Psychometrika (2026), **0**, 1–16
doi:10.1017/psy.2026.10136

PSYCHOMETRIC SOCIETY

THEORY AND METHODS

A-Optimal Designs for Rasch Poisson Gamma Counts Models

Heinz Holling[1] and Rainer Schwabe[2]

[1]Department of Psychology, University of Münster, Germany; [2]Institute for Mathematical Stochastics, Otto von Guericke University Magdeburg, Germany

Corresponding author: Heinz Holling; Email: holling@uni-muenster.de

(Received 6 March 2026 UTC; revised 18 June 2026 UTC; accepted 29 June 2026 UTC)

Abstract

In this article, we develop A-optimal designs for a longitudinal Rasch Poisson Gamma counts model (RPGCM). The model is particularly well suited for repeated testing designs in educational and psychological research when the outcomes are count data. Within the RPGCM framework, test scores depend on two types of parameters: respondents' abilities, which are modeled as gamma-distributed random effects, and item difficulties, which are treated as fixed parameters. To account for dependencies among repeated measurements, we incorporate a compound symmetry structure into the RPGCM. Parameter estimation is performed using maximum quasi-likelihood methods. Based on this framework, we derive A-optimal designs that allow for efficient estimation of mean abilities and commonly used functions thereof. In analogy to classical ANOVA models, no constraints are imposed on the mean structure. The resulting A-optimal designs are substantially less restrictive than commonly used D-optimal designs. Moreover, the derived designs are also valid for a more general class of longitudinal Poisson Gamma models with repeated measurements. These models can be applied to analyze growth curves for count data in other disciplines, including epidemiology, medicine, and criminology.

Keywords: A-optimal design; Rasch Poisson Gamma counts model; growth curve model

1. Introduction

The development of the optimal designs presented in this article was motivated by research on retest effects in cognitive abilities, which have been a central topic in educational research, psychology, and neuroscience for many years. Retest effects, also known as practice or test effects, denote systematic changes in performance that occur when individuals complete the same or equivalent cognitive ability tests repeatedly under comparable conditions. Several meta-analyses have demonstrated that repeated assessments of core cognitive abilities typically lead to substantial score improvements that follow a nonlinear trajectory characterized by diminishing gains across successive test administrations.

For example, a comprehensive longitudinal multilevel meta-analysis by Scharfen et al. (2018) examined retest effects across a large number of cognitive ability tests with up to four test administrations. Table 1 summarizes the results aggregated across tests assessing reasoning, memory, processing speed, and divergent thinking, the primary dimensions of intelligence according to the Berlin model of intelligence structure (BIS; Jäger, 1982). As shown in Table 1, statistically significant successive retest

© The Author(s), 2026. Published by Cambridge University Press on behalf of Psychometric Society.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

Table 1. Summary of retest effects reported by Scharfen et al. (2018)

Comparison	Number of effect sizes	Standardized mean change	*p*
1 vs. 2	566	0.327	< 0.001
1 vs. 3	181	0.495	< 0.001
1 vs. 4	39	0.516	< 0.001
2 vs. 3	181	0.169	< 0.001
3 vs. 4	39	0.021	0.319

effects were observed up to the third test administration. However, in some tests that measure more specific cognitive abilities such as verbal performance or memory, significant score increases were also found between the third and fourth test administrations.

Furthermore, the meta-analysis revealed substantial heterogeneity in retest effects across samples. Such individual differences must be taken into account to ensure valid interpretation of test outcomes, particularly when test-takers differ in their prior testing experience.

Longitudinal studies of cognitive abilities are typically time-consuming and resource-intensive. Recruiting participants is often difficult, and dropout rates tend to increase as studies progress, especially when tasks are tedious, cognitively demanding, or insufficiently motivating. From a practical and ethical perspective, it is therefore desirable to minimize the total number of test items administered across all measurement occasions while ensuring sufficient statistical efficiency for estimating the parameters of interest. To address this problem, we develop optimal designs for the longitudinal Rasch Poisson Gamma counts model (RPGCM).

The RPGCM is based on the Rasch Poisson counts model (RPCM), which was originally introduced by Rasch (1960) for the analysis of count data and is often regarded as one of the earliest item response models (Baghaei & Doebler, 2019). Like the more widely known logistic Rasch model, the RPCM includes both an ability parameter and an item difficulty parameter. Rasch applied this model to oral reading performance, measured by the number of words read and the number of errors, which followed a Poisson distribution. More generally, the model may be applicable to a wide range of ability test data, as three of the four BIS intelligence dimensions, memory, processing speed, and divergent thinking, yield count data.

The original RPCM was specified as a fixed-effects model. Subsequent developments introduced random effects for both person and item parameters. In particular, gamma-distributed person parameters have received considerable attention (e.g., Graßhoff et al., 2020; Jansen & van Duijn, 1992; Owen, 1969). As the conjugate prior of the Poisson distribution, the gamma distribution possesses several attractive statistical properties compared with alternative specifications such as the log-normal distribution (Jansen, 1986). In particular, the resulting marginal response distribution is a negative binomial distribution, which naturally accommodates the overdispersion often encountered in count data (e.g., Hung, 2012). In some studies, item parameters were also modeled as random effects, for example, by means of a Dirichlet distribution (Jansen, 1997) or a log-normal distribution (Jansen, 1986). A further line of development concerns the introduction of discrimination parameters. Myszkowski and Storme (2021) extended the RPCM beyond the one-parameter framework by incorporating discrimination effects. However, their approach retained the assumption of equidispersion inherent in the Poisson distribution. This limitation was addressed by Beisemann (2022), who proposed a Conway–Maxwell–Poisson (CMP) extension of the RPCM. In addition to item discrimination parameters, the CMP approach permits the modeling of both overdispersion and underdispersion, thereby substantially increasing the flexibility of the model.

The introduction of approaches using finite mixture models and latent classes represents a further extension of the RPCM. For example, van Duijn and Böckenholt (1995) demonstrated that repeated count observations may be explained by transitions between discrete latent classes. Similarly, Verhelst

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

and Kamphuis (2009) argued that speed-test data may require mixture formulations of the RPCM to adequately capture latent heterogeneity.

Another important development, which also does not assume a single continuous latent variable, is the development of multidimensional Poisson count models, which allow for the simultaneous modeling of multiple latent variables and their interrelationships. For example, Wedel et al. (2003) proposed multivariate factor-analytic Poisson models. Skrondal and Rabe-Hesketh (2004) embedded the RPCM and related extensions within the general framework of generalized linear latent and mixed models. This framework provides a unified likelihood-based approach to multilevel, longitudinal, and structural equation models for count data. To date, the most comprehensive multidimensional extension appears to be the multidimensional two-parameter CMP model proposed by Beisemann et al. (2025). This framework combines item discrimination parameters with flexible dispersion modeling and thus generalizes both multidimensional IRT models and generalized count models. A particularly innovative feature is the incorporation of rotation and regularization techniques to obtain sparse and interpretable loading structures. As a result, multidimensional count models become substantially more interpretable and consequently more useful for applied psychometric research.

To develop optimal designs for studies analyzing retest effects of cognitive abilities, we begin with a longitudinal version of the RPGCM. This model is particularly suitable when the item parameters are known, for example, when items are drawn from calibrated tests. Moreover, as outlined above, the gamma distribution constitutes an advantageous choice for the ability parameters. In longitudinal settings, the RPGCM must be extended, since dependencies between repeated measurements within individuals have to be taken into account. To this end, we adopt a longitudinal RPGCM based on the established concept of compound symmetry, following Parsamaram et al. (2026). An analogous framework for the analysis of metric repeated measurements using linear mixed models with multivariate normal random effects was proposed by Freise et al. (2023). Should more sophisticated Poisson count models be required for the analysis of retest effects in cognitive abilities, the development of corresponding optimal designs remains an important topic for future research.

