Article

Applied Psychological Measurement
2018, Vol. 42(6) 478–498

DOI: 10.1177/0146621617748325
journals.sagepub.com/home/apm
SAGE

Latent Class Analysis of Recurrent Events in Problem-Solving Items

Haochen Xu[1], Guanhua Fang[2], Yunxiao Chen[3], Jingchen Liu[2], and Zhiliang Ying[2]

Abstract

Computer-based assessment of complex problem-solving abilities is becoming more and more popular. In such an assessment, the entire problem-solving process of an examinee is recorded, providing detailed information about the individual, such as behavioral patterns, speed, and learning trajectory. The problem-solving processes are recorded in a computer log file which is a time-stamped documentation of events related to task completion. As opposed to cross-sectional response data from traditional tests, process data in log files are massive and irregularly structured, calling for effective exploratory data analysis methods. Motivated by a specific complex problem-solving item "Climate Control" in the 2012 Programme for International Student Assessment, the authors propose a latent class analysis approach to analyzing the events occurred in the problem-solving processes. The exploratory latent class analysis yields meaningful latent classes. Simulation studies are conducted to evaluate the proposed approach.

Keywords

computer-based assessment, complex problem-solving, process data, event history analysis, random effect, frailty, recurrent event, PISA 2012

Introduction

Complex Problem-Solving Ability and the Climate Control Item

In recent years, computer-based problem-solving items for evaluating students' complex problem-solving (CPS) abilities are becoming popular in education, due to their better reflection of real-life challenges in comparison with traditional test items (e.g., Guzdial et al., 1996). CPS items use simulation tasks to measure students' ability, typically in technology-rich environments. These tasks mimic problems in real life, such as identifying the ID of a specified club member and sending it to a correspondent by email, purchasing a train ticket satisfying certain constraints, and so on, which, to be successfully solved, require complex skills, such as defining goals and resolving impasses. The completion of these tasks involves a sequence of interactions between the examinee and the simulation environment. CPS items have been adopted in

[1]Fudan University, Shanghai, China
[2]Columbia University, New York, NY, USA
[3]Emory University, Atlanta, GA, USA

Corresponding Author:
Yunxiao Chen, Department of Psychology, Emory University, 36 Eagle Row #270, Atlanta, GA 30322, USA.
Email: yunxiao.chen@emory.edu

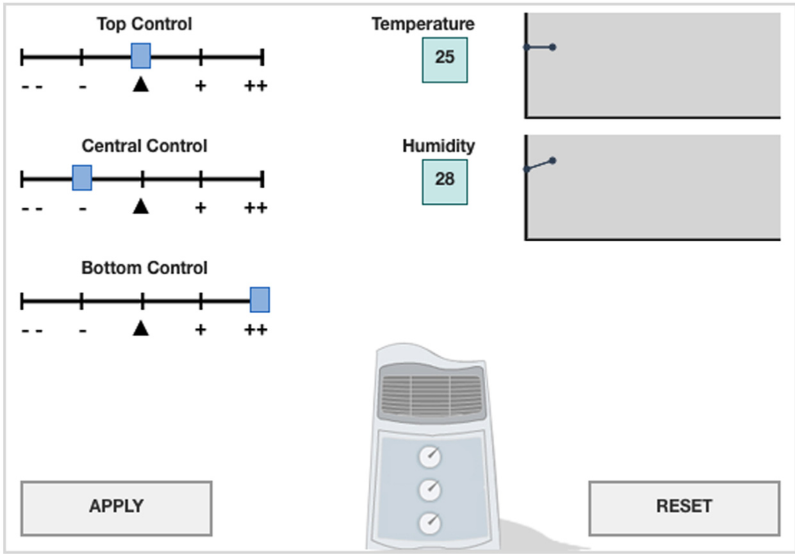

Figure 1. The Climate Control item from PISA 2012.
Note. PISA = Programme for International Student Assessment.

large-scale educational assessments, including the 2012 Programme for International Assessment of Adult Competencies (PIAAC; M. Goodman, Finnegan, Mohadjer, Krenzke, & Hogan, 2013), 2012 and 2015 Programmes for International Student Assessment (PISA; Organisation for Economic Co-Operation and Development [OECD], 2014, 2016).

In this study, the authors consider one CPS item, ''Climate Control,'' in PISA 2012, where PISA evaluates education systems worldwide by testing the skills and knowledge of 15-year-old students. The item can be found on the OECD website (http://www.oecd.org/pisa/test-2012/test-questions/question3/). As shown in the screenshot in Figure 1, in this item, examinees investigate a new air conditioner with no prior knowledge. Examinees are instructed to manipulate the panel (i.e., to move the top, central, and bottom control sliders; left side of Figure 1) and to answer how the input variables are related to the output variables, that is, temperature and humidity.

Specifically, the initial position of each control slider is indicated by a triangle ''▲.'' The examinees can change the top, central, and bottom controls on the left of Figure 1 by using the sliders. By clicking ''APPLY,'' they will see corresponding changes in temperature and humidity of the room. Figure 1 shows the outcome after clicking ''APPLY'' when the top, central, and bottom controls are set at ''▲,'' ''–,'' and ''++,'' respectively. The examinees can start over by clicking ''RESET.'' After exploration, the examinees are asked to draw lines in a diagram (Figure 2) to answer what each slider controls. Figure 2 shows the correct answer; that is, the top slider controls temperature and the middle and bottom sliders both influence humidity. The item is considered correctly answered if the diagram is correctly completed. This item assesses the examinees' ability to manipulate input variables (controls) to figure out their effects on output variables (temperature and humidity).

CPS Process Data

During an examinee's problem-solving process, a log file is automatically generated by the system in the form of a time-stamped documentation of events, recording examinees' actions on

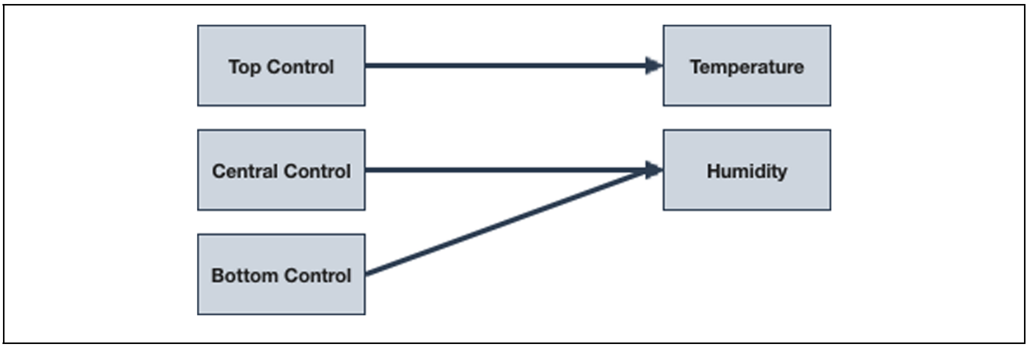

Figure 2. Climate Control item answer diagram.

computer. This temporally ordered data contain more detailed information about an examinee than the simple final outcome (i.e., correct or incorrect). Specifically, examinees' behavioral patterns, speed, and even learning trajectories may be reflected in the log file. Table 1 presents the process of one examinee solving the Climate Control item, and Table 2 explains the variables. The Climate Control log data and codebook can be downloaded online (http://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm).

