Psychology Science, Volume 46, 2004 (1), p. 7-34

Modeling intraindividual change over time in the absence of a "Gold Standard"

RAYMOND H. BAILLARGEON[1], RICHARD E. TREMBLAY[2], DOUG WILLMS[3], ELISA ROMANO[4], KYUNG-HYE LEE[5], HONG-XING WU[6]

Abstract

Looking at intra-individual change over time in a particular phenomenon may present some methodological challenges. The aim of this report was: 1. To show the effect of independent classification errors on the estimation of incidence and remission rates. 2. To show how a logitbased time-specific latent variables model can be used to model two distinct components of intraindividual change over time in the absence of a "gold standard", namely: (a) the continuity and discontinuity in the latent states over time; and (b) the strength of the association between the time-specific latent variables. 3. To illustrate this model using data on physical aggression from a representative sample of Canadian children assessed at 8-9 years of age and then again two years later at 10-11 years of age. The results showed that classification errors can yield either gross under or over estimates of the true incidence and remission rates. Furthermore, remission was far more sensitive than incidence to classification errors whereas incidence varied more drastically than remission depending on the amount of classification errors. We found that there was no association in the region off the main diagonal of the transition probability matrix beyond that expected by chance alone. In general, the stability of a 8-9 year-old child's latent physical aggression status (i.e., low-, medium- or high-aggressive) did not depend on its severity. Furthermore, the likelihood of changing from one latent physical aggression status to another was generally equal to the one of changing from the latter to the former.

Keywords: Classification errors, epidemiology, physical aggression, National Longitudinal Survey of Children and Youth, intraindividual change over time, latent class analysis.

[1] Raymond H. Baillargeon, Hôpital Sainte-Justine, Department of Psychiatry, 3175 Côte-Sainte-Catherine, Montréal, Québec, Canada, H3T 1C5, Phone: (514) 345-4931 (3892); Fax: (514) 345-4635; E-mail: rbaillargeon@justine.umontreal.ca.
[2] Université de Montréal, Canada
[3] University of New Brunswick, Fredericton, Canada
[4] Université d'Ottawa, Ottawa, Canada
[5] Asia University, Kyungsan-Si, South Korea
[6] Santé Canada, Ottawa, Canada

Modeling Intraindividual Change Over Time in the Absence of a "Gold Standard"

While data from cross-sectional studies are useful to study change at the aggregate level (i.e., net or mean change) data from longitudinal studies are needed to study change at the individual level. Consider the situation where one is following a birth cohort of children to understand the development of a particular mental disorder. The following two questions are of special interest: 1. To what extent will disorder-free children develop the disorder during a specified period of time (i.e., incidence)? 2. To what extent will children who suffer from the disorder become disorder-free during the same period (i.e., remission)? These questions are important independent of the fact that the prevalence of the disorder in question could decrease, increase, or remained unchanged during this period. The distinction between change at the aggregate level and change at the individual level is not always made, however. For instance, in the hypothetical situation described above, a decrease in the prevalence of the disorder with age does *not* imply that there is no incidence, merely that there are fewer children who develop the disorder for the first time than there are children who become disorder-free during this particular period of time. In fact, the incidence can be substantial and possibly represent a theoretically important aspect of the dynamic of the disorder in question. For example, a greater incidence among boys than girls might be responsible for the emergence of gender differences in the prevalence of the disorder in question (i.e., greater prevalence among boys). Similarly, an increase in the prevalence of the disorder with age does not imply that there is no remission, merely that children who become disorder-free are fewer than those who develop the disorder for the first time during this particular period of time. Hence, many different patterns of individual change can coexist within a given pattern of change at the aggregate level (Plewis, 1985). In addition, these questions are important whether or not there was a high positive or negative test-retest correlation between the repeated measures of the disorder. In fact, there can be substantial individual change over time despite perfect test-retest correlation (Alder & Scher, 1994). In sum, when studying the evolution of a particular phenomenon over time, it is important to distinguish between the strength of the association between repeated measures of the phenomenon in question, change at the aggregate level, and intraindividual change over time.

The first aim of this paper is to show the effect of independent classification errors on the estimation of incidence and remission rates. The second aim of this paper is to show how a logit-based time-specific latent variables model can be used to model intraindividual change over time in the absence of a "gold standard". The third aim of this paper is to illustrate this model using data from the National Longitudinal Survey of Children and Youth where physically aggressive behaviors were assessed in a representative sample of Canadian children at 8-9 years of age and then again two years later at 10-11 years of age.

Classification Errors and Intraindividual Change Over Time

Looking at intraindividual change over time may present some methodological challenges. Consider the situation where a less than perfectly sensitive and/or specific cutoff point is used to identify children who suffer from a particular mental disorder that is either present or absent. The cutoff point's sensitivity refers to the probability of a randomly selected disordered child in the population obtaining a score equal or above the cutoff point (i.e., true positive rate). The cutoff point's specificity refers to the probability of a randomly selected non-disordered child in the population obtaining a score below the cutoff point (i.e., true negative rate). If the cutoff point is not perfectly sensitive and/or specific the two categories of children created by this procedure, cases and non-cases, do not necessarily constitute homogeneous groups of children in the population. In fact, a number of non-disordered children will be classified as cases (i.e., [(1 - prevalence of the disorder in the population) (1 - specificity of the cutoff point)]); and, conversely, a number of disordered children will be classified as non-cases (i.e., [(prevalence of the disorder in the population) (1 - sensitivity of the cutoff point)]) yielding biased estimates of the true incidence and remission (Yanagawa & Gladen, 1984). Table 1 illustrates the effects of classification errors for selected values of the true prevalence at the first time point (i.e., .10), incidence and remission, and the cutoff point's sensitivity and specificity. We assumed that the cutoff point's characteristics did *not* vary over time and that the classification errors at the two time points were independent. For example, in the situation where there is *no* intraindividual change over time (i.e., true incidence = true remission = 0), the estimated incidence rate varies between .07 and .26. The difference between the true and the estimated remission rate is even larger with estimated rates above .5. Hence, remission seems affected much more than incidence by classification errors (especially to the lack of specificity of the cutoff point) whereas incidence seems to vary more drastically than remission depending on the cutoff's characteristics (e.g., the particular mix of disordered and non-disordered children in the case and non-case groups). Note that independent classification errors do not always yield over estimates of the true incidence and remission rates (or over estimates of the number of individuals changing states over time for that matter). Depending on the values of the true prevalence, incidence and remission, and of the cutoff point's characteristics, classification errors can either yield under or over estimates of the true incidence and remission rates. In sum, misclassifications can render hazardous the interpretation of incidence and remission estimates of a particular phenomenon obtained using a less that perfectly sensitive and/or specific cutoff point to categorize individuals[7].

A Time-Specific Latent Variables Model of Intraindividual Change Over Time

Even if there is no "gold standard" to perfectly classify individuals, it is nevertheless possible to obtain maximum likelihood estimates of the incidence and remission of a particular phenomenon of interest using latent class analysis. Let us consider the situation where several less than perfectly sensitive and/or specific symptoms, rated as either present or absent at two different time points, are used to identify children who suffer from a particular mental disorder. One latent class model for these data posits the existence of two time-specific latent variables denoted here as I and J. Each latent variable explains the interrelationships among the symptoms at a given point in time and it is assumed to have *no* effect on the ratings at the

[7] In addition, classification errors are likely to produce biased estimates of the association between the same characteristic measured at different points in time. The impact of misclassifications on the quantitative assessment of the strength of the association in a 2 x 2 table is well known, however (Cochran, 1968; Kraemer, 1979, 1985). Misclassifications do *not* increase the probability of rejecting the null hypothesis of *no* association when in fact it is true, but they do increase the probability accepting the null hypothesis when in fact it is false (see Table 1).