The remainder of the article is structured as follows. Section 2 describes the longitudinal RPGCM and outlines parameter estimation using maximum quasi-likelihood. Section 3 derives A-optimal designs for estimating mean abilities and common functions thereof under an unstructured mean model. Furthermore, these designs are illustrated using some growth patterns with four measurement occasions. In the final section, we briefly discuss the differences between the A-optimal designs derived in the present study and the corresponding D-optimal designs developed by Parsamaram et al. (2026). The article concludes with a summary and suggestions for future research.

2. Longitudinal Rasch Poisson Gamma counts model

This section outlines the longitudinal RPGCM and the estimation of its parameters. A more detailed technical description can be found in Parsamaram et al. (2026).

2.1. Model specification

Let Y_{ijk} denote the response of test person i to item k administered at the jth measurement occasion taking place at time t_j. The response is assumed to follow a Poisson distribution with mean $\theta_{ij}\sigma_{jk}$, where θ_{ij} denotes the ability of person i at time t_j, $i = 1, \ldots, N$, $j = 1, \ldots, J$, and σ_{jk} denotes the easiness of item k administered at time t_j, $k = 1, \ldots, n_j$.

The subjects are assumed to receive the same items with easiness σ_{jk} and, in particular, the same total number n_j at occasion j, $\sum_{j=1}^{J} n_j = n$, each. Furthermore, the item easiness parameters σ_{jk} are assumed to be known, which does not constitute a restriction when calibrated tests are employed, and the mean item easiness $\bar{\sigma} = \frac{1}{n}\sum_{j=1}^{J}\sum_{k=1}^{n_j}\sigma_{jk}$ is fixed. This condition is trivially satisfied if the same item parameters are used across time points. When automatically generated, rule-based items are administered, this assumption typically does not pose practical problems.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

The primary objective of the analysis is the estimation of the development of ability over time, represented by a common mean learning curve $\theta(t)$ evaluated at $t_1, \ldots, t_J$. Optimal designs allow $\theta(t)$ to be estimated with maximal precision as a function of the allocation of items across measurement occasions, that is, the proportions n_j/n and the items σ_{jk}, $k = 1, \ldots, n_j$. In empirical applications, the same number $n_j = n/J$ of items is often presented with common mean item easiness $\bar{\sigma}_j = \frac{1}{n_j}\sum_{k=1}^{n_j}\sigma_{jk}$ at each measurement occasion, that is, $n_j = n/J$ and $\bar{\sigma}_j = \bar{\sigma}$. This practice raises the question of whether such a design is efficient or whether alternative allocations can reduce the total number of items required without compromising statistical accuracy.

Abilities θ_{ij} are modeled as random effects. Specifically, θ_{ij} is expressed as a realization of

$$\Theta_{ij} = \Lambda_{ij}\,\theta(t_j),$$

where Λ_{ij} is a gamma-distributed latent variable capturing individual deviations from the mean trajectory $\theta(t)$.

To specify a longitudinal RPGCM based on compound symmetry, analogous to linear mixed models with multivariate normal random effects, a permanent person effect and time-specific person effects are introduced. Accordingly, the latent variable Λ_{ij} is decomposed as

$$\Lambda_{ij} = \Gamma_{i0} + \Gamma_{ij}, \qquad j = 1, \ldots, J,$$

where the time-specific effects $\Gamma_{i1}, \ldots, \Gamma_{iJ}$ are assumed to be independently and identically distributed as $\Gamma(\alpha, \tau)$, the permanent person effect $\Gamma_{i0} \sim \Gamma(\alpha_0, \tau)$ shares the same scale parameter τ, and all components are mutually independent.

Under these assumptions, $\Lambda_{ij} \sim \Gamma(\alpha_0 + \alpha, \tau)$, and the vector $\boldsymbol{\Lambda}_i = (\Lambda_{i1}, \ldots, \Lambda_{iJ})^\top$ follows a multivariate gamma distribution with correlated components induced by the permanent person effect.

By standardizing the latent effects by $\tau = 1/(\alpha_0 + \alpha)$ such that $\mathrm{E}(\Lambda_{ij}) = 1$, the marginal mean of the responses is given by

$$\mathrm{E}(Y_{ijk}) = \theta(t_j)\sigma_{jk}.$$

Thus, the growth curve $\theta(t)$ represents the marginal mean ability, satisfying $\mathrm{E}(\Theta_{ij}) = \theta(t_j)$.

Furthermore, $\mathrm{Var}(\Lambda_{ij}) = \tau$ and $\mathrm{Cov}(\Lambda_{ij}, \Lambda_{ij'}) = \rho\tau$, where $\rho = \alpha_0\tau$ denotes the intraclass correlation coefficient. Consequently,

$$\mathrm{Cov}(\boldsymbol{\Lambda}_i) = \tau\left[(1-\rho)\mathbf{I}_J + \rho\mathbf{1}_J\mathbf{1}_J^\top\right].$$

To model the mean structure, we adopt an unstructured one-way ANOVA-type specification $\theta(t_j) = \theta_j$ embedded in a generalized linear mixed model with a log link,

$$\theta_j = \exp(\beta_j), \qquad j = 1, \ldots, J,$$

with vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_J)$ of mean abilities over time.

In the model outlined above, the latent variables Γ_{ij} are assumed to be identically distributed across measurement occasions. Consequently, the variances of the ability parameters Θ_{ij} remain constant over time. For many studies investigating retest effects in cognitive abilities, this assumption appears plausible. We inspected the studies included in the meta-analysis by Scharfen et al. (2018) and found that the variances of the observed test scores were generally stable across repeated measurement occasions. However, gains in ability may not be strictly additive. For example, individuals with higher initial abilities may exhibit different growth trajectories than individuals with lower initial abilities. In such situations, the assumption of constant variance may be violated. To accommodate this possibility, we extend the model by allowing the latent variables to follow occasion-specific gamma distributions, $\Gamma_{ij} \sim \Gamma(\alpha_j, \tau), j = 1, \ldots, J$, thereby permitting the variances of the ability parameters to vary over time. The extended model, together with the corresponding optimal designs, is presented in the Appendix. It should be noted, however, that the resulting A-optimal designs depend on the variance parameters of the random effects, whereas the designs obtained under the assumption of identically distributed

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

Γ_{ij} do not. Thus, they require substantially more prior information, which may be difficult to obtain in practical applications.

2.2. Parameter estimation

In the generalized linear mixed model setup, the focus is on estimating the parameter vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)^\top$. Even if the parameters τ and ρ are assumed to be known, the likelihood involves integrals which are analytically difficult to handle. Consequently, we employ maximum quasi-likelihood estimation (Wedderburn, 1974), which relies only on the first- and second-order marginal moments of the response vector.

The maximum quasi-likelihood estimator $\hat{\boldsymbol{\beta}}_{\text{MQL}}$ is defined as the solution to

$$\sum_{i=1}^{N} \mathbf{D}_i^\top \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0},$$

where $\boldsymbol{\mu}_i = \mathrm{E}(\mathbf{Y}_i)$, $\mathbf{V}_i = \mathrm{Cov}(\mathbf{Y}_i)$, and $\mathbf{D}_i = \partial \boldsymbol{\mu}_i / \partial \boldsymbol{\beta}^\top$. The explicit expressions for these terms are provided in the Appendix.

The quasi-information matrix is given by

$$\mathbf{M}_{\text{Q}} = \sum_{i=1}^{N} \mathbf{D}_i^\top \mathbf{V}_i^{-1} \mathbf{D}_i.$$

Similar to the situation in maximum likelihood estimation, the asymptotic covariance matrix of the maximum quasi-likelihood estimator $\hat{\boldsymbol{\beta}}_{\text{MQL}}$ is proportional to the inverse $\mathbf{M}_{\text{Q}}^{-1}$ of the quasi-information matrix.