CPS Process Data Analysis

The log file can be viewed as an examinee's response to the item, but it is massive and irregularly structured as opposed to the cross-sectional response data from traditional tests. As a consequence, existing psychometric models, such as item response theory models (Embretson & Reise, 2000) and cognitive diagnosis models (Rupp, Templin, & Henson, 2010), are not directly applicable. Quantitative methods for analyzing CPS process data are starting to emerge. For example, Halpin and De Boeck (2013) and Halpin, von Davier, Hao, and Liu (2017) adopt a Hawkes process approach to analyze collaborative problem-solving items that measure examinees' problem-solving skills in collaborative scenarios. He and von Davier (2015, 2016) propose an N-gram method from natural language processing for analyzing problem-solving item in technology-rich environments from PIAAC, and Vista, Care, and Awwal (2017) propose a visualisation approach for exploring examinees' complex skillsets in online collaborative problem-solving tasks.

Cluster analysis plays an important role in exploratory data analysis and is widely used in social sciences, biomedical sciences, and engineering. In the analysis of problem-solving process data, clustering examinees who share common behavioral patterns can greatly facilitate the authors' understanding of the underlying cognitive processes of problem solving. Better understanding of examinee groups further helps researcher detect feature events that influence the task performance of each group. Existing clustering methods cannot be directly applied to log file data, due to the irregular data structure. To fill the gap, the authors propose a latent class model for log file data. The proposed approach combines two types of models that are widely used in social, behavioral, and health sciences, including latent class analysis (Clogg, 1995; L. A. Goodman, 1974a, 1974b; Lanza, Flaherty, & Collins, 2003; Lazarsfeld & Henry, 1968) and event history analysis (Aalen, Borgan, & Gjessing, 2008; Allison, 1984; Cook & Lawless, 2007; Hougaard, 2000; Tekle & Vermunt, 2012; Vermunt, 2009; Zeng & Lin, 2007). It is a general method for analyzing problem-solving process data. As the first model-based

Table 1. Log Data of an Examinee's Process of Solving the Climate Control Item.

event. number	event	time	event. type	top.setting	central. setting	bottom. setting	temp. value	humid. value	diag. state
1	START_ITEM	0.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	ACER_EVENT	62.80	apply	0	0	0	25	25	NULL
3	ACER_EVENT	71.30	reset	0	0	0	25	25	NULL
4	ACER_EVENT	88.60	apply	1	0	0	27	25	NULL
5	ACER_EVENT	94.30	apply	1	−1	0	29	24	NULL
6	ACER_EVENT	99.40	apply	1	−1	1	31	25	NULL
7	ACER_EVENT	130.80	Diagram	NULL	NULL	NULL	NULL	NULL	0
8	ACER_EVENT	132.60	Diagram	NULL	NULL	NULL	NULL	NULL	100000
9	ACER_EVENT	137.30	Diagram	NULL	NULL	NULL	NULL	NULL	100000
10	ACER_EVENT	139.40	Diagram	NULL	NULL	NULL	NULL	NULL	100000
11	ACER_EVENT	141.30	Diagram	NULL	NULL	NULL	NULL	NULL	100100
12	ACER_EVENT	142.00	Diagram	NULL	NULL	NULL	NULL	NULL	100100
13	ACER_EVENT	145.10	Diagram	NULL	NULL	NULL	NULL	NULL	100101
14	END_ITEM	159.20	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Table 2. Descriptions of Variables in Log Data of Climate Control Item.

Variable name	Description
event.number	The order of the current event.
event	The status of the current event. event = START_ITEM, if the student starts to answer the item; event = END_ITEM, if the student finishes answering the item; event = ACER_EVENT, if the student clicks "APPLY," "RESET," or draw lines.
time	The amount of time (in seconds) spent on the item upon the occurrence of the current event.
event.type	The type of the current event. event.type = apply, if the student clicks "APPLY" button; event.type = reset, if the student clicks "RESET" button; event.type = Diagram, if the student draws lines; event.type = NULL, if the student starts or ends.
top.setting central.setting bottom.setting	The exact position of the slider for each control, taking values in {–2, –1, 0, 1, 2}.
temp.value	The current temperature value.
humid.value	The current humidity value.
diag.state	The status of the diagram panel. diag.state = NULL, if event.type $\neq$ Diagram; diag.state is coded by six-digit 0/1 vector, if event.type = Diagram. The status of each control is coded by two digits, with (0,0), (1,0), (0,1), (1,1) representing the control being linked to "neither," "only Temperature," "only Humidity," and "both," respectively.

clustering method for CPS process data, it complements the existing quantitative methods discussed above. The proposed method is applied to the log data of the Climate Control item. This item has been analyzed in Greiff, Wüstenberg, and Avvisati (2015) which mainly focuses on the relationship between applying an effective exploration strategy (varying one slider at a time) and the performance on the item via correlation analysis. Their analysis is confirmatory because the event ''varying one slider at a time'' is predetermined by expert knowledge. In contrast, the proposed analysis is exploratory and thus does not require any prior knowledge. The analysis of this study discovers groups of examinees with similar behavioral patterns and uncovers informative event types.

The rest of the study is organized as follows. In ''Latent Class Analysis of Recurrent Events'' section, two latent class models are proposed for analyzing event history data. The proposed approach is evaluated via simulation studies in ''Simulation Study'' section and is applied to analyzing the Climate Control item in ''Analyzing Climate Control Item'' section. Discussions and extensions are provided in ''Discussion'' section. Details of an EM algorithm are provided in Appendix A.

Latent Class Analysis of Recurrent Events

In this section, the authors present two latent class models for analyzing recurrent events in problem-solving processes. The algorithm for parameter estimation is described in Appendix A.