other point in time. In addition, each latent variable is made up of two latent classes; namely, a nondisorder (i.e., i = j = 1) and a disorder latent class (i.e., i = j = 2) with each child being in one, and only one, latent class. Within a given latent class the symptoms are assumed to be statistically independent of one another (i.e., the assumption of local independence). We can estimate the probability of a randomly selected child in the population being disordered at the first time point (i.e., π^{I}_{2}; where $\pi^{I}_{1} + \pi^{I}_{2} = 1$). At each time point we can also estimate the conditional probability of any given symptom being absent (i.e., 1) among expected non-disordered children (i.e., symptom's specificity); and, we can estimate the conditional probability of any given symptom being present (i.e., 2) among expected disordered children (i.e., symptom's sensitivity). In Figure 1, $\pi^{I}_{A(1)|1}$ and $\pi^{J}_{A(1)|1}$ are symptom *A*'s specificity at the first and second time point, respectively. Note that $(\pi^{I}_{A(1)|1} + \pi^{I}_{A(2)|1}) = (\pi^{J}_{A(1)|1} + \pi^{J}_{A(2)|1}) = 1$. Similarly, $\pi^{I}_{A(2)|2}$ and $\pi^{J}_{A(2)|2}$ are symptom *A*'s sensitivity at the first and second time point, respectively. Note that $(\pi^{I}_{A(1)|2} + \pi^{I}_{A(2)|2}) = (\pi^{J}_{A(1)|2} + \pi^{J}_{A(2)|2}) = 1$. Further, we can estimate the elements of the transition probability matrix (TPM), that is, the probabilities of making the transition from one state, disordered or nondisordered, to the other, nondisordered or disordered, or to the same state, disordered or nondisordered, respectively. The two independent probabilities of the TPM are: (a) the conditional probability of a randomly selected child in the population being disordered at the second time point given that he or she was disorder-free earlier in time (i.e., incidence), and (b) the conditional probability of a randomly selected child in the population being disorder-free at the second time point given that he or she was disordered earlier in time (i.e., remission). In Figure 1, $\pi_{2|1}$ and $\pi_{1|2}$ are the disorder's incidence and remission, respectively. Note that $(\pi_{1|1} + \pi_{2|1}) = (\pi_{1|2} + \pi_{2|2}) = 1$. Finally, we can estimate the odds, $\Theta_{ij} = [\pi_{1|1} / \pi_{2|1}] / [\pi_{1|2} / \pi_{2|2}]$, that the children's disorder states stay the same rather than change over the two time points.

Modeling Two Distinct Components of Intraindividual Change Over Time

This time-specific latent variables model has been described elsewhere and applied in different contexts (e.g., Bassi, Hagenaars, Croon & Vermunt, 2000; Collins & Wugalter, 1992; van de Pol & Langeheine, 1990). However, it is not always clear how this model, when reformulated as a logit model with latent variables (Hagenaars, 1990, 1993, 1994; Vermunt, 1997a), can be used to model two distinct components of intraindividual change over time, namely: (a) the continuity and discontinuity in the latent states over time; and (b) the strength of the association between the time-specific latent variables.

Modeling the continuity and discontinuity in the latent states over time

One hypothesis of interest stipulates that the probability of changing from one latent state to another over time is equal to the probability of changing from the latter to the former (relative symmetry; Bishop, Fienberg & Holland, 1975, pp. 259-260). In the situation described above, one can test this hypothesis by having $\pi_{2|1} = \pi_{1|2}$. Under this *conditional symmetry submodel*, the incidence of the disorder is equal to its remission. Another hypothesis of interest stipulates that there is equal stability in the different latent states over time (relative persistence; Bishop, Fienberg & Holland, 1975, pp. 259-260). In the situation described

Table 1:
The Effect of Misclassifications on Incidence, Remission, and Odds Ratio Estimates

			True incidence	0	.01	.025	.05	.075	.1
			True remission	0	.09	.225	.45	.675	.9
			True odds ratio	--	1001	134.3	23.22	5.94	1.0
Cutoff's sensitivity / specificity									
	P	.32	I	.26	.26	.27	.29	.30	.32
.95 / .75	PVP	.30	R	.54	.56	.58	.61	.65	.68
	PVN	.99	Odds ratio	2.47	2.25	1.97	1.58	1.26	1.0
	P	.18	I	.11	.12	.13	.15	.16	.18
.9 / .9	PVP	.50	R	.50	.53	.58	.66	.74	.82
	PVN	.99	Odds ratio	8.11	6.65	4.96	3.04	1.81	1.0
	P	.12	I	.07	.08	.08	.10	.11	.12
.75 / .95	PVP	.63	R	.51	.55	.60	.70	.79	.88
	PVN	.97	Odds ratio	12.66	10.14	7.29	4.16	2.23	1.0

Note. I, R refer to the incidence and remission estimate, respectively. The true prevalence rate at the first time point was set at 10% and values of the true remission were chosen to give the same prevalence at the second time point. Prevalence refers to the prevalence estimate at the first time point; PVP (i.e., predictive value positive) and PVN (i.e., predictive value negative) refer to the percentage of individuals correctly identified as having and not having the disorder, respectively.

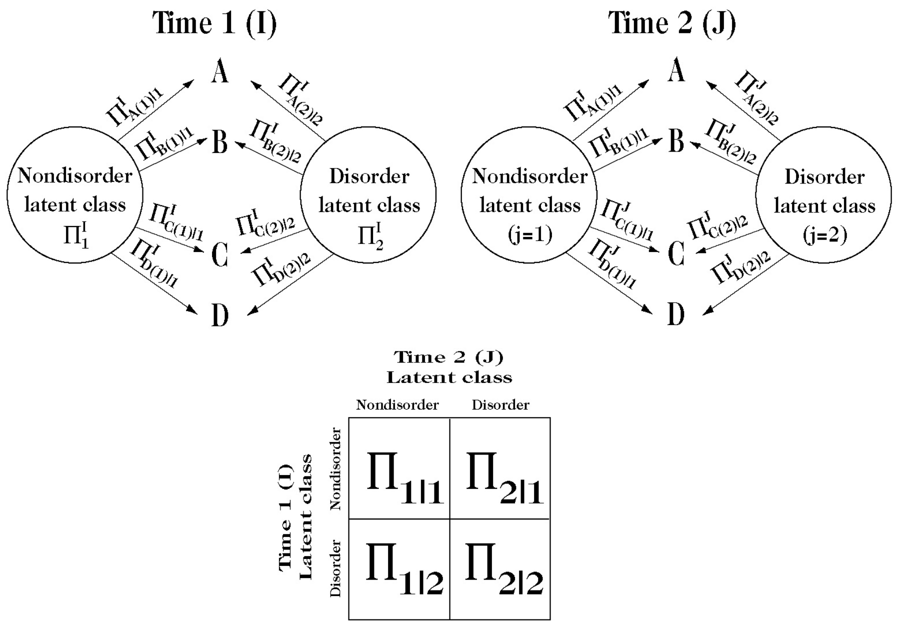

Figure 1:
Time-specific Latent Variables Model.

Note. π^I_1 and π^I_2 refer to the probability of a randomly selected child in the population being nondisordered and disordered, respectively, at the first time point. $\pi^I_{A(1)|1}$ and $\pi^J_{A(1)|1}$ refer to symptom *A*'s specificity at the first and second time point, respectively; $\pi^I_{A(2)|2}$ and $\pi^J_{A(2)|2}$ refer to symptom *A*'s sensitivity at the first and second time point, respectively. And, similarly, for the other symptoms. $\pi_{2|1}$ and $\pi_{1|2}$ refer to the disorder's incidence and remission, respectively.

above, one can test this hypothesis by having $\pi_{1|1} = \pi_{2|2}$. Under this *equal stability submodel* nondisordered and disordered children are equally likely to remain so during a specified period of time. Of course, if there are only two latent states at any given point in time, the conditional symmetry submodel and the equal stability submodel are equivalent. Note that because the marginal distribution of the latent variable at the first time point is fixed (i.e., it is considered an explanatory variable within a product-multinomial sampling scheme), it is not appropriate to test for marginal homogeneity or exact symmetry (Bishop, Fienberg & Holland, 1975; Fienberg, 1980).