As shown in Theorem 4.2 of Parsamaram et al. (2026), the quasi-information simplifies, under a saturated mean model and identical test conditions across individuals, to

$$\mathbf{M}_{\text{Q}} = N \left[\mathrm{diag}\left(\frac{1}{\theta_j \sum_{k=1}^{n_j} \sigma_{jk}} \right) + \tau \left((1-\rho)\mathbf{I}_J + \rho \mathbf{1}_J \mathbf{1}_J^\top \right) \right]^{-1}.$$

In the fixed-effects RPCM, the maximum quasi-likelihood estimator coincides with the common maximum likelihood estimator, and the likelihood-based Fisher information for $\boldsymbol{\beta}$ is obtained by formally setting $\tau = 0$ in the quasi-information.

3. A-optimal design

Parsamaram et al. (2026) developed locally D-optimal designs for estimating mean abilities in this model. According to their results, the D-optimal designs, which minimize the volume of the confidence ellipsoid of the parameters, require not only the specification of the parameter vector of interest, $\boldsymbol{\theta}$, but also prior knowledge of the variance τ and the correlation ρ. Such information, however, is rarely available in practice. In contrast, A-optimal designs, which minimize the sum of the variances of the parameter estimators or specified functions thereof, rely on substantially weaker assumptions and are therefore better suited to practical applications. We will discuss this topic in more detail in the fifth section.

The standardized (per response) quasi-information matrix is

$$\mathbf{M}_{\text{Q}}(\xi) = \left[\mathrm{diag}\left(\frac{1}{\theta_j w_j \tilde{\sigma}} \right) + n\tau \left((1-\rho)\mathbf{I}_J + \rho \mathbf{1}_J \mathbf{1}_J^\top \right) \right]^{-1},$$

where $\xi = (w_1, \ldots, w_J)$ denotes the vector of proportions $w_j = \sum_{k=1}^{n_j} \sigma_{jk} / \sum_{j=1}^{J} \sum_{k=1}^{n_j} \sigma_{jk}$ of item easiness allocated to the measurement occasions $j = 1, \ldots, J$, $\sum_{j=1}^{J} w_j = 1$.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

The vector ξ of proportions of easiness constitutes the test design and can be chosen to optimize certain characteristics of the estimation. These proportions of easiness can be interpreted as follows: When all items have the same easiness parameters $\sigma_{jk} = \bar{\sigma} = \frac{1}{n}\sum_{j=1}^{J}\sum_{k=1}^{n_j}\sigma_{jk}$, then $w_j = n_j/n$ is the proportion of items presented at occasion j. Alternatively, when the same number $n_j = n/J$ of items is presented at each occasion, then $w_j = \frac{1}{J\bar{\sigma}}\bar{\sigma}_j$ is proportional to the mean easiness $\bar{\sigma}_j = \frac{1}{n_j}\sum_{k=1}^{n_j}\sigma_{jk}$ of items to be administered at occasion j.

Note that, in the unstructured mean model, the proportions w_j are only involved in the first term in the representation of the quasi-information. In particular, the asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}$ equals $c_1\mathrm{diag}(1/(\theta_j w_j)) + \mathbf{C}_2$ for some generic constant terms $c_1 > 0$ and $\mathbf{C}_2$. This explains why the A-optimal weights neither depend on the variance parameters τ and ρ nor on $\bar{\sigma}$ and n.

We derive A-optimal designs for the most relevant parameters of the unstructured longitudinal RPGCM. We begin with the basic parameter vector $\boldsymbol{\beta}$ of the unstructured mean model and subsequently develop A-optimal designs for the mean ability vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_J)$ as well as for the square-root-transformed mean abilities $\boldsymbol{d} = (d_1, \ldots, d_J)$, $d_j = \sqrt{\theta_j}$.

In many applications, hypotheses are formulated in terms of contrasts rather than absolute parameter values. Accordingly, we also derive A-optimal designs for commonly used contrasts, including differences relative to a baseline (e.g., $\theta_j - \theta_1, j = 2, \ldots, J$) and successive differences (e.g., $\theta_j - \theta_{j-1}, j = 2, \ldots, J$).

All A-optimality criteria considered in the following can be reduced to the minimization of an objective function of the form

$$C(\xi) = \sum_{j=1}^{J}\frac{h_j}{w_j}, \qquad \text{for some constants } h_j > 0,$$

in the proportions of easiness w_j allocated to measurement occasion j. Minimization of this function yields the optimal weights

$$w_j^* = \frac{\sqrt{h_j}}{\sum_{k=1}^{J}\sqrt{h_k}}, \qquad j = 1, \ldots, J.$$

This general result considerably simplifies the derivation of the optimal designs presented below.

As an illustration, consider A-optimal designs for the parameters $\beta_j = \log(\theta_j), j = 1, \ldots, J$. In this case, the trace of the asymptotic covariance matrix has to be minimized. This is equivalent to minimizing

$$C_{\beta_1,\ldots,\beta_J}(\xi) = \sum_{j=1}^{J}\frac{1}{\theta_j w_j}.$$

Setting $h_j = 1/\theta_j$ immediately yields the A-optimal weights

$$w_{j,\beta_j}^* = \frac{\sqrt{1/\theta_j}}{\sum_{k=1}^{J}\sqrt{1/\theta_k}}, \quad j = 1, \ldots, J$$

for estimating the parameters $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)$.

3.1. A-optimal designs for β

For contrasts relative to the baseline, $\beta_j - \beta_1, j = 2, \ldots, J$, the objective function

$$C_{\beta_2-\beta_1,\ldots,\beta_J-\beta_1}(\xi) = \frac{J-1}{\theta_1 w_1} + \sum_{j=2}^{J}\frac{1}{\theta_j w_j}$$

has to be minimized, and for successive differences, $\beta_2 - \beta_1, \ldots, \beta_J - \beta_{J-1}$, the objective function

$$C_{\beta_2-\beta_1,\ldots,\beta_J-\beta_{J-1}}(\xi) = \frac{1}{\theta_1 w_1} + \sum_{j=2}^{J-1}\frac{2}{\theta_j w_j} + \frac{1}{\theta_J w_J}$$

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

has to be minimized. The resulting A-optimal weights for parameters and contrasts of $\boldsymbol{\beta}$ are listed in lines 1–3 of Table 2.

3.2. A-optimal designs for θ

In applications of longitudinal Poisson models, the primary parameter of interest is often the trajectory θ. In repeated testing, this may correspond, for example, to the expected number of correctly processed stimuli, in epidemiology to the expected number of rare disease cases, and in criminology to the expected number of aggressive acts.

For estimation of $\boldsymbol{\theta}$, the plug-in estimators $\hat{\theta}_j = \exp(\hat{\beta}_j)$ may be used. By the δ-method, the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$ equals $c_1 \text{diag}(\theta_j/w_j) + \mathbf{C}_2$. Again, applying the general objective function, the following A-optimal designs for $\boldsymbol{\theta}$ can be derived.

For mean abilities θ_j, $j = 1, \ldots, J$, the objective function is

$$C_{\theta_1,\ldots,\theta_J}(\xi) = \sum_{j=1}^{J} \frac{\theta_j}{w_j}.$$

For baseline differences $\theta_j - \theta_1$, $j = 2, \ldots, J$, the objective function is

$$C_{\theta_2-\theta_1,\ldots,\theta_J-\theta_1}(\xi) = \frac{(J-1)\theta_1}{w_1} + \sum_{j=2}^{J} \frac{\theta_j}{w_j}.$$

For successive differences $\theta_2 - \theta_1, \ldots, \theta_J - \theta_{J-1}$, the objective function is

$$C_{\theta_2-\theta_1,\ldots,\theta_J-\theta_{J-1}}(\xi) = \frac{\theta_1}{w_1} + \sum_{j=2}^{J-1} \frac{2\theta_j}{w_j} + \frac{\theta_J}{w_J}.$$

The corresponding A-optimal weights are given in lines 4–6 of Table 2.