Model 1

Consider m examinees responding to a CPS item, each of whom is from one of K latent classes. Specifically,

$$\alpha_i \sim \mathrm{Cat}(\{1, 2, \ldots, K\}, \mathbf{p}),$$

where α_i represents the latent class membership of individual i following a categorical distribution with support $\{1, 2, \ldots, K\}$. Here, $\mathbf{p} = (p_1, \ldots, p_K)$ represents the population proportions, satisfying $\sum_{k=1}^{K} p_k = 1$.

Given the latent class membership, the authors model the history of each event type as a Poisson process which is a probabilistic model for describing the recurrent process of a certain event type (e.g., clicking a button in solving the Climate Control item). Let τ_i be the total time (in seconds) that examinee i spends on solving the item. Suppose that there are J types of events. Let n_{ij} be the total number of type j events that occur to examinee i during the problem-solving process within the time interval $[0, \tau_i]$. It is assumed that

$$n_{ij} | \alpha_i = k, \tau_i \sim \mathrm{Pois}\left(\lambda_{jk}\tau_i\right). \tag{1}$$

The authors assume that different event types are conditionally independent given the latent class membership, which is known as the local independence assumption. Finally, the total time τ_i is assumed to be independent of the latent class membership. The unknown parameters in the model include $\Lambda = (\lambda_{jk})_{J\times K}$ that describe the relationship between the events and the latent classes and the population proportions $\mathbf{p}$.

The authors provide some remarks on this model. First, assuming n_{ij} to follow a Poisson distribution with τ_i as an offset can be induced by assuming that the process of the occurrence of event type j during the time interval $[0, \tau_i]$ follows a time-homogeneous Poisson process (Allison, 1984), an important modeling tool in event history analysis. The parameter λ_{jk} is known as the intensity of the Poisson process. The examinee-specific latent variable, α_i, is known as a frailty in event history analysis, and thus, the proposed model is also known as a frailty model (Duchateau & Janssen, 2007). Frailty models are extensions of classical event history analysis models that incorporate person-specific random effects to account for unobserved heterogeneity in the data. Second, assuming local independence can be induced by assuming that the Poisson processes of different event types are independent given the latent class membership. These two assumptions are simplifications of the practical situation and can be relaxed under the event history analysis framework. That is, the occurrence of each event type can be time-inhomogeneous and can be affected by the history of its own and other processes, which can be modeled by a multivariate time-inhomogeneous Poisson process, such as a multivariate Hawkes process (Hawkes, 1971a, 1971b; Hawkes & Oakes, 1974). Further discussion will be provided in Section 5 on modeling time-inhomogeneity. Finally, the termination time τ_i may depend on the latent class membership, and the authors therefore consider an extension below.

Model 2

Model 2 extends Model 1 by allowing for the total time τ_i to depend on examinee i's latent class membership. Specifically, it is assumed that

$$\tau_i | \alpha_i = k \sim \mathrm{Gamma}(\gamma_k, \beta_k),$$

which is a Gamma distribution with shape parameter γ_k and rate parameter β_k. The Gamma distribution is adopted, because it is one of the most commonly used parametric models for the duration of time (e.g., Kalbfleisch & Prentice, 2011) and it is conjugate to the Poisson distribution. Conjugacy brings computational advantage in fitting the model. Under the Gamma distribution, the random time has mean α_k/β_k and variance α_k/β_k^2. Specifically, the total time becomes independent of the latent class membership when $\gamma_1 = \cdots = \gamma_K$ and $\beta_1 = \cdots = \beta_K$.

The rest of the model setting is the same as Model 1. In addition to Λ and $\mathbf{p}$, $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_K)$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)$ are also estimated from the data.

Model-Based Classification

Given a model, an examinee can be classified according to the posterior mode of α_i, that is

$$\hat{\alpha}_i = \underset{k \in \{1,...,K\}}{\arg\max}\, P(\alpha_i = k | n_{i1}, \ldots, n_{iJ}, \tau_i; \Lambda, \mathbf{p}) \tag{2}$$

for Model 1 and

$$\hat{\alpha}_i = \underset{k \in \{1,...,K\}}{\operatorname{argmax}}\, P(\alpha_i = k | n_{i1}, \ldots, n_{iJ}, \tau_i; \Lambda, \mathbf{p}, \boldsymbol{\gamma}, \boldsymbol{\beta}) \tag{3}$$

for Model 2. The calculation of the posterior probabilities is given in Appendix A. In practice, the authors replace the model parameters in Equations 2 and 3 by their estimates.

Simulation Study

Three simulation studies are presented in this section. First, an example is provided to illustrate the use of the proposed models. Second, the performance of the two models is evaluated under settings with many latent classes. Finally, the authors provide an example for which Model 1 is inadequate while Model 2 works well.

Study 1

A data set is stimulated from Model 1 with $m = 500$ examinees, $K = 3$ latent classes, and population proportion $\mathbf{p} = (0.2,\ 0.3, 0.5)$. The total number of event types is $J = 15$, and the intensity parameters Λ are shown in Table 3. For ease of exposition, the authors let intensity λ_{jk} only take value 0.1 or 0.01. Specifically, $\lambda_{jk} = 0.1$ means that for an examinee from class k, event j occurs on average 0.1 times per second, that is, 6 times per minute. Thus, for examinees from class k, event type j happens more frequently than event type j' for which $\lambda_{j'k} = 0.01$ (i.e., on average 0.6 times per minute). The total time τ_i is assumed to follow Gamma$(4, 0.04)$ distribution, not depending on the latent class membership. It means that on average each examinee takes $4/0.04 = 100$ s to finish. The authors point out that Model 2 is also a true model, as τ_i's are assumed to follow a Gamma distribution.

Both models were fit to this simulated data set. Results of parameter estimation are shown in Tables 4 and 5. Model 1 does not model the distribution of termination time τ_i, and thus, only the population proportion $\mathbf{p}$ and intensity matrix Λ are estimated. For Model 2, in addition to Λ and $\mathbf{p}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\beta}$ are estimated that characterize the distribution of termination time given the latent classes. For both fitted Models 1 and 2, 0.6% of examinees are misclassified based on the posterior mode estimate of the latent class membership, where the posterior mode estimator is described in Section 2.3.

Study 2

In this study, the performance of Models 1 and 2 is evaluated when there exist many latent classes ($K = 10,\ 20$). The authors consider sample sizes $m = 500,\ 1000$ and types of events $J = 50,\ 100$, yielding eight conditions in total. For each condition, the population proportions

Table 3. Simulation Study 1: True Λ.

	k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Λ^T	1	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	2	0.10	0.10	0.10	0.10	0.10	0.01	0.01	0.01	0.01	0.01	0.10	0.10	0.10	0.10	0.10
	3	0.01	0.01	0.01	0.01	0.01	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10

Table 4. Simulation Study 1: $\hat{\Lambda}$ and $\hat{\mathbf{p}}$ From Fitting Model 1.