Modeling the strength of the association between the time-specific latent variables

The association between two discrete ordinal variables I (i = 1, ..., I) and J (j = 1, ... , J) can be described in terms of a basic set of odds ratios in 2 x 2 subtables formed from adjacent rows (i.e., i and i + 1) and adjacent columns (i.e., j and j + 1) (Clogg & Shihadeh, 1994). This basic set includes a total of (I - 1) (J - 1) local odds ratios, $\Theta_{ij(i+1\ j+1)}$, with each odds ratio describing the association present in a particular region of the I x J table. Whenever I and J represent repeated measures of the same phenomenon it may be useful to distinguish between two different regions of the I x J table. First, the region on the main diagonal of the transition probability matrix (TPM) where change is occurring between adjacent latent states. The association in this region can be described by a subset of (I - 1) local odds ratios, $\Theta_{ij(i+1\ j+1)}$, where i = j. Second, the region off the main diagonal of the TPM where some if not all change is occurring between non-adjacent latent states. The association in this region can be described by a subset of (I - 1) (I - 2) local odds ratios, $\Theta_{ij(i+1\ j+1)}$, where $i \neq j$. One hypothesis of interest stipulates that the association (if any) present in the region on and off the main diagonal of the TPM is the same. One can test this hypothesis by constraining the (I - 1) (J - 1) local odds ratios to be equal. Under this *uniform association submodel* change occurring in the region on the main diagonal is as predictable as change occurring in the region off the main diagonal of the TPM. Another hypothesis of interest stipulates that the association (if any) present in the region off the main diagonal of the TPM is symmetric (quasi-symmetry; Bishop, Fienberg & Holland, 1975, pp. 286-288). One can test this hypothesis by having $\Theta_{ij(i+1\ j+1)} = \Theta_{i+1\ j+1(ij)}$ for all $i \neq j$. Under this *symmetric association submodel* change occurring in the region above the main diagonal is as predictable as change occurring in the region below the main diagonal of the TPM. A special case of this submodel occurs whenever $\Theta_{ij(i+1\ j+1)} = \Theta_{i+1\ j+1(ij)} = 1$ for all $i \neq j$. Under this *off diagonal, null association submodel* there is *no* association in the region off the main diagonal of the TPM beyond that expected by chance alone. Note that this submodel is *not* equivalent to the quasi-independence model (Goodman, 1965, 1968) that is often considered for the analysis of change in discrete variables (e.g., Clogg, Eliason & Grego, 1990). Finally, another hypothesis of interest is that the association (if any) present in the region on the main diagonal of the TPM is uniform. One can test this hypothesis by having $\Theta_{ij(i+1\ j+1)} = \Theta_{i+1\ j+1(i+2\ j+2)}$ for all i = j. Under this *main diagonal uniform association submodel* change occurring between any two adjacent latent states is equally predictable for the different latent states. Note that the hypothesis of null association between the two time-specific latent variables could also be considered, but it is rarely of interest.

Modeling Intraindividual Change in Physical Aggression in the Canadian Population of Children Between 8-9 and 10-11 Years of Age

One objective of this study is to illustrate the logit-based time-specific latent variables model described above using data on physical aggression from a prospective population-based survey of children between 8-9 and 10-11 years of age. More specifically, we examined the following questions: 1. How best to characterize the association between physical aggression at 8-9 and 10-11 years of age? Does the association vary depending on the region of the TPM? If so, is the association present in the region on the main diagonal of the TPM uniform? And, is the association present in the region off the main diagonal of the TPM symmetric? If so, is it different from the one expected by chance alone? Does the association vary between boys and girls? 2. How best to characterize the continuity and discontinuity in children's latent physical aggression status between 8-9 and 10-11 years of age? Does the likelihood of continuity vary depending on the child's latent physical aggression status at 8-9 years of age? Is there symmetry in the likelihood of discontinuity? In other words, is the likelihood of changing from a particular latent physical aggression status to another equal to the likelihood of changing from the latter to the former?

Method

Participants

The National Longitudinal Survey of Children and Youth (NLSCY; NLSC project team, 1995a; NLSCY project team, 1997a) is the first nation-wide household survey on child health in Canada. It was developed conjointly by Human Resources Development Canada and Statistics Canada. For the first data collection cycle (i.e., 1994-1995), a representative sample of 22,831 newborn to 11 year-old children from the 10 Canadian provinces living in private households were surveyed. Children living in hospitals and residential facilities for more than six months, on Indian reserves, and in the Yukon and Northwest Territories were not targeted by the survey. The children in each household were selected at random, up to a maximum of four children per household. A sample of 1,789 and 1,725 8-9 year-old boys and girls, respectively, was surveyed at the first data collection cycle. Two years later at the second data collection cycle (i.e., 1996-1997), 519 (497) and 520 (493) 10 (11) year-old boys and girls, respectively, from the original sample were surveyed. We eliminated cases with missing values on any of the three behavior items used to assess physical aggression (see below). This resulted in a loss of less than 3% of participants equally distributed between boys and girls and the two age groups (i.e., 8 and 9 year olds). We assumed that data were missing at random. In a probability sample such as the NLSCY, each child is assigned a longitudinal weight that stands for the number of individuals in the population that he or she "represents". These weights were used to obtain unbiased estimates of the parameters of the logit-based time-specific latent variables model. We divided each child's longitudinal weight by the mean of the weights for his or her group (e.g., 8 year-old girl) to get appropriate statistical tests. In order to take into account the NLSCY's design effect which increases the risk of falsely rejecting the null hypothesis we have used a conservative alpha level (i.e., $\alpha = .01$).

Instrument

A number of behavior items from the NLSCY (NLSC project team, 1995b; NLSCY project team, 1997b) can be used to assess children's aggression. The following three behavior items pertaining specifically to physical aggression were chosen: (a) Gets into many fights?; (b) Physically attacks people?; and (c) Kicks, bites, hits other children?. These items have been used before to estimate the prevalence of physical aggression in the Canadian population of 2 to 11 year-old children (Baillargeon, Tremblay & Willms, 1999). The primary respondent was the person most knowledgeable about the child, which in the majority of cases was the mother. Each behavior item was rated using a three-point Likert scale: never or not true (1), sometimes or somewhat true (2), and often or very true (3).

Maximum likelihood estimation

Maximum likelihood estimates of the parameters for the different logit-based time-specific latent variables models considered in this study were obtained using a computer program for the analysis of categorical data written by Jeroen K. Vermunt (1997b). This program is called lEM ("log-linear and event history analysis with missing data using the expectation maximization algorithm") and is distributed freely. Maximum likelihood estimates of the parameters of latent variables models are computed via the expectation maximization (EM) algorithm with either the iterative proportional fitting or the uni-dimensional Newton algorithm applied in the M step. First, we ran the EM algorithm more than 1,000 times using different starting values and compared the different solutions to identify a global maximum solution (i.e., the parameter estimates that maximize the likelihood function and its natural log) (Clogg, 1981). Each time we let the EM algorithm iterate to full convergence (the minimum convergence criterion was set at .00000000001) or up to 5,000 times, whichever occurred first. Then, we used the parameter estimates as starting values for the different submodels considered in this study. The lEM-code for some logit-based time-specific latent variables submodels considered in this study is presented in the appendix.

Results

Table 2 presents the goodness-of-fit test statistics for the different time-specific latent variables models considered in this study. First, we considered a time-specific latent variables model with *no* restrictions on the parameter estimates. This unconstrained model assumes that the ratings on the three physical aggression items can be explained by two time-specific latent variables each made up of three latent classes: (a) a low-aggressive latent class whose members tend to be rated in category 1 (never or not true), (b) a medium-aggressive latent class whose members tend to be rated in category 2 (sometimes or somewhat true), and (c) a high-aggressive latent class whose members tend to be rated in category 3 (often or very true). In other words, it is assumed that the behavior items tend to elicit ratings in the same category for a given child. The value of the likelihood-ratio chi-square (L^2) associated with this model is 1,037.03 with 2,824 degrees of freedom.[8] These results suggest that a time-specific latent variables model provides an acceptable fit to the physical aggression data.

[8] Note that the Pearson chi-square (X^2) suggests that this model does *not* fit the data, however. This is likely due to a relatively large number of cells with zero observed frequencies. These cells receive zero weight in the summary of fit the L^2 provides. Note that the Cressie-Read statistic, which represents a middle ground with a weight (of 2/3) that is neither 1 (as in the X^2) nor 0 (as in the L^2) leads to the same conclusion as the L^2.