Ratios of expected values are often more stable and interpretable than differences. Again, plug-in estimators can be used and the asymptotic covariance is obtained by the δ-method. For ratios relative to the baseline, θ_j/θ_1, $j = 2, \ldots, J$, the objective function is given by

$$C_{\theta_2/\theta_1,\ldots,\theta_J/\theta_1}(\xi) = \frac{1}{\theta_1^2}\left(\frac{\sum_{j=2}^{J}\theta_j^2/\theta_1}{w_1} + \sum_{j=2}^{J}\frac{\theta_j}{w_j}\right),$$

and for successive ratios, it is

$$C_{\theta_2/\theta_1,\ldots,\theta_J/\theta_{J-1}}(\xi) = \frac{\theta_2^2/\theta_1^3}{w_1} + \sum_{j=2}^{J-1}\frac{\theta_{j+1}^2/\theta_j^3 + \theta_j/\theta_{j-1}^2}{w_j} + \frac{\theta_J/\theta_{J-1}^2}{w_J}.$$

The A-optimal weights are in lines 7 and 8 of Table 2.

3.3. A-optimal designs for $d = \sqrt{\theta}$

Poisson variables with large expected values are well approximated by a normal distribution after square root transformation, $d_j = \sqrt{\theta_j}$, with $\text{E}(\hat{d}_j) \approx \sqrt{\theta_j}$ and $\text{Var}(\hat{d}_j) \approx 1/4$. Even for smaller values of θ_j, this variance-stabilizing transformation substantially reduces heteroscedasticity.

Applying the δ-method once more, the asymptotic covariance matrix of the plug-in estimator $\hat{\boldsymbol{d}} = (\hat{d}_1, \ldots, \hat{d}_J)$ reduces to $c_1 \text{diag}(1/w_j) + \mathbf{C}_2$. Ignoring the constant terms, the resulting covariance structure is identical to that in a one-way ANOVA. Consequently, designs that are A-optimal in the one-way layout are also A-optimal for $\boldsymbol{d}$ and related linear contrasts (see lines 9–11 in Table 2). In particular, for $\boldsymbol{d}$ itself, the optimal design is uniform on the measurement occasions, that is, $w_j^* = 1/J$, $j = 1, \ldots, J$. For ratios of $\boldsymbol{d}$, A-optimal designs are given in Table 2, lines 12 and 13.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

Table 2. A-optimal weights for parameters and contrasts

	Parameter	w_1^*	$w_j^*, j = 2, \ldots, J-1$	w_J^*
1	β_j		$\frac{\sqrt{1/\theta_j}}{\sum_{k=1}^{J}\sqrt{1/\theta_k}}$	
2	$\beta_j - \beta_1$	$\frac{\sqrt{(J-1)/\theta_1}}{\sqrt{(J-1)/\theta_1+\sum_{k=2}^{J}\sqrt{1/\theta_k}}}$	$\frac{\sqrt{1/\theta_j}}{\sqrt{(J-1)/\theta_1+\sum_{k=2}^{J}\sqrt{1/\theta_k}}}$	
3	β_j/β_{j-1}	$\frac{\sqrt{1/\theta_1}}{\sqrt{1/\theta_1+\sum_{k=2}^{J-1}\sqrt{2/\theta_k}+\sqrt{1/\theta_J}}}$	$\frac{\sqrt{2/\theta_j}}{\sqrt{1/\theta_1+\sum_{k=2}^{J-1}\sqrt{2/\theta_k}+\sqrt{1/\theta_J}}}$	$\frac{\sqrt{1/\theta_J}}{\sqrt{1/\theta_1+\sum_{k=2}^{J-1}\sqrt{2/\theta_k}+\sqrt{1/\theta_J}}}$
4	θ_j		$\frac{\sqrt{\theta_j}}{\sum_{k=1}^{J}\sqrt{\theta_k}}$	
5	$\theta_j - \theta_1$	$\frac{\sqrt{(J-1)\theta_1}}{\sqrt{(J-1)\theta_1+\sum_{k=2}^{J}\sqrt{\theta_k}}}$	$\frac{\sqrt{\theta_j}}{\sqrt{(J-1)\theta_1+\sum_{k=2}^{J}\sqrt{\theta_k}}}$	
6	$\theta_j - \theta_{j-1}$	$\frac{\sqrt{\theta_1}}{\sqrt{\theta_1}+\sum_{k=2}^{J-1}\sqrt{2\theta_k}+\sqrt{\theta_J}}$	$\frac{\sqrt{2\theta_j}}{\sqrt{\theta_1}+\sum_{k=2}^{J-1}\sqrt{2\theta_k}+\sqrt{\theta_J}}$	$\frac{\sqrt{\theta_J}}{\sqrt{\theta_1}+\sum_{k=2}^{J-1}\sqrt{2\theta_k}+\sqrt{\theta_J}}$
7	θ_j/θ_1	$\frac{\sqrt{\sum_{k=2}^{J}\theta_k^2/\theta_1}}{\sqrt{\sum_{k=2}^{J}\theta_k^2/\theta_1+\sum_{k=2}^{J}\sqrt{\theta_k}}}$	$\frac{\sqrt{\theta_j}}{\sqrt{\sum_{k=2}^{J}\theta_k^2/\theta_1+\sum_{k=2}^{J}\sqrt{\theta_k}}}$	
8	θ_j/θ_{j-1}	$\frac{\sqrt{\theta_2^2/\theta_1^3}}{\sqrt{\theta_2^2/\theta_1^3+\sum_{k=2}^{J-1}\sqrt{\theta_{k+1}^2/\theta_k^3+\theta_k/\theta_{k-1}^2}+\sqrt{\theta_J/\theta_{J-1}^2}}}$	$\frac{\sqrt{\theta_{j+1}^2/\theta_j^3+\theta_j/\theta_{j-1}^2}}{\sqrt{\theta_2^2/\theta_1^3+\sum_{k=2}^{J-1}\sqrt{\theta_{k+1}^2/\theta_k^3+\theta_k/\theta_{k-1}^2}+\sqrt{\theta_J/\theta_{J-1}^2}}}$	$\frac{\sqrt{\theta_J/\theta_{J-1}^2}}{\sqrt{\theta_2^2/\theta_1^3+\sum_{k=2}^{J-1}\sqrt{\theta_{k+1}^2/\theta_k^3+\theta_k/\theta_{k-1}^2}+\sqrt{\theta_J/\theta_{J-1}^2}}}$
9	d_j		$1/J$	
10	$d_j - d_1$	$\frac{\sqrt{J-1}}{\sqrt{J-1}+J-1}$	$\frac{1}{\sqrt{J-1}+J-1}$	
11	$d_j - d_{j-1}$	$\frac{1}{2+(J-2)\sqrt{2}}$	$\frac{\sqrt{2}}{2+(J-2)\sqrt{2}}$	$\frac{1}{2+(J-2)\sqrt{2}}$
12	d_j/d_1	$\frac{\sqrt{\sum_{k=2}^{J}\theta_k/\theta_1}}{\sqrt{\sum_{k=2}^{J}\theta_k/\theta_1}+J-1}$	$\frac{1}{\sqrt{\sum_{k=2}^{J}\theta_k/\theta_1}+J-1}$	
13	d_j/d_{j-1}	$\frac{\sqrt{\theta_2/\theta_1^2}}{\sqrt{\theta_2/\theta_1^2}+\sum_{k=2}^{J-1}\sqrt{\theta_{k+1}/\theta_k^2+1/\theta_{k-1}}+\sqrt{1/\theta_{J-1}}}$	$\frac{\sqrt{\theta_{j+1}/\theta_j^2+1/\theta_{j-1}}}{\sqrt{\theta_2/\theta_1^2}+\sum_{k=2}^{J-1}\sqrt{\theta_{k+1}/\theta_k^2+1/\theta_{k-1}}+\sqrt{1/\theta_{J-1}}}$	$\frac{\sqrt{1/\theta_{J-1}}}{\sqrt{\theta_2/\theta_1^2}+\sum_{k=2}^{J-1}\sqrt{\theta_{k+1}/\theta_k^2+1/\theta_{k-1}}+\sqrt{1/\theta_{J-1}}}$

3.4. Numerical results

All A-optimal designs for $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, and their contrasts are locally optimal designs. They depend on the relative change in the mean trajectory between measurement occasions, but not on the overall mean. The latter is in accordance with the property that the combination of ability and easiness may be scaled at will. For $\boldsymbol{d}$ and the corresponding difference contrasts, however, the optimal designs are independent of the expected values and, hence, globally optimal for all parameter values. They are functions of the number J of measurement occasions only, which substantially simplifies experimental planning.