	K j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\hat{\mathbf{p}}$
$\hat{\Lambda}^T$	1	1.0×10^{-1}	1.0×10^{-1}	9.9×10^{-2}	9.9×10^{-2}	1.0×10^{-1}	9.7×10^{-2}	1.1×10^{-1}	9.4×10^{-2}	9.5×10^{-2}	9.7×10^{-2}	8.7×10^{-3}	1.0×10^{-2}	8.1×10^{-3}	1.2×10^{-2}	1.1×10^{-2}	2.0×10^{-1}
	2	9.8×10^{-2}	9.7×10^{-2}	9.8×10^{-2}	9.8×10^{-2}	1.0×10^{-1}	9.1×10^{-3}	1.0×10^{-2}	9.6×10^{-3}	9.1×10^{-3}	1.1×10^{-2}	9.7×10^{-2}	1.0×10^{-1}	1.0×10^{-1}	9.8×10^{-2}	1.0×10^{-1}	2.8×10^{-1}
	3	9.6×10^{-3}	1.1×10^{-2}	1.0×10^{-2}	1.0×10^{-2}	9.9×10^{-3}	9.7×10^{-2}	9.6×10^{-2}	9.7×10^{-2}	9.9×10^{-2}	9.7×10^{-2}	1.0×10^{-1}	9.6×10^{-2}	1.0×10^{-1}	9.7×10^{-2}	1.0×10^{-1}	5.2×10^{-1}

Table 5. Simulation Study 1: $\hat{\Lambda}$, $\hat{\mathbf{p}}$, $\hat{\gamma}$, $\hat{\beta}$ From Fitting Model 2.

	k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\hat{\mathbf{p}}$	$\hat{\gamma}$	$\hat{\beta}$
$\hat{\Lambda}^T$	1	9.9×10^{-2}	1.1×10^{-1}	1.0×10^{-1}	1.0×10^{-1}	1.0×10^{-1}	1.0×10^{-1}	1.0×10^{-1}	9.9×10^{-2}	1.0×10^{-1}	1.0×10^{-1}	1.1×10^{-2}	1.0×10^{-2}	1.1×10^{-2}	1.1×10^{-2}	1.2×10^{-2}	1.9×10^{-1}	3.4	3.2×10^{-2}
	2	1.0×10^{-1}	9.6×10^{-2}	9.8×10^{-2}	9.5×10^{-2}	9.6×10^{-2}	9.1×10^{-3}	1.0×10^{-2}	1.0×10^{-2}	9.4×10^{-3}	1.0×10^{-2}	9.9×10^{-2}	9.8×10^{-2}	9.6×10^{-2}	1.0×10^{-1}	9.8×10^{-2}	2.9×10^{-1}	3.8	3.7×10^{-2}
	3	9.8×10^{-3}	1.1×10^{-2}	9.0×10^{-3}	9.1×10^{-3}	9.7×10^{-3}	9.9×10^{-2}	1.0×10^{-1}	1.0×10^{-1}	9.7×10^{-2}	9.9×10^{-2}	1.0×10^{-1}	1.0×10^{-1}	9.8×10^{-2}	1.0×10^{-1}	1.0×10^{-1}	5.2×10^{-1}	4.7	4.6×10^{-2}

Table 6. Simulation Study 2: Results on the Accuracy Index γ (Unit: %).

	$m=500$				$m=1000$			
	$J=50$		$J=100$		$J=50$		$J=100$	
	$K=10$	$K=20$	$K=10$	$K=20$	$K=10$	$K=20$	$K=10$	$K=20$
Model 1	3.74	3.95	3.70	3.86	2.01	3.13	1.98	3.06
Model 2	3.71	3.89	3.71	3.85	2.00	3.09	1.96	3.06
Bayes	3.20	3.84	3.19	3.84	1.70	2.89	1.69	2.86

are assumed to be uniform, that is, $p_k = 1/K$. In addition, termination time τ_i is assumed to follow Gamma$(4, 0.04)$, independent of the latent class membership. The settings of Λ are provided in Appendix B. For each simulation condition, 100 independent replications are generated.

The performance of the two models were evaluated based on their classification accuracy. Let $(\hat{\alpha}_1, \ldots, \hat{\alpha}_m)$ be the posterior mode estimates of the latent class membership and $(\alpha_1, \ldots, \alpha_m)$ be the true membership. $\hat{\alpha}_i$ and α_i are not directly comparable, because the labels of latent classes may have been switched when fitting the model. Ideally, the authors would like to find a permutation of the latent classes $1, \ldots, K$, denoted by $\kappa(1), \ldots, \kappa(K)$, such that $(\kappa(\hat{\alpha}_1), \ldots, \kappa(\hat{\alpha}_m))$ is most close to $(\alpha_1, \ldots, \alpha_m)$ and then compare under a standard 0/1 loss function. This is how the classification errors are computed in Studies 1 and 3. However, finding the optimal permutation requires searching over all possible permutations of $\{1, \ldots, K\}$, which is computationally infeasible for large values of K. To avoid this issue, the authors measure the classification accuracy by the following index

$$\gamma = \frac{2}{m(m-1)} \sum_{i \neq j} |1_{\{\alpha_i \neq \alpha_j\}} - 1_{\{\hat{\alpha}_i \neq \hat{\alpha}_j\}}|,$$

which takes value between 0 and 1 and is invariant under label switching. This index calculates the proportion of pairs of examinees who are inconsistently classified based on the estimated and the true latent class memberships. The smaller the γ, the more accurate the classification.

Results are shown in Table 6. The average of γ over 100 replications is reported for both fitted Models 1 and 2. They are slightly larger than the oracle mean γ values under the true model (“Bayes” in Table 6). Finally, the authors point out that if one randomly assigns examinees into K latent classes, the mean of γ is 50%. The results indicate that the proposed two models perform reasonably well under these settings when the number of latent classes is large.

Study 3

Data are finally generated from Model 2 where the distribution of termination time depends on the latent class membership. Specifically, τ_i follow Gamma$(1, 1)$, Gamma$(1, 0.01)$, and Gamma$(1, 0.005)$, given $\alpha_i = 1, 2$, and 3, respectively. The rest of the setting is the same as in Study 1. Under this setting, the average termination time for classes 1, 2, and 3 is 1, 100, and 200 s, respectively. For example, Class 1 may be those who skip the item, Class 2 may be a group of fast examinees, and Class 3 may be a group of slow ones. To compare the two models, 100 independent data sets are generated.