Table 2:
Goodness-of-fit Test Statistics Associated with Different Time-specific Latent Variables Models

		Pearson chi-square (X^2)	p	Likelihood-ratio chi-square (L^2)	p	Cressie-Read (CR)	p	Degrees of freedom (df)	Akaike's Information Criterion (AIC)
Modeling the strength of the association between the time-specific latent variables									
Model									
1.Unconstrained	boy	2427.14	.00	642.23	1.0	1186.68	1.0	1412	-2181.77
	girl	9922.45	.00	394.80	1.0	1378.02	.74	1412	-2429.20
	total	12349.59	.00	1037.03	1.0	2564.70	1.0	2824	-4610.98
2. Homogeneous association		12016.62	.00	1064.28	1.0	2541.50	1.0	2828	-4591.72
3. Uniform association		12426.35	.00	1064.24	1.0	2597.87	1.0	2830	-4595.76
4. Symmetric association		12350.27	.00	1039.09	1.0	2578.39	1.0	2826	-4612.91
5. Main diagonal uniform association		14917.16	.00	1043.92	1.0	2724.99	.91	2826	-4608.08
Model 4 and 5		15232.81	.00	1052.45	1.0	2763.71	.80	2828	-4603.55
6. Off diagonal null association		12162.84	.00	1047.66	1.0	2575.72	1.0	2828	-4608.35

Modeling the continuity and discontinuity in the latent physical aggression status over time									
7. Equal stability	boy	6613.47	.00	682.84	1.0	1785.70	.00	1413	-2143.16
Low = medium	girl	10242.87	.00	395.01	1.0	1404.32	.56	1413	-2430.99
8. Equal stability	boy	3354.73	.00	657.67	1.0	1336.76	.93	1413	-2168.33
Medium = high	girl	9118.15	.00	395.54	1.0	1331.54	.94	1413	-2430.46
9. Equal stability	boy	2397.58	.00	644.38	1.0	1193.60	1.0	1413	-2181.62
Low = high	girl	9572.49	.00	396.21	1.0	1373.66	.77	1413	-2429.79
10. Conditional	boy	6665.69	.00	683.33	1.0	1793.86	.00	1413	-2431.20
symmetry	girl	9689.52	.00	395.62	1.0	1360.65	.84	1413	-2430.38
M|L = L|M									
11. Conditional	boy	2403.23	.00	642.51	1.0	1183.03	1.0	1413	-2183.49
symmetry	girl	--	--	--	--	--	--	--	--
H|L = L|H									
12. Conditional	boy	2430.47	.00	644.28	1.0	1200.53	1.0	1413	-2181.73
symmetry	girl	8668.78	.00	396.06	1.0	1303.48	.98	1413	-2429.94
H|M = M|H									

Note. Low = medium refers to equal stability for low- and medium-aggressive; Medium = high refers to equal stability for medium- and high-aggressive; Low = high refers to equal stability for low- and high-aggressive; M|L = L|M refers to equal likelihood of changing from low- to medium-aggressive as from medium- to low-aggressive; H|L = L|H refers to equal likelihood of changing from low- to high-aggressive as from high- to low-aggressive; H|M = M|H refers to equal likelihood of changing from medium- to high-aggressive as from high- to medium-aggressive. The lEM-code for some time-specific latent variables models is presented in the appendix.

Second, we considered a submodel which postulates that the association between the time-specific latent variables is the same for boys and girls. This homogeneous association submodel represents an increase of 27.26 in L^2 with a corresponding increase of only 4 in the degrees of freedom (L^2 = 1064.28 - 1037.03 = 27.26; df = 2828 - 2824 = 4; p = .00002) from the unconstrained model. The large increase in L^2 compared to the increase in the degrees of freedom suggests that the association between physical aggression at 8-9 and 10-11 years of age is *not* the same for boys and girls. In other words, it appears that there are gender differences in the stability of interindividual differences in physical aggression between 8-9 and 10-11 years of age. In turn, this implies that the continuity and discontinuity in children's latent physical aggression status between 8-9 and 10-11 years of age is *not* the same for boys and girls.

Third, we considered the uniform association submodel which postulates that the association between the time-specific latent variables can be described using a single odds ratio (i.e., one for boys and one for girls). This submodel represents an increase of 27.21 in L^2 with a corresponding increase of only 6 in the degrees of freedom (L^2 = 1064.24 - 1037.03 = 27.2; df = 2830 - 2824 = 6; p = .0001) from the unconstrained model. This suggests that change in latent physical aggression status occurring in the region on the main diagonal of the TPM is either less or more predictable than change occurring in the region off the main diagonal.

Fourth, we considered the symmetric association submodel and the main diagonal uniform association submodel. Neither the symmetric association submodel (L^2 = 1039.09 - 1037.03 = 2.07; df = 2826 - 2824 = 2; p = .36), nor the main diagonal uniform association submodel (L^2 = 1043.92 - 1037.03 = 6.89; df = 2826 - 2824 = 2; p = .03) represents a statistically significant decrease of fit over the unconstrained model. Note, however, that a model that combines both types of restrictions represents a statistically significant decrease of fit over the unconstrained model (L^2 = 1052.45 - 1037.03 = 15.42; df = 2828 - 2824 = 4; p = .004). Since these two submodels have the same degrees of freedom we used Akaike's Information Criterion (AIC; AIC = L^2 - [2 degrees of freedom]) to compare them. The symmetric association submodel has the smallest AIC value, and therefore should be preferred over the main diagonal uniform association submodel. These results suggest that change in latent physical aggression status occurring in the region above the main diagonal is as predictable as change occurring in the region below the main diagonal of the TPM. But, is it at all predictable? To answer this question we considered the off diagonal null association submodel which postulates that there is *no* association in the region off the main diagonal of the TPM beyond that expected by chance alone. This submodel represents an increase of 8.56 in L^2 with a corresponding increase of 2 in the degrees of freedom from the symmetric association submodel (p = .014), suggesting that change in latent physical aggression status occurring in the region off the main diagonal of the TPM is *not* predictable.[9] However, because of boundary parameter estimates converging to either 1 or 0 it was not possible to estimate the association between the time-specific latent variables under the off diagonal null association submodel; therefore, the results presented next will be drawn from the unconstrained model.

---

[9] Note, however, that the AIC associated with the symmetric association submodel is lower than the one associated with the off diagonal, null association submodel.

Prevalence of physical aggression in the Canadian population of 8-9 year-old boys and girls

Table 3 presents estimates of the conditional probability of a rating in category k (k = 1, 2, 3) to behavior item j (j = A, B, C) given membership in latent class l (l = 1, 2, 3) under the unconstrained time-specific latent variables model. These estimates allow for a clear characterization of the three latent classes. Members of the first latent class, which we shall refer to as low-aggressive, tend *not* to manifest physically aggressive behaviors. Members of the second latent class, which we shall refer to as medium-aggressive, tend to manifest physically aggressive behaviors on an occasional basis. Members of the third latent class, which we shall refer to as high-aggressive, tend to manifest physically aggressive behaviors on a frequent basis or at least, to show a propensity to do so that is much higher than medium-aggressive children. Note that these estimates are not the same at 8-9 and 10-11 years of age, but the characterization of the three latent classes is essentially the same. The majority of children in the population were estimated to belong to the low-aggressive latent class at 8-9 years of age (i.e., 69.5% of boys and 79.7% of girls). In contrast, a small but nonetheless substantial percentage of children were estimated to belong to the high-aggressive latent class at 8-9 years of age (i.e., 5.9% of boys and 2.5% of girls).[10] The prevalence of physical aggression was higher among 8-9 year-old boys. For instance, the odds of belonging to the high- rather than low-aggressive latent class were 2.66 [99% confidence interval: 1.19-5.99] times higher among 8-9 year-old boys than girls.

Continuity and discontinuity in children's latent physical aggression status from 8-9 to 10-11 years of age

Table 4 presents estimates of the conditional probability of a randomly selected 10-11 year-old child in the Canadian population belonging to the low-, medium- or high-aggressive latent class, given his or her latent physical aggression status at 8-9 years of age. For instance, 83.7% and 69.6% of high-aggressive 8-9 year-old boys and girls, respectively, were estimated *not* to have changed status at 10-11 years of age. Hence, the majority of high-aggressive children at 8-9 years of age were still displaying an elevated propensity to manifest physically aggressive behaviors two years later.

To compare the stability of the different latent physical aggression statuses, we considered three equal stability submodels: (a) a submodel which stipulates equal stability for low- and medium-aggressive, (b) a submodel which stipulates equal stability for medium- and high-aggressive, and (c) a submodel which stipulates equal stability for low- and high-aggressive. For girls, none of these submodels represents a statistically significant decrease of fit over the unconstrained model; and moreover, each submodel has a lower AIC value than the unconstrained model (see Table 2). In fact, a submodel which combines all three submodels represents an increase of only 1.65 in L^2 with a corresponding increase of 2 in the degrees of freedom from the unconstrained model (p = .44). Hence, the stability of the latent

---

[10] At 8-9 years of age, the estimated proportion of medium-aggressive boys and girls in the Canadian population was 24.6% and 17.8%, respectively.