Table 3 illustrates the derived A-optimal designs for four example growth patterns, each with $J = 4$ repeated measurements and a baseline with a mean of 4. The first example corresponds to a constant value for the growth curve, which may be considered as a null hypothesis setup of no retest effects. The second example exhibits growth increments motivated by the meta-analysis of retest effects reported by Scharfen et al. (2018). The third example represents a linear increase in expected values, whereas the fourth example follows an exponential growth curve, as commonly observed in other application domains such as epidemiology.

The results demonstrate once again that A-optimal designs for square-root-transformed parameters and their differences are independent of the actual expected values. In general, a uniform allocation of items across measurement occasions is suboptimal and can lead to a loss of efficiency, in particular, when a baseline measurement serves as a reference. This topic will be analyzed in more detail in the next section.

It should be noted that alternative expected values, growth rates, or numbers of measurement occasions may yield different A-optimal designs. However, the corresponding optimal weights can always be obtained directly from the closed-form expressions derived above.

4. A-efficiency

In practice, experimenters frequently employ a uniform allocation of observations across measurement occasions, either for convenience or due to limited prior knowledge about expected values. The previously derived A-optimal designs for the parameter vectors $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, and $\boldsymbol{d} = \sqrt{\boldsymbol{\theta}}$ provide a minimal estimator variance. In this section, we analyze the gain of these optimal designs compared to uniform designs $\xi_{\text{uniform}} = (1/J, \ldots, 1/J)$. Commonly used uniform designs allocate an equal proportion of observations with equal mean easiness to each measurement occasion. Specifically, we consider the A-efficiency of the uniform design compared to the previously obtained A-optimal designs for the parameters $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, and $\boldsymbol{d}$ as well as the contrasts thereof presented in Table 2.

The efficiency, which will be considered here, is defined by

$$\text{Eff}_A(\xi_{\text{uniform}}; \boldsymbol{\delta}) = \frac{C_{\boldsymbol{\delta}}(\xi^*_{\boldsymbol{\delta}})}{C_{\boldsymbol{\delta}}(\xi_{\text{uniform}})},$$

for the uniform design ξ_{uniform}, where $\boldsymbol{\delta}$ stands for the parameter vector of interest and $\xi^*_{\boldsymbol{\delta}}$ is the corresponding A-optimal design.

The A-efficiency ranges between 0 and 1 and quantifies the amount of within-test-person information of the uniform design compared to the A-optimal design. Based on the efficiency, the corresponding A-deficiency can be computed as $\text{Def}_A(\xi_{\text{uniform}}; \boldsymbol{\delta}) = \text{Eff}_A(\xi_{\text{uniform}}; \boldsymbol{\delta})^{-1} - 1$, which represents the proportional increase in the total number n of items required when using design ξ_{uniform} instead of the A-optimal design $\xi^*_{\boldsymbol{\delta}}$ to achieve the same estimation precision while keeping the mean easiness $\bar{\sigma}$ fixed. In the case of the fixed-effects RPCM, the above A-deficiency may be interpreted likewise as the proportional increase in the number N of test persons required when the design ξ_{uniform} is used instead of the A-optimal design $\xi^*_{\boldsymbol{\delta}}$ to obtain the same value for the trace of the covariance matrix while the total easiness $n\bar{\sigma}$ per test person is kept fixed.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

Table 3. A-optimal weights for four sets of growth patterns with four measurement occasions each

Set	θ	$w^*_{\beta_j}$	$w^*_{\beta_j-\beta_1}$	$w^*_{\beta_j-\beta_{j-1}}$	$w^*_{\theta_j}$	$w^*_{\theta_j-\theta_1}$	$w^*_{\theta_j-\theta_{j-1}}$	$w^*_{\theta_j/\theta_1}$	$w^*_{\theta_j/\theta_{j-1}}$	$w^*_{d_j}$	$w^*_{d_j-d_1}$	$w^*_{d_j-d_{j-1}}$	$w^*_{d_j/d_1}$	$w^*_{d_j/d_{j-1}}$
1	4	0.25	0.37	0.21	0.25	0.37	0.21	0.37	0.21	0.25	0.37	0.21	0.37	0.21
	4	0.25	0.21	0.29	0.25	0.21	0.29	0.21	0.29	0.25	0.21	0.29	0.21	0.29
	4	0.25	0.21	0.29	0.25	0.21	0.29	0.21	0.29	0.25	0.21	0.29	0.21	0.29
	4	0.25	0.21	0.21	0.25	0.21	0.21	0.21	0.21	0.25	0.21	0.21	0.21	0.21
2	4	0.29	0.41	0.24	0.22	0.33	0.18	0.41	0.27	0.25	0.37	0.21	0.41	0.25
	5.31	0.25	0.20	0.29	0.25	0.22	0.29	0.19	0.30	0.25	0.21	0.29	0.20	0.30
	5.84	0.24	0.20	0.28	0.26	0.23	0.31	0.20	0.25	0.25	0.21	0.29	0.20	0.27
	6.06	0.23	0.19	0.19	0.27	0.23	0.22	0.20	0.17	0.25	0.21	0.21	0.20	0.18
3	4	0.36	0.49	0.30	0.16	0.25	0.13	0.51	0.36	0.25	0.37	0.21	0.50	0.33
	8	0.25	0.20	0.30	0.23	0.21	0.27	0.13	0.32	0.25	0.21	0.29	0.17	0.31
	12	0.21	0.16	0.25	0.28	0.25	0.33	0.16	0.21	0.25	0.21	0.29	0.17	0.23
	16	0.18	0.14	0.15	0.33	0.29	0.27	0.19	0.12	0.25	0.21	0.21	0.17	0.13
4	4	0.39	0.53	0.33	0.14	0.22	0.12	0.59	0.33	0.25	0.37	0.21	0.56	0.33
	8	0.28	0.21	0.33	0.20	0.18	0.23	0.09	0.33	0.25	0.21	0.29	0.15	0.33
	16	0.20	0.15	0.23	0.28	0.25	0.33	0.13	0.23	0.25	0.21	0.29	0.15	0.23
	32	0.14	0.11	0.12	0.39	0.35	0.33	0.18	0.12	0.25	0.21	0.21	0.15	0.12

Table 4. A-efficiency of the uniform design relative to the A-optimal design for the four sets of growth patterns with *four* measurement occasions as outlined in Table 3 (left) and with *seven* measurement occasions as outlined in the text (right)