The estimates of **p** from the two models are shown in Tables 7, where the size of the first latent class is significantly underestimated by Model 1. In addition, the classification error

Table 7. Simulation Study 3: Means of $\hat{p}_1$, $\hat{p}_2$, and $\hat{p}_3$ From Models 1 and 2 Over 100 Replications and Their Standard Errors.

	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$
Model 1	$6.6\times10^{-3}(3.8\times10^{-2})$	$3.7\times10^{-1}(6.1\times10^{-3})$	$6.2\times10^{-1}(1.5\times10^{-2})$
Model 2	$1.7\times10^{-1}(6.8\times10^{-3})$	$3.2\times10^{-1}(2.3\times10^{-3})$	$5.1\times10^{-1}(1.9\times10^{-3})$

(averaged over 100 relications) for Model 1 is 21.6%, which is much higher than that of Model 2 (6.3%). Model 1 does not perform well in this example due to its failure to capture the difference in the termination time between the latent classes.

Analyzing Climate Control Item

The log data of the Climate Control item are analyzed using the two proposed models. This data set contains 16,920 15-year-old students' problem-solving processes, 54.4% of whom answer correctly. This data set is cleaned from the entire data released by OECD, and the students in the data are from 42 countries and economies. The histogram of total time spent on this item is shown in Figure 3. Event types considered in the current analysis are all possible configurations of the three control sliders when the ''APPLY'' button is clicked (i.e., all events with ''event.type = apply'' in Table 2). Each event type is coded by a vector of three entries, with each entry taking a value in $\{-2,-1,0,1,2\}$. The value of each entry represents the position of the corresponding control slider. For example, (1, 0, 0) denotes the event that the top control slider is set at ''+,'' and the rest two are set at ''▲'' (see Figure 1).

Different choices of the number of latent classes, $K = 2, 3, ..., 15$, are investigated. It turns out that the two models are robust in the sense that for each model, when $K \geq 5$, the latent classes of a substantial size ($\hat{p}_k > 0.1$) are similar across different K's. In Tables 8 and 9, results of fitting Models 2 and 1 when $K = 15$ are presented, including the estimates of the parameters associated with the latent classes of a substantial size ($\hat{p}_k > 0.1$). For the intensity matrix Λ, only events that most frequently occur in the top latent classes are presented.

The results of fitting Model 2 are shown in Table 8. The first part of Table 8 is the estimated intensities of each event type. The three estimated intensities with highest values in each class are highlighted in bold font. The other model parameters, as well as the correct rates, are also presented in these tables. For example, in Table 8, the estimated intensity 1.8×10^{-3} is bolded, which corresponds to the event $(1,1,1)$. It means that according to the estimated model, on average, examinees in Class 3 apply action $(1,1,1)$ 1.8×10^{-3} times per second, that is, 0.11 times per minute. The four classes in the table constitute 74% of the population. The top three event types that most frequently occur in Class 1 are the ones with only one control at Position 2 and the other two at Position 0, which are known to be an efficient exploration strategy (Greiff et al., 2015). In fact, among those who are classified into Class 1, 81.4% of them answer the item correctly. Examinees in Class 2 also tend to vary one slider at a time, but the slider is set at Position 1. Within this group, 67.7% of the examinees answer the item correctly, which is still above the overall correct rate 54.4%. Examinees in classes 3 and 4 take inefficient actions frequently, such as applying (1,1,1), (1, –1, 1), (0, 0, 0), and (2, 2, 2), which possibly explains their low correct rates (14.1% and 35.3%). The authors also notice that the expected total time of Classes 3 and 4 are larger than those of Classes 1 and 2, indicating that it takes the examinees in Classes 3 and 4 more time to finish.

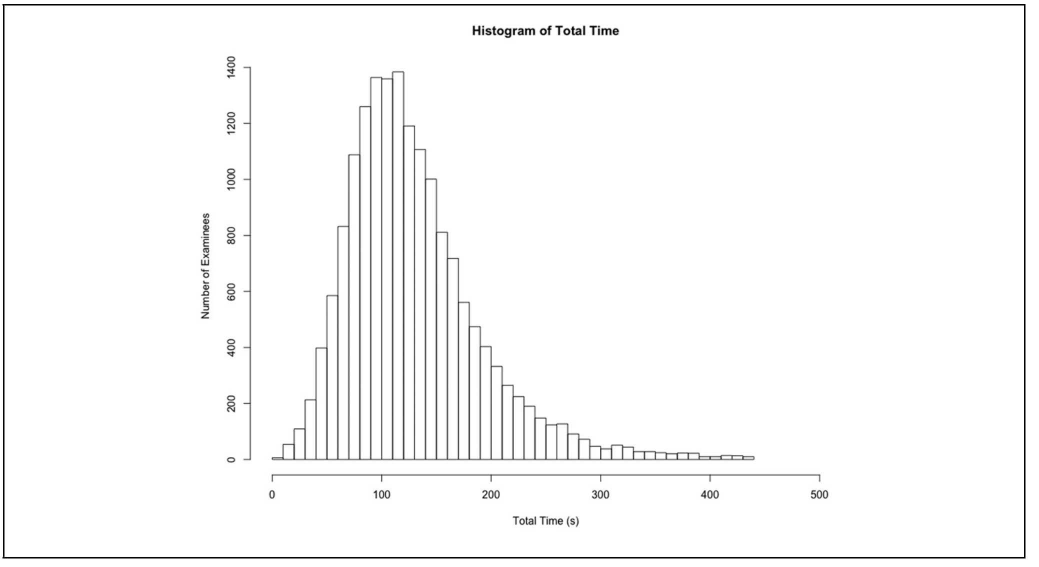

Figure 3. Climate Control item analysis: Histogram of termination time.

The results of fitting Model 1 are shown in Table 9, which can be read similarly as Table 8. The four classes in Table 9 constitute 67.2% of the population. Based on the estimated intensities, Classes 1, 2, and 3 in Table 9 share similar behavior as the first three classes of Model 2, though the estimated population proportions are different. The examinees in Class 4 of Model 1 have relatively high intensities on all the event types of varying only one control at a time, which possibly explains the high correct rate of this class (86.6%).