Table 3:
Conditional Probability Estimates Under the Unconstrained Time-specific Latent Variables Model

Boy Latent class (l = 1, 2, 3)	Low-aggressive (l = 1)		Medium-aggressive (l = 2)		High-aggressive (l = 3)	
	Age (years)		Age (years)		Age (years)	
	8-9	10-11	8-9	10-11	8-9	10-11
Conditional probability estimate						
$\pi_{A(1)l l}$	.818 (.021)	.843 (.016)	.161 (.039)	--	.120 (.054)	.239 (.061)
$\pi_{A(2)l l}$	.178 (.021)	.152 (.016)	.784 (.040)	.986 (.044)	.359 (.074)	.101 (.070)
$\pi_{A(3)l l}$	.004 (.005)	.004 (.003)	.055 (.021)	.014 (.044)	.520 (.076)	.661 (.078)
$\pi_{B(1)l l}$	.962 (.011)	.952 (.008)	.404 (.051)	.315 (.076)	.253 (.073)	.035 (.035)
$\pi_{B(2)l l}$	.038 (.011)	.048 (.008)	.596 (.051)	.685 (.076)	.461 (.076)	.799 (.054)
$\pi_{B(3)l l}$	--	--	--	--	.286 (.067)	.166 (.049)
$\pi_{C(1)l l}$	.963 (.011)	.969 (.007)	.497 (.048)	.409 (.077)	.011 (.028)	.170 (.052)
$\pi_{C(2)l l}$	.037 (.011)	.031 (.007)	.495 (.047)	.552 (.074)	.791 (.059)	.673 (.063)
$\pi_{C(3)l l}$	--	--	.008 (.007)	.038 (.021)	.198 (.056)	.157 (.047)

Girl

Latent class ($\underline{l}$ = 1, 2, 3)	Low-aggressive ($\underline{l}$ = 1)		Medium-aggressive ($\underline{l}$ = 2)		High-aggressive ($\underline{l}$ = 3)	
	Age (years)		Age (years)		Age (years)	
	8-9	10-11	8-9	10-11	8-9	10-11
Conditional probability estimate						
$\pi_{A(1)\|l}$	.823 (.021)	.923 (.048)	.114 (.044)	.420 (.052)	.919 (.065)	.364 (.073)
$\pi_{A(2)\|l}$	.171 (.020)	.071 (.044)	.727 (.048)	.530 (.052)	--	.465 (.073)
$\pi_{A(3)\|l}$	.007 (.005)	.006 (.008)	.160 (.034)	.050 (.021)	.082 (.065)	.171 (.054)
$\pi_{B(1)\|l}$	.983 (.007)	.996 (.007)	.554 (.058)	.892 (.029)	.484 (.126)	.456 (.075)
$\pi_{B(2)\|l}$	.017 (.007)	.005 (.007)	.446 (.058)	.108 (.029)	--	.427 (.073)
$\pi_{B(3)\|l}$	--	--	--	--	.516 (.126)	.117 (.046)
$\pi_{C(1)\|l}$	.989 (.005)	.982 (.006)	.708 (.046)	--	.384 (.112)	--
$\pi_{C(2)\|l}$	.011 (.005)	.018 (.006)	.287 (.045)	--	.526 (.113)	.949 (.031)
$\pi_{C(3)\|l}$	--	--	.005 (.005)	--	.090 (.059)	.051 (.031)

Note. Standard errors appear in parentheses. A dash indicates a boundary estimate. A, B and C refer to the first (i.e., Gets into many fights?), second (i.e., Physically attacks people) and third (i.e., Kicks, bites, hits other children?) behavior item, respectively. $\pi_{j(k)|l}$ refers to the conditional probability of a rating in category k (k = 1, 2, 3) to behavior item j (j = A, B, C) given membership in latent class l (l = 1, 2, 3).

Table 4:
Estimates of the Conditional Probability of a Randomly Selected 10-11 Year-old Child in the Canadian Population Belonging to the Low-, Medium- or High-aggressive Latent Class Given His or Her Latent Physical Aggression Status at 8-9 Years of Age

Latent physical aggression status at 10-11 years of age	Boy			Girl		
Latent physical aggression status at 8-9 years of age	Low-aggressive	Medium-aggressive	High-aggressive	Low-aggressive	Medium-aggressive	High-aggressive
Low-aggressive	.971 (.011) [.675]	.004 (.009)b [.003]	.025 (.008)a [.017]	.863 (.08) [.688]	.137 (.08)b [.109]	--
Medium-aggressive	.623 (.053) [.153]	.377 (.053) [.093]	--	--	.82 (.035) [.146]	.18 (.035)a [.032]
High-aggressive	.05 (.055)b [.003]	.112 (.087)b [.007]	.837 (.101) [.05]	--	.304 (.114)b [.008]	.696 (.114) [.018]

Note. Standard errors appear in parentheses and unconditional probability estimates in brackets. A dash indicates a boundary estimate.
a The estimate's coefficient of variation (i.e., standard error of the estimate / estimate) is greater than 16.5% but less than 33.3%.
b The estimate's coefficient of variation (i.e., standard error of the estimate / estimate) is greater than 33.3%.

physical aggression status of a randomly selected 8-9 year-old girl in the Canadian population does *not* seem to depend on its severity. In contrast, for boys, both the first (L^2 = 682.84 - 642.23 = 40.61; df = 1413 - 1412 = 1; p = .000) and second (L^2 = 657.67 - 642.23 = 15.44; df = 1413 - 1412 = 1; p = .000) (but not the third) submodels represent a statistically significant decrease of fit over the unconstrained model, suggesting that medium-aggressive 8-9 year-old boys are less likely to remain so two years later than their low- and high-aggressive counterparts (see Table 4).

To better characterize change in children's latent physical aggression status, we considered three conditional symmetry submodels: (a) a submodel which stipulates that the likelihood of changing from low- to medium-aggressive is equal to the one of changing from medium- to low-aggressive, (b) a submodel which stipulates that the likelihood of changing from low- to high-aggressive is equal to the one of changing from high- to low-aggressive, and (c) a submodel which stipulates that the likelihood of changing from medium- to high-aggressive is equal to the one of changing from high- to medium-aggressive. For girls, none of the two relevant submodels (the second submodel was not relevant because of boundary parameter estimates) represents a statistically significant decrease of fit over the unconstrained model; and moreover, each submodel has a lower AIC value than the unconstrained model (see Table 2). In fact, a conditional symmetry submodel which combines the first and the third submodels represents an increase of only 2.06 in L^2 with a corresponding increase of 2 in the degrees of freedom from the unconstrained model (p = .36). These results suggest that for girls the likelihood of changing from one latent physical aggression status to another between 8-9 and 10-11 years of age is equal to the one of changing from the latter to the former. For boys, the first (but not the second or third)[11] submodel represents a statistically significant decrease of fit over the unconstrained model (L^2 = 683.33 - 642.23 = 41.10; df = 1413 - 1412 = 1; p = .000), suggesting that the likelihood of a randomly selected 8-9 year-old medium-aggressive boy in the Canadian population becoming low-aggressive two years later is much higher than the one of a low-aggressive boy to become medium-aggressive during the same period (see Table 4). In sum, the results suggest that the development of physical aggression between 8-9 and 10-11 years of age can be characterized by an increase in the propensity to manifest physically aggressive behaviors on a frequent basis. In other words, because there are fewer high- than medium-aggressive 8-9 year-old children; and, in turn, fewer medium- than low-aggressive 8-9 year-old children, there seems to be more children in the Canadian population who start manifesting physically aggressive behaviors on a frequent basis than children who stop doing so between 8-9 and 10-11 years of age. Note, however, that the results suggest the inverse pattern for the propensity to manifest physically aggressive behaviors on an occasional basis, at least for boys.

[11] In fact, a conditional symmetry submodel that combines these two submodels represents an increase of only 2.26 in L^2 with a corresponding increase of 2 in the degrees of freedom from the unconstrained model (p = .32).

Discussion

Information about intraindividual change in a particular phenomenon over time is arguably the single most important piece of information coming out of a longitudinal study. For instance, in the field of developmental psychopathology, one is often interested to know the extent to which a randomly selected child in the population can develop for the first time a particular behavior problem over a specified period of time; or conversely, the extent to which a disturbed child can become problem-free (e.g., Bennett, Lipman, Racine & Offord, 1998; Bennett et al., 1999; Bennett & Offord, 2001). However, current research strategies in developmental psychopathology but also in other related fields often involve categorizing individuals using a less than perfectly sensitive and/or specific cutoff point. Hence, classification errors can not be totally ruled out. For example, the percentages of cases who are correctly classified is likely to be less than 70% in most epidemiological studies of mental disorders (Zarin & Earls, 1993). One aim of this paper was to show how a logit-based time-specific latent variables model could be used to model intraindividual change in physical aggression in the Canadian population of children between 8-9 and 10-11 years of age.