Parameter	$J = 4$				$J = 7$			
	Set 1	Set 2	Set 3	Set 4	Set 1	Set 2	Set 3	Set 4
β_j	1.00	0.99	0.93	0.87	1.00	0.94	0.89	0.70
$\beta_j - \beta_1$	0.93	0.88	0.76	0.70	0.85	0.62	0.56	0.43
$\beta_j - \beta_{j-1}$	0.97	0.98	0.94	0.89	0.98	0.97	0.92	0.75
θ_j	1.00	0.99	0.94	0.87	1.00	0.97	0.93	0.70
$\theta_j - \theta_1$	0.93	0.97	0.99	0.94	0.85	0.98	0.96	0.75
$\theta_j - \theta_{j-1}$	0.97	0.96	0.92	0.89	0.98	0.94	0.92	0.75
θ_j/θ_1	0.93	0.88	0.73	0.61	0.85	0.60	0.51	0.24
θ_j/θ_{j-1}	0.97	0.96	0.88	0.89	0.98	0.83	0.81	0.75
d_j	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$d_j - d_1$	0.93	0.93	0.93	0.93	0.85	0.85	0.85	0.85
$d_j - d_{j-1}$	0.97	0.97	0.97	0.97	0.98	0.98	0.98	0.98
d_j/d_1	0.93	0.88	0.75	0.67	0.85	0.62	0.54	0.32
d_j/d_{j-1}	0.97	0.97	0.91	0.89	0.98	0.91	0.87	0.75

For the four data examples introduced in Table 3, we compute the A-efficiency of the uniform design relative to the A-optimal design for the parameters $\boldsymbol{\theta}$ and d, as well as for the contrasts considered above. The results are summarized in Table 4 (left). Furthermore, we compute the A-efficiency for another example of four sets of analogous growth patterns, but each with seven measurement occasions. The first set of $\boldsymbol{\theta}$ values is constant 4, the second consists of (4, 8, 11, 13, 14, 14.5, 14.75), the third of (4, 8, 12, 16, 20, 24, 28), and the fourth of (4, 8, 16, 32, 64, 128, 256); see Table 4 (right).

The results indicate that a uniform allocation of observations is generally suboptimal. A-optimal designs lead to reductions in variance, although these gains are small in several of the examples considered. In such situations, practitioners may rely on uniform designs with only minor loss in efficiency. However, substantially larger reductions are observed for parameters representing contrasts with respect to a reference category. For such contrasts, which are frequently of practical interest, A-optimal designs should therefore be preferred. For any specific situation arising in an empirical study, the formulas presented above allow the derivation of A-optimal designs and the corresponding A-efficiency.

Overall, these findings highlight the practical relevance of A-optimal design strategies for longitudinal and experimental studies when contrasts relative to a reference category are of primary interest.

5. D-optimal versus A-optimal designs for the longitudinal RPGCM

Recently, Parsamaram et al. (2026) developed locally D-optimal designs for estimating mean abilities in the RPGCM under compound symmetry of the covariance matrix of the random effects. This work motivated us to derive A-optimal rather than D-optimal designs here, although the latter constitute the most prominent class of optimal designs. While a D-optimal design maximizes the determinant of the Fisher information matrix and thereby minimizes the volume of the asymptotic confidence ellipsoid for the parameter vector, an A-optimal design minimizes the trace of the inverse information matrix, this means that it minimizes the average asymptotic variance of the parameter estimates. A

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

major reason for the predominance of D-optimal designs is the invariance of the D-criterion under nonsingular reparameterizations of the model; in particular, the designs do not depend on the scaling of the parameters. Hence, Parsamaram et al. (2026) needed to derive D-optimal designs only for $\boldsymbol{\beta} = \log(\boldsymbol{\theta})$; the resulting designs are also D-optimal for the parameters $\boldsymbol{\theta}$ and $\boldsymbol{d}$, which, as shown above, does not hold for A-optimal designs. However, the derived D-optimal designs require knowledge of the unknown parameter vector $\boldsymbol{\beta}$ and, in addition, of τ and ρ. The A-optimal designs derived above, in contrast, only require the specification of the parameter of interest, $\boldsymbol{\beta}$ or $\boldsymbol{\theta}$, while the A-optimal designs for $\boldsymbol{d}$ do not presuppose any a priori knowledge at all. Moreover, the A-criterion is easier to interpret, as it is directly aligned with the substantive research objective, namely, the precise estimation of the parameters themselves, without taking the covariances of the estimators into account. Altogether, both types of designs have their merits; for practical purposes, however, A-optimal designs may be considered more advantageous for the longitudinal RPGCM and are only feasible because of their parsimonious requirements regarding prior knowledge.

In addition to the unstructured model considered in the present study, Parsamaram et al. (2026) derived general conditions for D-optimal designs for models with structured functional relationships describing the evolution of the mean ability $\theta(t)$ over time, such as generalized linear and nonlinear models. Extending the present A-optimality results to such models would be substantially more challenging because the invariance property of the D-criterion no longer applies and the A-optimal designs additionally depend on the variance parameters. The development of A-optimal designs for these settings therefore may remain a topic for future research. It should be noted that an unstructured model may often be adequate, in particular, when only a few measurement occasions are analyzed. As the meta-analysis of Scharfen et al. (2018) shows, increases in the abilities are to be expected in most studies only to the third or fourth measurement occasion. Furthermore, of the studies included in this meta-analysis, most studies (75%) analyzed retest effects for two occasions, while 19% and 4% of the studies included three and four repetitions, respectively. Only 2% of the studies included in this meta-analysis had more than four measurement points. Nevertheless, the analysis of ability changes may be interesting for a larger set of repetitions, in particular, in clinical settings.

A direct comparison between the D-optimal designs derived by Parsamaram et al. (2026) and the A-optimal designs developed in the present study is difficult. As noted above, nearly all D-optimal designs for the RPGCM require knowledge not only of the parameters of interest, such as $\boldsymbol{\beta}$, but also of additional quantities, including τ and ρ. In contrast, the derivation of the A-optimal designs considered here does not depend on such prior information.

One exception is the unstructured model with equal means considered in Example 5.1 by Parsamaram et al. (2026). In this case, the D-optimal design yields uniform weights, $w_j = 1/J$, for the parameters β_j. Owing to the invariance property of the criterion, the same weights are obtained for θ_j and d_j. These weights coincide with those obtained from the corresponding A-optimal designs for these parameters (see Table 2).

6. Summary

In this article, we developed A-optimal designs for longitudinal RPGCMs. The proposed designs also apply to more general longitudinal Poisson Gamma models and are therefore relevant across disciplines in which repeated count data are analyzed.

The designs specify optimal item allocation proportions across measurement occasions to achieve the most precise estimation of parameters such as those of mean abilities and commonly used contrasts. By minimizing estimator variances, these designs enhance statistical efficiency while potentially reducing participant burden.

An important direction for future research is the development of A-optimal designs for comparing multiple growth curves, as many applications focus on contrasts between trajectories, such as treatment and control groups. In addition, models imposing structured functional relationships on the evolution

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

of the mean ability over time should be considered, particularly when data are collected at a large number of time points.

Acknowledgements. We are grateful to the editor, the associate editor, and two anonymous referees for their helpful comments and suggestions.

Data availability statement. No data were used or generated in this study.

Funding statement. This work was supported by the German Research Foundation under Grants HO1286/6 and SCHW531/15.

Competing interests. The authors have no competing interests to declare relevant to this article's content.

References

Baghaei, P., & Doebler, P. (2019). Introduction to the Rasch Poisson counts model: An R tutorial. *Psychological Reports*, *122*, 1967–1994.

Beisemann, M. (2022). A flexible approach to modelling over-, under- and equidispersed count data in IRT: The two-parameter Conway–Maxwell–Poisson model. *British Journal of Mathematical and Statistical Psychology*, *75*, 411–443.

Beisemann, M., Holling, H., & Doebler, P. (2025). Every trait counts: Marginal maximum likelihood estimation for novel multidimensional count data item response models with rotation or ℓ_1–regularization for simple structure. *Psychometrika*, *90*, 304–330.

Freise, F., Holling, H., & Schwabe, R. (2023). Optimal design for estimating the mean ability over time in repeated item response testing. *Journal of Statistical Planning and Inference*, *225*, 266–282.

Graßhoff, U., Holling, H., & Schwabe, R. (2020). D-optimal design for the Rasch counts model with multiple binary predictors. *British Journal of Mathematical and Statistical Psychology*, *73*, 541–555.