Discussion

In this study, the authors propose two latent class models for analyzing recurrent events in problem-solving processes. They apply the proposed model to analyze the Climate Control item in PISA 2012. By combining latent class modeling and event history analysis, the proposed models effectively cluster examinees who share similar problem-solving patterns into the same group. In the analysis of Climate Control item, meaningful latent classes of examinees are found. Examinees in these latent classes tend to take certain types of events frequently, such as actions of moving one slider at a time. It also confirms the finding in Greiff et al. (2015) that taking the strategy of moving one slider at a time is positively correlated with correctly answering the item. In the current analysis, a country effect is not considered. The relationship between the students' latent class membership and their nationality is worth investigating in future research.

The proposed method is a general tool for exploring the structure of CPS process data. Although only applied to the Climate Control item in this study, it is applicable to other CPS items as long as recurrent event types can be defined. It helps education researchers extract prevalent behavioral patterns. However, the authors point out that domain knowledge is needed to obtain good understanding and interpretation of the results.

As the authors' first step toward modeling and understanding problem-solving process data, there are problems worth future investigation. One problem is the identifiability of the proposed models. The proposed models are only identifiable up to a label switching, which is a common

Table 8. Climate Control Item Analysis: Parameter Estimates From Model 2 Associated With the Top Four Latent Classes.

top.setting	central.setting	bottom.setting	Class 1	Class 2	Class 3	Class 4
1	1	1	5.0×10^{-4}	2.5×10^{-3}	$\underline{\mathbf{1.8 \times 10^{-3}}}$	2.1×10^{-3}
2	2	2	1.3×10^{-3}	1.6×10^{-4}	6.9×10^{-4}	3.9×10^{-3}
−2	−2	−2	3.3×10^{-4}	4.7×10^{-5}	2.0×10^{-4}	2.7×10^{-3}
1	−1	1	1.4×10^{-4}	5.7×10^{-4}	$\underline{\mathbf{1.3 \times 10^{-3}}}$	1.2×10^{-3}
0	0	0	6.3×10^{-3}	5.5×10^{-3}	$\underline{\mathbf{5.5 \times 10^{-3}}}$	5.3×10^{-3}
0	2	0	$\underline{\mathbf{1.0 \times 10^{-2}}}$	8.1×10^{-4}	9.9×10^{-5}	1.7×10^{-3}
1	0	0	5.1×10^{-3}	$\underline{\mathbf{1.0 \times 10^{-2}}}$	8.8×10^{-4}	2.9×10^{-3}
2	0	0	$\underline{\mathbf{9.8 \times 10^{-3}}}$	8.5×10^{-4}	2.0×10^{-4}	2.8×10^{-3}
0	0	2	$\underline{\mathbf{9.7 \times 10^{-3}}}$	9.7×10^{-4}	1.5×10^{-4}	1.8×10^{-3}
0	1	0	4.6×10^{-3}	$\underline{\mathbf{9.6 \times 10^{-3}}}$	2.4×10^{-4}	1.9×10^{-3}
0	0	1	4.1×10^{-3}	$\underline{\mathbf{8.7 \times 10^{-3}}}$	2.4×10^{-4}	2.0×10^{-3}
$\hat{\mathbf{p}}$			2.8×10^{-1}	1.8×10^{-1}	1.7×10^{-1}	1.1×10^{-1}
$\hat{\gamma}$			6.6	5.8	3.4	5.2
$\hat{\beta}$			5.3×10^{-2}	4.5×10^{-2}	2.4×10^{-2}	3.3×10^{-2}
Correct rate			81.4%	67.7%	14.1%	35.3%

Note. The bold and underlined values indicate the three largest intensities for each of the top three classes.

Table 9. Climate Control Item Analysis: Parameter Estimates From Model 1 Associated With the Top Four Latent Classes.

top.setting	central.setting	bottom.setting	Class 1	Class 2	Class 3	Class 4
1	1	1	2.4×10^{-3}	5.6×10^{-4}	$\underline{\mathbf{1.9 \times 10^{-3}}}$	6.7×10^{-4}
2	2	2	2.1×10^{-4}	2.8×10^{-3}	8.0×10^{-4}	3.7×10^{-4}
1	−1	1	5.7×10^{-4}	1.2×10^{-4}	$\underline{\mathbf{1.2 \times 10^{-3}}}$	2.7×10^{-4}
0	0	0	5.0×10^{-3}	3.8×10^{-3}	$\underline{\mathbf{5.4 \times 10^{-3}}}$	1.1×10^{-2}
0	2	0	1.1×10^{-3}	$\underline{\mathbf{1.0 \times 10^{-2}}}$	2.8×10^{-5}	8.6×10^{-3}
1	0	0	$\underline{\mathbf{1.0 \times 10^{-2}}}$	2.9×10^{-3}	8.9×10^{-4}	8.3×10^{-3}
2	0	0	1.1×10^{-3}	$\underline{\mathbf{1.1 \times 10^{-2}}}$	1.2×10^{-4}	7.8×10^{-3}
0	0	2	1.3×10^{-3}	$\underline{\mathbf{9.8 \times 10^{-3}}}$	6.7×10^{-5}	8.5×10^{-3}
0	1	0	$\underline{\mathbf{9.3 \times 10^{-3}}}$	1.8×10^{-3}	2.3×10^{-4}	9.0×10^{-3}
0	0	1	$\underline{\mathbf{8.5 \times 10^{-3}}}$	1.7×10^{-3}	2.2×10^{-4}	7.8×10^{-3}
$\hat{\mathbf{p}}$			1.9×10^{-1}	1.9×10^{-1}	1.7×10^{-1}	1.2×10^{-1}
Correct proportion			68.4%	73.3%	13.6%	86.6%

Note. The bold and underlined values indicate the three largest intensities for each of the top three classes.

feature of latent class models (e.g., Allman, Matias, & Rhodes, 2009).[1] It is believed that the generic identifiability results for latent class models that are established in Allman et al. (2009) can be extended to the authors' models under mild regularity conditions. In addition, in the current models, given the latent class membership, the processes of different event types are modeled as independent time-homogeneous Poisson processes, which ignores the dynamic feature of the problem processes. This assumption can be relaxed by assuming the Poisson processes to be time-inhomogeneous. In that case, instead of having an intensity parameter λ_{jk} for event j and class membership k, the authors estimate an intensity function $\lambda_{jk}(t)$ depending on time t that characterizes the instantaneous probability of the event type j occurring at t. The dependence of one event at time t on the event history of its own and other events can be incorporated by properly modeling the intensity functions. The authors refer the readers to Cook and Lawless (2007) for more details. Another way to capture the dynamics of the processes is to include events of taking certain consecutive actions in the analysis. Such an event type is known as an N-gram, where N is the number of consecutive actions (He & von Davier, 2015, 2016). Finally, developing a way to choose the number of latent classes is left for future investigation.