Methodological Discussion

This section consists of a few remarks concerning the different submodels proposed in this paper and how they compare to other models that have been described in the vast literature on the analysis of square contingency tables having ordered categories.

The off diagonal null association submodel

The off diagonal null association submodel is different from the quasi-independence model (Goodman, 1965, 1968). Whereas the later is conditional on those individuals whose status changed over time (i.e., the main diagonal cell entries of the TPM are fitted exactly) the former is not. Hence, one advantage of the (more parsimonious) off diagonal null association submodel is that inferences can be drawn for a particular data set without having to ignore the main diagonal cells where a large proportion of the data are often observed. Moreover, when analyzing the association present in the region off the main diagonal of the TPM this model can provide a useful baseline (null model) for comparative purposes with the symmetric association submodel. Note that the off diagonal null association submodel is not a special case of the row-column association model (Goodman, 1979a; Clogg, 1982a, 1982b); and, therefore, it can not be used to draw inferences about adjacent row and/or column categories distinguishability. On the other hand, the off diagonal null association submodel can provide useful information about latent status “barriers” pertaining to change from one latent status to another in the region on the main diagonal of the TPM. In fact, it is equivalent to the crossings parameter model for the full TPM including the main diagonal cells, a symmetric association model introduced by Goodman (1972). Hence, the greater the odds ratio describing the association in particular 2 x 2 subtable the region on the main diagonal of the TPM the larger the barrier or distance between the two adjacent latent statuses.

Main diagonal uniform association submodel

When analyzing the association present in the region on the main diagonal of the TPM the main diagonal uniform association submodel can provide a useful baseline (null model). Moreover, it can provide useful information about latent status "inertia" pertaining to change from one latent status to another in the region off the main diagonal of the TPM. In fact, in the particular situation where the time-specific latent variables are trichotomous it is equivalent to the diagonals parameter model for the full TPM including the main diagonal cells, a nonsymmetric association model introduced by Goodman (1972). (Otherwise, when the time-specific latent variables have more than three categories the diagonals parameter model is a special case of the main diagonal uniform association submodel.) Hence, if the odds ratio describing the association present in the region above the main diagonal of the TPM is larger (smaller) than the one describing the association present in the region below the main diagonal of the TPM that would suggest a greater (smaller) latent status inertia for change occurring above than below the main diagonal of the TPM.

Submodels for analyzing simultaneously the association present in both regions of the TPM

Consider the particular submodel that consists in having the symmetric association submodel and the main diagonal uniform association submodel fitted simultaneously. In the situation where the time-specific latent variables are trichotomous this hybrid submodel is equivalent to Goodman (1972)'s diagonals-absolute parameter model for the full TPM including the main diagonal cells, a symmetric association model. (Otherwise, when the time-specific variables have more than three categories then the diagonals-absolute parameter model is a special case of this hybrid submodel.) Under this hybrid submodel latent status inertia in the region off the main diagonal of the TPM is the same whether change is occurring above or below the main diagonal of the TPM.

Consider the particular submodel that consists in having the off diagonal null association submodel and the main diagonal uniform association submodel fitted simultaneously. This hybrid submodel is equivalent to a special case of the crossings parameter model that is obtained when all crossing parameters are equal in magnitude (Goodman, 1979b). Under this submodel latent status barrier in the region on the main diagonal of the TPM is uniform. In other words, the distance between any two adjacent latent statuses is the same. Other (less restrictive) hybrid submodels may also be of interest. For instance, one could consider a submodel that combines the main diagonal uniform association submodel with a partial off diagonal null association submodel that applies either above or below the main diagonal of the TPM.

Conditional symmetry submodel

While it is not appropriate to test for exact symmetry and marginal homogeneity when the row/explanatory latent variable marginal needs to be fitted exactly the conditional symmetry submodel can nonetheless provide some useful information about (exact) asymmetry and marginal inhomogeneity. Consider the situation where the time-specific latent variables are dichotomous. Unless the row/explanatory latent variable has a perfectly rectangular marginal distribution conditional symmetry implies that the odds that an observation will fall in one of the minor diagonal cell (i, j) rather than in the corresponding (j, i) cell is not equal to one (i.e., exact symmetry); and therefore, that there are differences in the marginal distributions of the time-specific latent variables. For instance, in the particular situation where the prevalence of a disorder is less than 50% conditional symmetry implies that there are fewer children who became disorder-free during a specified period of time than there are children who developed the disorder for the first time with a corresponding increase in the prevalence of the disorder in question over time.

Symmetric association submodel

The symmetric association submodel is different from the log-multiplicative symmetric association submodel (Goodman, 1979a; Clogg, 1982a, 1982b). The latter is not fitting the main diagonal cells exactly and therefore is more parsimonious. The difference in the likelihood ratio chi-squares for the log-multiplicative submodel and the symmetric association submodel can be used to test the hypothesis that the former model is true under the assumption that there is symmetric association (Goodman, 1979a).

There are other models that may be congruent with this type of data where the row marginal is given (and therefore needs to be fitted exactly by the model), but are more parsimonious than the association models considered in this paper. These are log-linear models for the analysis of the dependence of one column/response variable on a row/explanatory variable where the column marginal is not fitted exactly (although the row marginal is always fitted exactly) (Goodman, 1983). These models would allow to model not only the association between the time-specific latent variables but also the marginal distribution of the column/response latent variable (e.g., uniform versus truncated geometric). Incidentally, the uniform association submodel considered in this paper is a special case of one dependence model, namely, the parallel log-odds model where the adjacent rows are equally spaced (Goodman, 1986, 1987; for details about the relationship between dependence models and association models see Goodman, 1981). In fact, under the uniform association submodel the ratio of the odds that an observation will fall in column j rather than j + 1 in the i^{th} row category and the corresponding odds in the next row category is a constant. Generally, models for the analysis of association are a subset of models for the analysis of dependence.

Substantive Discussion

Our results show that a majority of high-aggressive children at 8-9 years of age will continue two years later to manifest physically aggressive behaviors with a propensity that is much higher than the one of the other children in the Canadian population. From a public policy perspective these results suggest that it may be possible to detect many aggressive children before they enter adolescence where antisocial behaviors involving physical aggression can have very serious consequences. This is especially important since serious forms of antisocial behavior have been found to be highly resistant to treatment in school-aged children and adolescents (Kazdin, 1987). In addition, the cost of violence to society is enormous (World Health Organization, 1999). School-aged aggressive children are at higher risk of alcohol and drug abuse, accidents, violent crimes, depression, suicide attempts, spouse abuse, and neglectful and abusive parenting (Tremblay & LeMarquand, 2001). Note that the possibilities for prevention and treatment by early detection may be further improved by the identification of the factors associated with the persistence of and recovery from aggressive behavior over time as well as the factors associated with the development and non-development of aggressive behavior over time. Because the NLSCY allows to examine the natural history of aggressive behavior within the general population, it should provide us with a unique opportunity to identify these factors.

Longitudinal studies of aggressive behavior in males and females have relied almost exclusively on test-retest correlation coefficients (Olweus, 1979, 1980, 1981, 1984a, 1984b). One limitation of test-retest correlation coefficients is they do *not* directly assess intraindividual change over time but rather the consistency of interindividual differences over time (i.e., the stability of an individual's relative position within a group). The proposed time-specific latent variables model can provide a framework for describing the continuity and discontinuity in children's propensity to manifest physically aggressive behaviors over time. This kind of information has remained concealed in the test-retest correlation coefficients produced by standard methods (Bergman & Magnusson, 1997; Broverman, 1962; Carlson, 1971; Magnusson, 1985, 1995). Note that this is also the case for the semiparametric group based approach (Nagin, 1999) that is becoming increasingly popular for the analysis of developmental trajectories of physical aggression and other behaviors (e.g., Nagin & Tremblay, 1999; Broidy et al., 2003). The few longitudinal studies that investigated intraindividual change in aggressive behavior in children over time used a cutoff point to distinguish between aggressive and non-aggressive children (e.g., Loeber, 1982 who reanalyzed Elliott & Huizinga's (1980) data; Patterson, 1982 who reanalyzed Lefkowitz, Eron, Walder & Huesmann's (1977) data; Loeber, Tremblay, Gagnon & Charlebois, 1989; Shaw, Gilliom & Giovannelli, 2000; Tremblay, Loeber, Gagnon, Charlebois, Larivée & LeBlanc, 1991). However, because none of these studies reported the predictive value (positive or negative) of the cutoff point used the rates of false negative and false positive cannot be determined, and consequently, the impact of classification errors on their assessment of intraindividual change in aggressive behavior over time cannot be determined.

