/01621459.1979.10481038
- Cleveland, W. S., and Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *J. Am. Stat. Assoc.* 83, 596–610. doi: 10.1080/01621459.1988.10478639
- Cox, D. R., and Hinkley, D. V. (1974). *Theoretical Statistics*. Boca Raton, FL: Chapman & Hall.
- Fitzmaurice, G., and Molenberghs, G. (2008). “Advances in longitudinal data analysis: a historical perspective,” in *Longitudinal Data Analysis: A Handbook of Modern Statistical Methods*, eds G. M. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs (London: Chapman & Hall/CRC Press), 3–30. doi: 10.1201/9781420011579.pt1
- Fitzmaurice, G., Laird, N., and Ware, J. (2004). *Applied Longitudinal Analysis*. New York, NY: A John Wiley & Son, Inc.
- Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models, 3rd Edn*. Thousand Oaks, CA: Sage Publications.
- Hastie, T., and Tibshirani, R. J. (1986). Generalized additive models. *Stat. Sci.* 1, 297–310. doi: 10.1214/ss/1177013604
- Hurvich, C. M., and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika* 76, 297–307. doi: 10.1093/biomet/76.2.297
- Kuk, A. Y. C., and Cheng, Y. W. (1997). The Monte Carlo Newton-Raphson algorithm. *J. Stat. Comput. Simul.* 59, 233–250. doi: 10.1080/00949657708811858
- Laird, N. M., and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* 38, 963–974. doi: 10.2307/2529876
- Long, J. D., Loeber, R., and Farrington, D. P. (2009). Marginal and random intercepts models for longitudinal binary data with examples from criminology. *Multivar. Behav. Res.* 44, 28–58. doi: 10.1080/00273170802620071
- Long, J. D., and Ryoo, J. (2010). Using fractional polynomials to model nonlinear trends in longitudinal data. *Br. J. Math. Stat. Psychol.* 63, 177–203. doi: 10.1348/000711009X431509
- McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *J. Am. Stat. Assoc.* 92, 162–170. doi: 10.1080/01621459.1997.10473613
- Molenberghs, G., and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. New York, NY: Springer.
- Morrell, C. H., Pearson, J. D., and Brant, L. J. (1997). Linear transformations of linear mixed-effects models. *Am. Stat.* 51, 338–343. doi: 10.1080/00031305.1997.10474409
- Nelder, J. A., and Wedderburn, R. W. M. (1972). Generalized linear models. *J. R. Stat. Soc. Ser. A* 135, 370–384. doi: 10.2307/2344614
- Pinheiro, J. C., and Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed effects model. *J. Comput. Graph. Stat.* 4, 12–35. doi: 10.1080/10618600.1995.10474663
- Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-effects Models in S and S-plus*. New York, NY: Springer-Verlag.
- Raftery, A. E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika* 83, 251–266. doi: 10.1093/biomet/83.2.251
- Royston, P., and Altman, D. G. (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modeling. *Appl. Stat.* 43, 429–453. doi: 10.2307/2986270
- Ryoo, J. (2011). Model selection with the linear mixed model for longitudinal data. *Multivar. Behav. Res.* 46, 598–624. doi: 10.1080/00273171.2011.589264
- Ryoo, J. H., Konold, T. R., Long, J. D., Molfese, V. J., and Zhou, X. (2017). Nonlinear growth mixture models with fractional polynomials: an illustration with early childhood mathematics ability. *Struct. Equ. Model. Multidisc. J.* doi: 10.1080/10705511.2017.1335206. [Epub ahead of print].
- Schall, R. (1991). Estimation in generalized linear models with random effects. *Biometrika* 78, 719–727. doi: 10.1093/biomet/78.4.719
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136
- Sterba, S. K. (2014). Fitting nonlinear latent growth models with individually-varying time points. *Struct. Equat. Model.* 21, 630–647. doi: 10.1080/10705511.2014.919828
- Stiratelli, R., Laird, N. M., and Ware, J. H. (1984). Random effects models for serial observations with binary response. *Biometrics* 40, 961–971. doi: 10.2307/2531147
- Wahba, G. (1990). *Spline Models for Observational Data*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman & Hall/CRC.
- Zeger, S. L., and Karim, M. R. (1991). Generalized linear models with random effects: a Gibbs sampling approach. *J. Am. Stat. Assoc.* 86, 79–86. doi: 10.1080/01621459.1991.10475006

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ryoo, Long, Welch, Reynolds and Swearer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.