Appendix A

Expectation–Maximization (EM) Algorithm

In this section, the authors present the details of parameter estimation using the EM algorithm. There are two steps in the EM algorithm, an expectation step (E-step) and a maximization step (M-step). Given the observed data X (such as the number of events n_{ij}, the total time τ_i), the unobserved data Y (the latent class membership α_i in this case), and a vector of unknown parameters $\boldsymbol{\theta}$ (population proportions $\mathbf{p}$, intensity matrix Λ, etc.), the authors define the complete-data likelihood as

$$L(\boldsymbol{\theta}; X, Y) = p(X, Y|\boldsymbol{\theta}),$$

and the log likelihood as $l(\boldsymbol{\theta}; X, Y) = \log L(\boldsymbol{\theta}; X, Y)$.

The EM algorithm iteratively applies the two steps until convergence. Given parameter values $\boldsymbol{\theta}^{(n)}$ obtained in the nth iteration, the $n+1$th iteration proceeds as follows:

1. (E-step) The authors calculate the expectation of the log likelihood $l(\boldsymbol{\theta}; X, Y)$ with respect to the conditional distribution of Y given X and under the current parameter $\theta^{(n)}$,

$$Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}\right) = \mathrm{E}_{Y|X,\boldsymbol{\theta}^{(n)}} l(\boldsymbol{\theta}; X, Y).$$

2. (M-step) The authors find the maximizer of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)})$ as a function of $\boldsymbol{\theta}$,

$$\boldsymbol{\theta}^{(n+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(n)}).$$

The explicit form of the terms in the EM algorithm for the two models are given in Appendix A.

Model I

Let $n = (n_{ij})_{m\times J}$, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$. The complete-data likelihood is

$$L(\Lambda, \mathbf{p}; n, \boldsymbol{\tau}, \boldsymbol{\alpha}) = \prod_{j=1}^{J} \prod_{i=1}^{m} \lambda_{j\alpha_i}^{n_{ij}} e^{-\lambda_{j\alpha_i}\tau_i} p_{\alpha_i},$$

and the log likelihod is

$$l(\Lambda, \mathbf{p}; n, \boldsymbol{\tau}, \boldsymbol{\alpha}) = \log L(\Lambda, \mathbf{p}; n, \boldsymbol{\tau}, \boldsymbol{\alpha}) = \sum_{j=1}^{J} \sum_{i=1}^{m} \sum_{k=1}^{K} I\{\alpha_i = k\} \left(n_{ij} \log\lambda_{jk} - \lambda_{jk}\tau_i + \log p_k\right).$$

Here, the observed data are $n, \boldsymbol{\tau}$, the unobserved data are $\boldsymbol{\alpha}$, and the parameters include Λ, $\mathbf{p}$. By Bayes formula, the conditional distribution of α_i given the observed data and parameters is

$$\tilde{p}_{ik} = p(\alpha_i = k | n, \boldsymbol{\tau}, \Lambda, \mathbf{p}) = \frac{(\prod_{j=1}^{J} \lambda_{jk}^{n_{ij}} e^{-\lambda_{jk}\tau_i}) p_k}{\sum_{l=1}^{K} (\prod_{j=1}^{J} \lambda_{jl}^{n_{ij}} e^{-\lambda_{jl}\tau_i}) p_l},$$

which suggests

$$\alpha_i | n, \tau, \Lambda, \mathbf{p} \sim \mathrm{Cat}(\{1, 2, \ldots, K\}, \tilde{\mathbf{p}}_i),$$

where $\mathrm{Cat}(\{1, 2, \ldots, K\}, \tilde{\mathbf{p}}_i)$ denotes the categorical distribution with support $\{1, 2, \ldots, K\}$ and $\tilde{\mathbf{p}}_i = (\tilde{p}_{i1}, \ldots, \tilde{p}_{iK})$.

Model 2

The complete-data likelihood under Model 2 is

$$L(\Lambda, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\gamma}; n, \boldsymbol{\tau}, \boldsymbol{\alpha}) = \prod_{i=1}^{m} (\prod_{j=1}^{J} \lambda_{j\alpha_i}^{n_{ij}} e^{-\lambda_{j\alpha_i}\tau_i}) \frac{\beta_{\alpha_i}^{\gamma_{\alpha_i}}}{\Gamma(\gamma_{\alpha_i})} \tau_i^{\gamma_{\alpha_i}-1} e^{-\beta_{\alpha_i}\tau_i} p_{\alpha_i},$$

and the log likelihood is

$$\begin{aligned} l(\Lambda, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\gamma}; n, \boldsymbol{\tau}, \boldsymbol{\alpha}) &= \log L(\Lambda, \mathbf{p}, \boldsymbol{\beta}, \gamma; n, \boldsymbol{\tau}, \boldsymbol{\alpha}) \\ &= \sum_{j=1}^{J} \sum_{i=1}^{m} \sum_{k=1}^{K} I\{\alpha_i = k\} \big(n_{ij} \log \lambda_{jk} - \lambda_{jk}\tau_i + \gamma_k \log \beta_k \\ &\quad -\log \Gamma(\gamma_k) + (\gamma_k - 1) \log \tau_i - \beta_k \tau_i + \log p_k\big). \end{aligned}$$

The observed and unobserved data are the same as in Model 1, but the parameters further include $\boldsymbol{\beta}, \boldsymbol{\gamma}$ in the distributions of total time. Similarly, the conditional distribution of α_i given the observed data and parameters is

$$\tilde{p}_{ik} = p(\alpha_i = k | n, \tau, \Lambda, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \frac{\left(\prod_{j=1}^{J} \lambda_{jk}^{n_{ij}} e^{-\lambda_{jk}\tau_i}\right) \frac{\beta_k^{\gamma_k}}{\Gamma(\gamma_k)} \tau_i^{\gamma_k - 1} e^{-\beta_k \tau_i} p_k}{\sum_{l=1}^{K} \left(\prod_{j=1}^{J} \lambda_{jl}^{n_{ij}} e^{-\lambda_{jl}\tau_i}\right) \frac{\beta_l^{\gamma_l}}{\Gamma(\gamma_l)} \tau_i^{\gamma_l - 1} e^{-\beta_l \tau_i} p_l},$$

meaning that

$$\alpha_i | n, \boldsymbol{\tau}, \Lambda, \mathbf{p}, \boldsymbol{\beta}, \boldsymbol{\gamma} \sim \mathrm{Cat}(\{1, 2, ..., K\}, \tilde{\mathbf{p}}_i),$$

where $\tilde{\mathbf{p}}_i = (\tilde{p}_{i1}, \ldots, \tilde{p}_{iK})$.