Hung, L. (2012). A negative binomial regression model for accuracy tests. *Applied Psychological Measurement*, *36*, 88–103.

Jäger, A. O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen: Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells [Multimodal classification of intelligence achievement: Experimentally controlled, further development of a descriptive intelligence structure model]. *Diagnostica*, *28*, 195–225.

Jansen, M. G. (1986). A Bayesian version of Rasch's multiplicative Poisson model for the number of errors of an achievement test. *Journal of Educational Statistics*, *11*, 147–160.

Jansen, M. G. (1997). Applications of Rasch's Poisson counts model to longitudinal count data. In R. Langeheine, & J. Rost (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 380–388). Waxmann.

Jansen, M. G., & van Duijn, M. (1992). Extensions of Rasch's multiplicative Poisson model. *Psychometrika*, *57*, 405–414.

Myszkowski, N., & Storme, M. (2021). Accounting for variable task discrimination in divergent thinking fluency measurement: An example of the benefits of a 2-parameter Poisson counts model and its bifactor extension over the Rasch Poisson counts model. *The Journal of Creative Behavior*, *55*, 800–818.

Owen, R. J. (1969). A Bayesian analysis of Rasch's multiplicative Poisson model for misreadings. *ETS Research Bulletin Series*, *1969*(2), i–22.

Parsamaram, P., Holling, H., & Schwabe, R. (2026). Optimal design in repeated testing for count data. *Journal of Statistical Planning and Inference*, *241*, 106334.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research.

Scharfen, J., Peters, J. M., & Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence*, *67*, 44–66.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. CRC.

van Duijn, M., & Böckenholt, U. (1995). Mixture models for the analysis of repeated count data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *44*, 473–485.

Verhelst, N., & Kamphuis, F. (2009). *A Poisson-gamma model for speed tests*. Cito.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, *61*, 439–447.

Wedel, M., Böckenholt, U., & Kamakura, W. A. (2003). Factor models for multivariate count data. *Journal of Multivariate Analysis*, *87*, 356–369.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

Appendix

Extended longitudinal RPGCM with varying time-specific variance

In this appendix, we extend the longitudinal RPGCM by allowing the variances of the random effects to vary across measurement occasions. We first describe the changes in the latent variable Λ_{ij}. On this basis, we derive the corresponding quasi-information matrix. A-optimal weights for the model parameters and contrasts are then obtained from this matrix using the same procedure as for the weights reported in Table 2. The resulting weights for the extended model are given in Table A1.

The time-specific effects Γ_{ij} as part of the latent variable Λ_{ij}, which is decomposed as

$$\Lambda_{ij} = \Gamma_{i0} + \Gamma_{ij}, \qquad j = 1, \ldots, J,$$

are now assumed to be distributed as $\Gamma(\alpha_j, \tau)$, $j = 1, \ldots, J$. Again, the permanent person effect $\Gamma_{i0} \sim \Gamma(\alpha_0, \tau)$ shares the same scale parameter τ, and all components are mutually independent.

Then, the latent variable Λ_{ij} is distributed as $\Gamma(\alpha_0 + \alpha_j, \tau)$ with mean $\mathrm{E}(\Lambda_{ij}) = (\alpha_0 + \alpha_j)\tau$, and the marginal mean of the responses is given by

$$\mathrm{E}(Y_{ijk}) = (\alpha_0 + \alpha_j)\tau\theta(t_j)\sigma_{jk}.$$

The vector $\mathbf{\Lambda}_i = (\Lambda_{i1}, \ldots, \Lambda_{iJ})^\top$ is distributed according to a multivariate gamma distribution with $\mathrm{Var}(\Lambda_{ij}) = (\alpha_0 + \alpha_j)\tau^2$ and $\mathrm{Cov}(\Lambda_{ij}, \Lambda_{ij'}) = \alpha_0\tau^2$. Thus,

$$\mathrm{Cov}(\mathbf{\Lambda}_i) = \tau^2\left[\mathrm{diag}(\alpha_j) + \alpha_0 \mathbf{1}_J \mathbf{1}_J^\top\right].$$

For deriving the representation of the quasi-information matrix of this extended approach $\mathbf{M}_\mathrm{Q}$, we denote by $\mathbf{Y}_{ij} = (Y_{ij1}, \ldots, Y_{ijn_j})^\top$ the vector of responses of test person i at time t_j and by $\mathbf{Y}_i = (\mathbf{Y}_{i1}^\top, \ldots, \mathbf{Y}_{iJ}^\top)^\top$ the vector of all responses of test person i over all J testing occasions with first- and second-order moments

$$\begin{aligned}
\mu_{jk} &= \mathrm{E}(Y_{ijk}) = (\alpha_0 + \alpha_j)\tau\theta(t_j)\sigma_{jk},\\
\boldsymbol{\mu}_j &= \mathrm{E}(\mathbf{Y}_{ij}) = (\alpha_0 + \alpha_j)\tau\theta(t_j)(\sigma_{j1}, \ldots, \sigma_{jn_j})^\top,\\
\boldsymbol{\mu} &= \mathrm{E}(\mathbf{Y}_i) = (\boldsymbol{\mu}_1^\top, \ldots, \boldsymbol{\mu}_J^\top)^\top, \quad \text{and}\\
\mathbf{V} &= \mathrm{Cov}(\mathbf{Y}_i) = \mathrm{diag}(\mu_{jk}) + \mathrm{diag}\left(\frac{\alpha_j}{(\alpha_0 + \alpha_j)^2}\boldsymbol{\mu}_j\boldsymbol{\mu}_j^\top\right) + \frac{\alpha_0}{(\alpha_0 + \alpha_j)^2}\boldsymbol{\mu}\boldsymbol{\mu}^\top.
\end{aligned}$$

For the characterization of the inverse covariance matrix $\mathbf{V}^{-1}$, we introduce

$$\begin{aligned}
\mathbf{b}_j &= \frac{\alpha_0 + \alpha_j}{(\alpha_0 + \alpha_j)^2 + \alpha_j\sum_{k=1}^{n_j}\mu_{jk}}\mathbf{1}_{n_j},\\
\mathbf{b} &= (\mathbf{b}_1^\top, \ldots, \mathbf{b}_J^\top)^\top, \quad \text{and}\\
q_j &= \frac{(\alpha_0 + \alpha_j)\sum_{k=1}^{n_j}\mu_{jk}}{(\alpha_0 + \alpha_j)^2 + \alpha_j\sum_{k=1}^{n_j}\mu_{jk}}.
\end{aligned}$$

Then

$$\mathbf{V}^{-1} = \mathrm{diag}\left(\frac{1}{\mu_{jk}}\right) - \mathrm{diag}\left(\frac{\alpha_j}{\alpha_0 + \alpha_j + \alpha_j\sum_{k=1}^{n_j}\mu_{jk}}1_{n_j}1_{n_j}^\top\right) - \frac{\alpha_0}{1 + \alpha_0}\sum_{j=1}^{J}\frac{q_j}{\alpha_0 + \alpha_j}\mathbf{b}\mathbf{b}^\top.$$

With $\mathbf{D} = \mathrm{diag}(\boldsymbol{\mu}_j)$ and $\mathbf{q} = (q_1, \ldots, q_J)^\top$, the quasi-information matrix becomes

$$\begin{aligned}
\mathbf{M}_\mathrm{Q} &= N\mathbf{D}^\top\mathbf{V}^{-1}\mathbf{D}\\
&= N\left[\mathrm{diag}((\alpha_0 + \alpha_j)q_j) - \frac{\alpha_0}{1 + \alpha_0\sum_{j=1}^{J}q_j/(\alpha_0 + \alpha_j)}\mathbf{q}\mathbf{q}^\top\right],
\end{aligned}$$

and the representation of

$$\mathbf{M}_\mathrm{Q}^{-1} = \frac{1}{N}\left[\mathrm{diag}\left(\frac{1}{n(\alpha_0 + \alpha_j)\tau\theta_j w_j\tilde{\sigma}}\right) + \mathrm{diag}\left(\frac{\alpha_j}{(\alpha_0 + \alpha_j)^2}\right) + \alpha_0\mathbf{a}\mathbf{a}^\top\right],$$

where $\mathbf{a} = (1/(\alpha_0 + \alpha_1), \ldots, 1/(\alpha_0 + \alpha_J))^\top$, follows by standard arguments of matrix inversion.