Some of us have suggested that extreme antisocial behavior like physical aggression is due to arrested socialization (e.g., Patterson, 1982). According to this hypothesis aggressive children are simply children who have no grown up, have not learned not to aggress, with their behavior representing that which is normative for younger children. Contrary to the hypothesis of arrested socialization, our results suggest that the development of aggressive behavior during late childhood is characterized by more children learning to aggress on a frequent basis than learning not to. It is often believed that there is a decreasing trend in the prevalence of physical aggression with age (e.g., Tremblay et al., 1999; Tremblay, 2000). Our results suggest that while this may be the case for physically aggressive behaviors manifested on an occasional basis, this is not true for physically aggressive behaviors manifested on a frequent basis, at least for boys. For girls, our results suggest that there is an increase not a decrease of both occasional and frequent physically aggressive behaviors between 8-9 and 10-11 years of age. There is at least another implication of our results for the conceptualization of the development of aggressive behavior during late childhood. According to Rolf

Loeber's (1982; Loeber & Stouthamer-Loeber, 1998) severity hypothesis "... the more extreme the antisocial behavior, the more stable it will be over time." (p. 431). Our results show that the stability of a child's latent physical aggression status is not related to its severity. In fact, contrary to the severity hypothesis, there seems to be a tendency such that the higher a child's propensity to manifest physically aggressive behaviors, the least likely it will remain stable from 8-9 to 10-11 years of age.

References

1. Alder, G. A., & Scher, S. J. (1994). Using growth curve analyses to assess personality change and stability in adulthood. In T. F. Heatherton & J. L. Weinberger (Eds.), Can personality change? (pp. 149-173). Washington, DC: American Psychological Association.
2. Baillargeon, R. H., Tremblay, R. E., & Willms, J. D. (1999). The Prevalence of Physical aggression in Canadian Children: A Multi-Group Latent Class Analysis of Data from the National Longitudinal Survey of Children and Youth (NLSCY). Research report. Applied Research Branch. Human Resources and Development Canada (HRDC).
3. Bassi, F., Hagenaars, J. A., Croon, M. A., & Vermunt, J. K. (2000). Estimating true changes when categorical panel data are affected by uncorrelated and correlated classification errors. An application to Unemployment data. Sociological Methods and Research, 29, 230-268.
4. Bennett, K. J., Lipman, E. L., Brown, S., Racine, Y., Boyle, M. H., & Offord, D. R. (1999). Predicting conduct problems: Can high-risk children be identified in kindergarten and grade 1? Journal of Consulting and Clinical Psychology, 67, 47-480.
5. Bennett, K. J., Lipman, E. L., Racine, Y., & Offord, D. R. (1998). Annotation: Do measures of externalising behaviour in normal populations predict later outcome?: Implications for targeted interventions to prevent conduct disorder. Journal of Child Psychology and Psychiatry, 39, 1059-1070.
6. Bennett, K. J., & Offord, D. R. (2001). Screening for conduct problems: Does the predictive accuracy of conduct disorder symptoms improve with age? Journal of the American Academy of Child and Adolescent Psychiatry, 40, 1418-1425.
7. Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. Development and Psychopathology, 9, 291-319.
8. Bishop, Y. M. M., Fienberg, S. E., Holland, P. W. (1975). Discrete multivariate analysis. Cambridge, MA: The MIT press.
9. Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., Fergusson, D. M., Horwood, J., Loeber, R., Laird, R., Lynam, D., Moffitt, T. E., Pettit, G. S., & Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. Developmental Psychology, 39(2), 222-245.
10. Broverman, D. M. (1962). Normative and ipsative measurement in psychology. Psychological Bulletin, 69, 295-305.
11. Carlson, R. (1971). Where is the person in personality research? Psychological Bulletin, 75, 203-219.
12. Clogg, C. C. (1981). New developments in latent structure analysis. In D. M. Jackson & E. F. Borgatta (Eds.). Factor analysis and measurement in sociological research (pp. 215-246). Beverley Hills, CA: Sage.
13. Clogg, C. C. (1982a). Some models for the analysis of association in multi-way cross-classifications having ordered categories. Journal of the American Statistical Association, 77, 803-815.

14. Clogg, C. C. (1982b). Using association models in sociological research: Some examples. American Journal of Sociology, 88, 114-134.
15. Clogg, C. C., Eliason, S. R., & Grego, J. M. (1990). Models for the analysis of change in discrete variables. In A. von Eye (Ed.), Statistical methods in longitudinal research (Vol. 2, pp. 409-441). San Diego, CA: Academic Press.
16. Clogg, C. C. & Shihadeh, E. S. (1994). Statistical models for ordinal variables. Thousand Oaks, CA: Sage publications.
17. Cochran, W. G. (1968). Errors of measurement in statistics. Technometrics, 10, 637-666.
18. Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. Multivariate Behavioral Research, 27, 131-157.
19. Elliott, D. S., & Huizinga, D. (1980). Defining pattern delinquency: A conceptual typology of delinquent offenses. Paper presented at the meeting of the American Society of Criminology, San Francisco, November.
20. Fienberg, S. E. (1980). The analysis of cross-classified categorical data (2nd ed.). The MIT Press.
21. Goodman, L. A. (1965). On the statistical analysis of mobility tables. American Journal of Sociology, 70, 564-585.
22. Goodman, L. A. (1968). The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries. Journal of the American Statistical Association, 63, 1091-1131.
23. Goodman, L. A. (1972). Some multiplicative models for the analysis of cross classified data. In L. M. Le Cam, J. Neyman, & E. L. Scott (Eds.), Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Theory and Statistics (pp. 649-696). Berkeley, CA: University of California Press.
24. Goodman, L. A. (1979a). Simple models for the analysis of association in cross-classifications having ordered categories. Journal of the American Statistical Association, 74, 537-552.
25. Goodman, L. A. (1979b). Multiplicative models for the analysis of occupational mobility tables and other kinds of cross-classification tables. American Journal of Sociology, 84, 804-819.
26. Goodman, L. A. (1981). Three elementary views of log-linear models for the analysis of cross-classification having ordered categories. In S. Leinhardt (Ed), Sociological Methodology (vol. 12, pp. 193-239). London: Jossey-Bass Publishers.
27. Goodman, L. A. (1983). The analysis of dependence in cross-classifications having ordered categories, using log-linear models for frequencies and log-linear models for odds. Biometrics, 39, 149-160.
28. Goodman, L. A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. International Statistical Review, 54, 243-270.
29. Goodman, L. A. (1987). New methods for analyzing the intrinsic character of qualitative variables using cross-classified data. American Journal of Sociology, 93, 529-583.
30. Hagenaars, J. A. (1990). Categorical longitudinal data. Log-linear panel, trend, and cohort analysis. Newbury Park, CA: Sage publications.
31. Hagenaars, J. A. (1993). Loglinear models with latent variables. Newbury Park, CA: Sage.
32. Hagenaars, J. A. (1994). Latent variables in log-linear models of repeated observations. In A. von Eye & C. C. Clogg (Eds.), Latent variables analysis: Applications for developmental research (pp. 329-352). Thousand Oaks, CA: Sage publications.
33. Kazdin, A. E. (1987). Conduct disorders in childhood and adolescence. Newbury Park, CA: Sage Publications.
34. Kraemer, H. C. (1979). Ramifications of a population model for k as a coefficient of reliability. Psychometrika, 44, 461-472.