Appendix B

Intensity Matrix in Simulation Study 2

The settings of $\Lambda = (\lambda_{jk})_{J \times K}$ in simulation Study 2 are shown. Specifically, Tables B1, B2, B3, and B4 correspond to the four simulation settings, where a and b in the tables take values 0.1 and 0.01, respectively.

Table B1. $\tilde{\Lambda}$ With $J = 50, K = 10$ in Simulation Study 2.

k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b
5	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a
6	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a
7	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a
8	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
9	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
10	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b

k j	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
5	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
6	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
7	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
8	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b
9	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b
10	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a

Table B2. $\tilde{\Lambda}$ With $J = 50$, $K = 20$ in Simulation Study 2.

k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
5	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
6	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
7	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
8	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
9	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
10	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
11	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
12	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
13	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
14	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b
15	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b
16	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b
17	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b
18	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b
19	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	b	b	b
20	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a

Table B3. $\tilde{\Lambda}$ With $J = 100,\ K = 10$ in Simulation Study 2.

k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b
5	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a
6	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a
7	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
8	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
9	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
10	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b

k j	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
5	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
6	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
7	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
8	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
9	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b
10	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a

Table B4. $\tilde{\Lambda}$ With $J = 100$, $K = 10$ in Simulation Study 2.

k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
5	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b
6	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b
7	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b
8	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b
9	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
10	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a
11	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a
12	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a
13	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a
14	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
15	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
16	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
17	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
18	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
19	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
20	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b

k j	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
5	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
6	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
7	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
8	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
9	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
10	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b

(continued)

Table B4. (continued)

k j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
11	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
12	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
13	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
14	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
15	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
16	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
17	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
18	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b
19	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b
20	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by NSF grant SES-1323977, NSF grant IIS-1633360, Army Research Office grant W911NF-15-1-0159, and NIH grant R01GM047845.

Note

1. However, label switching that is often an issue in Bayesian analysis where parameters are estimated by posterior mean (e.g. Stephens, 2000) is not a concern in this analysis. This is because under the frequentist setting, the likelihood function has multiple global modes (corresponding to permutations of labels) and the EM algorithm tries to reach one of them.

References

Aalen, O., Borgan, O., & Gjessing, H. (2008). *Survival and event history analysis: A process point of view*. New York, NY: Springer.

Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data*. Beverly Hills, CA: Sage.

Allman, E. S., Matias, C., & Rhodes, J. A. (2009). Identifiability of parameters in latent structure models with many observed variables. *The Annals of Statistics*, *37, 6A*, 3099-3132.

Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311-359). New York, NY: Plenum Press.

Cook, R. J., & Lawless, J. (2007). *The statistical analysis of recurrent events*. New York, NY: Springer.

Duchateau, L., & Janssen, P. (2007). *The frailty model*. New York, NY: Springer.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.

Goodman, L. A. (1974a). The analysis of systems of qualitative variables when some of the variables are unobservable: Part I. A modified latent structure approach. *The American Journal of Sociology*, *79*, 1179-1259.

Goodman, L. A. (1974b). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, *62*, 215-231.

Goodman, M., Finnegan, R., Mohadjer, L., Krenzke, T., & Hogan, J. (2013). *Literacy, numeracy, and problem solving in technology-rich environments among U.S. adults: Results from the program for the international assessment of adult competencies 2012. First look* (NCES 2014-008). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education*, *91*, 92-105.

Guzdial, M., Kolodner, J., Hmelo, C., Narayanan, H., Carlson, D., Rappin, N., & Newstetter, W. (1996). Computer support for learning through complex problem solving. *Communications of the ACM*, *39*(4), 43-46.

Halpin, P. F., & De Boeck, P. (2013). Modelling dyadic interaction with Hawkes processes. *Psychometrika*, *78*, 793-814.

Halpin, P. F., von Davier, A. A., Hao, J., & Liu, L. (2017). Measuring student engagement during collaboration. *Journal of Educational Measurement*, *54*, 70-84.

Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society, Series B*, *33*, 438-443.

Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, *58*, 83-90.

Hawkes, A. G., & Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, *11*, 493-503.

He, Q., & von Davier, M. (2015). Identifying feature sequences from process data in problem-solving items with N-grams. In L. A. van der Ark, D. M. Bolt, W.-C. Wang, J. A. Douglas, & M. Wiberg (Eds.), *Quantitative psychology research* (pp. 173-190). New York, NY: Springer.

He, Q., & von Davier, M. (2016). Analyzing process data from problem-solving items with n-grams: Insights from a computer-based large-scale assessment. In Y. Rosen, S. Ferrara, & M. Mosharraf (Eds.), *Handbook of research on technology tools for real-world skill development* (pp. 750-777). Hershey, PA: IGI Global.

Hougaard, P. (2000). *Analysis of multivariate survival data*. New York, NY: Springer.

Kalbfleisch, J. D., & Prentice, R. L. (2011). *The statistical analysis of failure time data*. Hoboken, NJ: John Wiley.

Lanza, S. T., Flaherty, B. P., & Collins, L. M. (2003). Latent class and latent transition analysis. In I. B. Weiner, J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of psychology* (Vol. 2. pp. 663-685). Hoboken, NJ: John Wiley.

Lazarsfeld, P., & Henry, N. (1968). *Latent structure analysis*. Boston, MA: Houghton Mifflin.

Organisation for Economic Co-Operation and Development. (2014). *PISA 2012 technical report*. Retrieved from http://www.oecd.org/pisa/pisaproducts/pisa2012technicalreport.htm

Organisation for Economic Co-Operation and Development. (2016). *PISA 2015 results in focus*. Retrieved from https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York, NY: Guilford Press.

Stephens, M. (2000). Dealing with label switching in mixture models. *Journal of the Royal Statistical Society, Series B*, *62*, 795-809.

Tekle, F. B., & Vermunt, J. K. (2012). Event history analysis. In H. Cooper (Ed.), *APA handbook of research methods in psychology* (Vol. 3, pp. 267-290). Washington, DC: American Psychological Association.

Vermunt, J. K. (2009). Event history analysis. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology* (pp. 658-674). Thousand Oaks, CA: Sage.

Vista, A., Care, E., & Awwal, N. (2017). Visualising and examining sequential actions as behavioural paths that can be interpreted as markers of complex behaviours. *Computers in Human Behavior*, *76*, 656-671.

Zeng, D., & Lin, D. (2007). Maximum likelihood estimation in semiparametric regression models with censored data. *Journal of the Royal Statistical Society, Series B*, *69*, 507-564.