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press

Table A1. A-optimal weights for parameters and contrasts, general shape parameters $a_j = \alpha_0 + \alpha_j$

	Parameter $\boldsymbol{\delta}$	w_1^*	$w_j^*, j = 2, \ldots, J-1$	w_J^*	c_δ
1	β_j		$\sqrt{\frac{1}{\theta_j a_j}}/c_{\beta_j}$		$c_{\beta_j} = \sum_{k=1}^{J}\sqrt{\frac{1}{\theta_k a_k}}$
2	$\beta_j - \beta_1$	$\sqrt{\frac{J-1}{\theta_1 a_1}}/c_{\beta_j-\beta_1}$	$\sqrt{\frac{1}{\theta_j a_j}}/c_{\beta_j-\beta_1}$		$c_{\beta_j-\beta_1} = \sqrt{\frac{J-1}{\theta_1 a_1}} + \sum_{k=2}^{J}\sqrt{\frac{1}{\theta_k a_k}}$
3	$\beta_j - \beta_{j-1}$	$\sqrt{\frac{1}{\theta_1 a_1}}/c_{\beta_j-\beta_{j-1}}$	$\sqrt{\frac{2}{\theta_j a_j}}/c_{\beta_j-\beta_{j-1}}$	$\sqrt{\frac{1}{\theta_J a_J}}/c_{\beta_j-\beta_{j-1}}$	$c_{\beta_j-\beta_{j-1}} = \sqrt{\frac{1}{\theta_1 a_1}} + \sum_{k=2}^{J-1}\sqrt{\frac{2}{\theta_k a_k}} + \sqrt{\frac{1}{\theta_J a_J}}$
4	θ_j		$\sqrt{\frac{\theta_j}{a_j}}/c_{\theta_j}$		$c_{\theta_j} = \sum_{k=1}^{J}\sqrt{\frac{\theta_k}{a_k}}$
5	$\theta_j - \theta_1$	$\sqrt{\frac{(J-1)\theta_1}{a_1}}/c_{\theta_j-\theta_1}$	$\sqrt{\frac{\theta_j}{a_j}}/c_{\theta_j-\theta_1}$		$c_{\theta_j-\theta_1} = \sqrt{\frac{(J-1)\theta_1}{a_1}} + \sum_{k=2}^{J}\sqrt{\frac{\theta_k}{a_k}}$
6	$\theta_j - \theta_{j-1}$	$\sqrt{\frac{\theta_1}{a_1}}/c_{\theta_j-\theta_{j-1}}$	$\sqrt{\frac{2\theta_j}{a_j}}/c_{\theta_j-\theta_{j-1}}$	$\sqrt{\frac{\theta_J}{a_J}}/c_{\theta_j-\theta_{j-1}}$	$c_{\theta_j-\theta_{j-1}} = \sqrt{\frac{\theta_1}{a_1}} + \sum_{k=2}^{J-1}\sqrt{\frac{2\theta_k}{a_k}} + \sqrt{\frac{\theta_J}{a_J}}$
7	θ_j/θ_1	$\sqrt{\frac{\sum_{j=2}^{J}\theta_j^2}{\theta_1 a_1}}/c_{\theta_j/\theta_1}$	$\sqrt{\frac{\theta_j}{a_j}}/c_{\theta_j/\theta_1}$		$c_{\theta_j/\theta_1} = \sqrt{\frac{\sum_{k=2}^{J}\theta_k^2}{\theta_1 a_1}} + \sum_{k=2}^{J}\sqrt{\frac{\theta_k}{a_k}}$
8	θ_j/θ_{j-1}	$\sqrt{\frac{\theta_2^2}{\theta_1^3 a_1}}/c_{\theta_j/\theta_{j-1}}$	$\sqrt{\frac{\theta_{j+1}^2}{\theta_j^3 a_j} + \frac{\theta_j}{\theta_{j-1}^2 a_j}}/c_{\theta_j/\theta_{j-1}}$	$\sqrt{\frac{\theta_J}{\theta_{J-1}^2 a_J}}/c_{\theta_j/\theta_{j-1}}$	$c_{\theta_j/\theta_{j-1}} = \sqrt{\frac{\theta_2^2}{\theta_1^3 a_1}} + \sum_{k=2}^{J-1}\sqrt{\frac{\theta_{k+1}^2}{\theta_k^3 a_k} + \frac{\theta_k}{\theta_{k-1}^2 a_k}} + \sqrt{\frac{\theta_J}{\theta_{J-1}^2 a_J}}$
9	d_j		$\sqrt{\frac{1}{a_j}}/c_{d_j}$		$c_{d_j} = \sum_{k=1}^{J}\sqrt{\frac{1}{a_k}}$
10	$d_j - d_1$	$\sqrt{\frac{J-1}{a_1}}/c_{d_j-d_1}$	$\sqrt{\frac{1}{a_j}}/c_{d_j-d_1}$		$c_{d_j-d_1} = \sqrt{\frac{J-1}{a_1}} + \sum_{k=2}^{J}\sqrt{\frac{1}{a_k}}$
11	$d_j - d_{j-1}$	$\sqrt{\frac{1}{a_1}}/c_{d_j-d_{j-1}}$	$\sqrt{\frac{2}{a_j}}/c_{d_j-d_{j-1}}$	$\sqrt{\frac{1}{a_J}}/c_{d_j-d_{j-1}}$	$c_{d_j-d_{j-1}} = \sqrt{\frac{1}{a_1}} + \sum_{k=2}^{J-1}\sqrt{\frac{2}{a_k}} + \sqrt{\frac{1}{a_J}}$
12	d_j/d_1	$\sqrt{\sum_{k=2}^{J}\frac{\theta_k}{\theta_1 a_1}}/c_{d_j/d_1}$	$\sqrt{\frac{1}{a_j}}/c_{d_j/d_1}$		$c_{d_j/d_1} = \sqrt{\sum_{k=2}^{J}\frac{\theta_k}{\theta_1 a_1}} + \sum_{k=2}^{J}\sqrt{\frac{1}{a_k}}$
13	d_j/d_{j-1}	$\sqrt{\frac{\theta_2}{\theta_1^2 a_1}}/c_{d_j/d_{j-1}}$	$\sqrt{\frac{\theta_{j+1}}{\theta_j^2 a_j} + \frac{1}{\theta_{j-1} a_j}}/c_{d_j/d_{j-1}}$	$\sqrt{\frac{1}{\theta_{J-1} a_J}}/c_{d_j/d_{j-1}}$	$c_{d_j/d_{j-1}} = \sqrt{\frac{\theta_2}{\theta_1^2 a_1}} + \sum_{k=2}^{J-1}\sqrt{\frac{\theta_{k+1}}{\theta_k^2 a_k} + \frac{1}{\theta_{k-1} a_k}} + \sqrt{\frac{1}{\theta_{J-1} a_J}}$

Based on this quasi-information matrix, the A-optimal weights for the model parameters and contrasts were derived for the extended model and are represented in Table A1. It should be noted that the results in Sections 2 and 3 follow by setting formally $\alpha_0 + \alpha_j = 1$.

Cite this article: Holling, H. and Schwabe, R. (2026). A-Optimal Designs for Rasch Poisson Gamma Counts Models. *Psychometrika*, 1–16. https://doi.org/10.1017/psy.2026.10136

https://doi.org/10.1017/psy.2026.10136 Published online by Cambridge University Press