35. Kraemer, H. C. (1985). The robustness of common measures of 2 x 2 association to bias due to misclassifications. The American Statistician, 39, 286-290.
36. Lefkowitz, M. M., Eron, L. D., Walder, L. O., Huesmann, L. R. (1977). Growing up to be violent: A longitudinal study of the development of aggression. Pergamon Press.
37. Loeber, R. (1982). The stability of antisocial and delinquent child behavior: A review. Child Development, 53, 1431-1446.
38. Loeber, R., Stouthamer-Loeber, M. (1998). Development of juvenile aggression and violence. Some common misconceptions and controversies. American Psychologist, 53, 242-259.
39. Loeber, R., Tremblay, R. E., Gagnon, C., & Charlebois, P. (1989). Continuity and desistance in disruptive boys' early fighting at school. Development and Psychopathology, 1, 39-50.
40. Magnusson, D. (1985). Implications of an interactional paradigm for research on human development. International Journal of Behavioral Development, 8, 115-137.
41. Magnusson, D. (1995). Individual development: A holistic, integrated model. In P. Moen, G. H. Elder Jr. & K. Lüscher (Eds.), Examining lives in context: Perspectives on the ecology of human development (pp. 19-60). Washington, DC: American Psychological Association.
42. Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. Psychological Methods, 4, 139-157.
43. Nagin, D. S., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child Development, 70, 1181-1196.
44. NLSC Project Team (1995a). National Longitudinal Survey of Children: Overview of survey instruments for 1994-95, data collection, cycle 1 (Catalogue No. 95-02). Statistics Canada and Human Resources Development Canada.
45. NLSC Project Team (1995b). National Longitudinal Survey of Children: Survey instruments for 1994-95, data collection, cycle 1 (Catalogue No. 95-01). Statistics Canada and Human Resources Development Canada.
46. NLSCY Project Team (1997a). National Longitudinal Survey of Children & Youth: Overview of survey instruments for 1996-97 data collection, cycle 2 (Catalogue No. 89FOO78XPE). Statistics Canada and Human Resources Development Canada.
47. NLSCY Project Team (1997b). National Longitudinal Survey of Children & Youth: Cycle 2 survey instruments 1996-97 (Book 1 - Parent & child) (Catalogue No. 89FOO77XPE). Statistics Canada and Human Resources Development Canada.
48. Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. Psychological Bulletin, 86, 852-875.
49. Olweus, D. (1980). The consistency issue in personality psychology revisited-with special reference to aggression. British Journal of Social and Clinical Psychology, 19, 377-390.
50. Olweus, D. (1981). Continuity in aggressive and withdrawn, inhibited behavior patterns. Psychiatry and Social Science, 1, 141-159.
51. Olweus, D. (1984a). Development of stable aggressive reaction patterns in males. In R. J. Blanchard & D. C. Blanchard (Eds.), Advances in the study of aggression, Vol. 1 (pp. 103-137). Orlando, FL: Academic Press.
52. Olweus, D. (1984b). Stability in aggressive and withdrawn, inhibited behavior patterns. In R. M. Kaplan, V. J. Konecni, & R. W. Novaco (Eds.), Aggression in children and youth (pp. 104-137). The Hague: Martinus Nijhoff.
53. Patterson, G. R. (1982). A social learning approach. Vol. 3 Coercive family process. Eugene, OR: Castalia.
54. Plewis, I. (1985). Analyzing change: Measurement and explanation using longitudinal data. New York: John Wiley & Sons.

55. Shaw, D. S., Gilliom, M., & Giovannelli, J. (2000). Aggressive behavior disorders. In C. H. Zeanah Jr. (Ed.), Handbook of infant mental health (2nd ed) (pp. 397-411). New York: Guilford Press.
56. Tremblay, R. E. (2000). The development of aggressive behaviour during childhood: What have we learned in the past century? International Journal of Behavioral Development, 24, 129-141.
57. Tremblay, R. E., Japel, C., Pérusse, D., Boivin, M., Zoccolillo, M., Montplaisir, J., & McDuff, P. (1999). The search for the age of «onset» of physical aggression: Rousseau and Bandura revisited. Criminal Behavior and Mental Health, 9, 8-23.
58. Tremblay, R. E., & LeMarquand, D. (2001). Individual risk and protective factors. In R. Loeber & D. Farrington (Eds.), Child delinquents: Development, interventions and service needs (pp. 137-164). Thousand Oaks, CA: Sage Publications.
59. Tremblay, R. E., Loeber, R., Gagnon, C., Charlebois, P., Larivée, S., & LeBlanc, M. (1991). Disruptive boys with stable and unstable high fighting behavior patterns during junior elementary school. Journal of Abnormal Child Psychology, 19, 285-300.
60. van de Pol, F., & Langeheine, R. (1990). Mixed Markov latent class models. In C.C. Clogg (Ed.), Sociological methodology 1990 (pp. 213 - 247). Oxford: Blackwell.
61. Vermunt, J. K. (1997a). Log-linear models for event histories. Thousand Oaks, CA: Sage publications.
62. Vermunt, J. K. (1997b). lEM: A general program for the analysis of categorical data [computer program]. Tilburg University.
63. World Health Organization (1999).WHO information series in school health. Document three. Violence prevention: An important element of a health-promoting school. Geneva.
64. Yanagawa, T., & Gladen, B. C. (1984). Estimating disease rates from a diagnostic test. American Journal of Epidemiology, 119, 1015-1023.
65. Zarin, D. A., & Earls, F. (1993). Diagnostic decision making in psychiatry. American Journal of Psychiatry, 150, 197-206.

Appendix

lEM-code for some Time-specific Latent Variables Models presented in Table 2

Model 3. Uniform association

```
LEM: log-linear and event history analysis with missing data.
Developed by Jeroen Vermunt (c), Tilburg University, The Neth-
erlands. Version 1.0 (September 18, 1997).
* X = Latent physical aggression variable at 8-9 years of age
* Y = Latent Physical aggression variable at 10-11 years of
age
* In 1994-1995
* abecq6g:  Gets into many fights?
* abecq6aa: Physically attacks people?
* abecq6nn: Kicks, bites, hits other children?
* In 1996-1997
* bbecq6g:  Gets into many fights?
```

```
* bbecq6aa: Physically attacks people?
* bbecq6nn: Kicks, bites, hits other children?
* A = abecq6g  (i.e., 1 = never; 2 = sometimes; 3 = often)
* B = abecq6aa (i.e., 1 = never; 2 = sometimes; 3 = often)
* C = abecq6nn (i.e., 1 = never; 2 = sometimes; 3 = often)
* D = bbecq6g  (i.e., 1 = never; 2 = sometimes; 3 = often)
* E = bbecq6aa (i.e., 1 = never; 2 = sometimes; 3 = often)
* F = bbecq6nn (i.e., 1 = never; 2 = sometimes; 3 = often)
* G = Age (in years) of the child in 1994-1995 (i.e., 1 = 8; 2
= 9)
* S = Sex of the child (i.e., 1 = girl; 2 = boy)
lat 2
man 8
dim 3  3  2  2  3  3  3  3  3  3
lab X  Y  S  G  A  B  C  D  E  F
mod GS X|S Y|XS {YS, spe(XY,1b,S,c)}
A|XS
B|XS
C|XS
D|YS
E|YS
F|YS
```

Model 4. Symmetric association

```
lat 2
man 8
dim 3  3  2  2  3  3  3  3  3  3
lab X  Y  S  G  A  B  C  D  E  F
mod GS X|S Y|XS {YS,fac(YXS,6)}
A|XS
B|XS
C|XS
D|YS
E|YS
F|YS
des [1 0 5
     0 0 0
     5 0 2

     3 0 6
     0 0 0
     6 0 4]
```

Model 5. Main diagonal uniform association

```
lat 2
man 8
dim 3   3   2   2   3   3   3   3   3   3
lab X   Y   S   G   A   B   C   D   E   F
mod GS  X|S  Y|XS  {YS, fac(YXS,6)}
A|XS
B|XS
C|XS
D|YS
E|YS
F|YS
des  [1  0  2
      0  0  0
      5  0  1

      4  0  3
      0  0  0
      6  0  4]
```

Model 6. Off diagonal null association

```
lat 2
man 8
dim 3   3   2   2   3   3   3   3   3   3
lab X   Y   S   G   A   B   C   D   E   F
mod GS  X|S  Y|XS  {YS,fac(YXS,4)}
A|XS
B|XS
C|XS
D|YS
E|YS
F|YS
des  [1  0  0
      0  0  0
      0  0  2

      3  0  0
      0  0  0
      0  0  4]
```

Model 7. Low = medium

```
lat 2
man 7
dim 3  3  2  3  3  3  3  3  3
lab X  Y  G  A  B  C  D  E  F
mod G X Y|X eq2
A|X
B|X
C|X
D|Y
E|Y
F|Y
des [2 0 0
     0 2 0
     0 0 0]
```

Model 10. M|L = L|M

```
lat 2
man 7
dim 3  3  2  3  3  3  3  3  3
lab X  Y  G  A  B  C  D  E  F
mod G X Y|X eq2
A|X
B|X
C|X
D|Y
E|Y
F|Y
des [0 2 0
     2 0 0
     0 0 0]
```