Applications of confirmatory latent class analysis in developmental psychology

Olav Laudy
Department of Methodology and Statistics, Utrecht University, The Netherlands

Mark Zoccolillo
Department of Psychiatry, McGill University, Canada

Raymond H. Baillargeon
Department of Developmental Psychiatry, University of Montréal, Canada

Jan Boom
Department of Developmental Psychology, Utrecht University, The Netherlands

Richard E. Tremblay
Canada Research Chair in Child Development, University of Montréal, Canada

Herbert Hoijtink
Department of Methodology and Statistics, Utrecht University, The Netherlands

In the field of developmental, psychology researchers may have several competing theories with respect to their research subject. In this paper an approach will be proposed that can be used to select the best of these theories. It will be shown that a theory can be translated in a constrained latent class model using inequality constraints. This can be done for several (possibly competing) theories. Subsequently, fit-measures can be used to determine which model (and thus which theory) is supported most by the data. The approach will be introduced using data with respect to self-reported child and adult antisocial behaviour. It will be further illustrated using data obtained using the figural intersection task.

Exploratory latent class analysis (ELCA)(Clogg, 1981; Goodman, 1974; Haberman, 1988; Vermunt, 1996) is a statistical technique that can be used to divide a sample of respondents into homogeneous subgroups, also called

Correspondence should be addressed to Olav Laudy, Utrecht University, Faculty of Social Sciences, Department of Methodology and Statistics, Postbus 80140, NL-3508 TC, Utrecht, The Netherlands. Email: o.laudy@fss.uu.nl

latent classes. In developmental psychology, for example, data with respect to some cognitive ability of children, like responses to items on Piaget's balance scale task, Piaget's water level task or the figural intersection task can be subjected to an ELCA. This can result in groups of children using different strategies (e.g., Boom, Hoijtink, & Kunnen, 2001; Jansen & van der Maas, 1997; Pascual-Leone & Baillargeon, 1994) and the transitions between the groups (Hoben & Hettmansperger, 2001). For a methodological overview and applications see von Eye and Clogg (1994).

A key question in ELCA is: into how many homogeneous subgroups should the sample be divided? Usually fit measures (Everitt, 1988; Lin & Dayton, 1997) are used to determine which number of classes is optimal. Furthermore, the resulting classes have to be interpreted. To illustrate this, consider an ELCA of the responses of 2001 women to several items with respect to self-reported child and adult antisocial behaviours. In order to assess maternal antisocial disorder, the mothers were asked about five childhood antisocial behaviours and four adult behaviours. In Table 1 the items (responses to either "yes" or "no" questions) and the response percentages can be found. The women were mothers of 5-month-old infants

TABLE 1
Items indicative of child and adult antisocial behaviour

j	Item phrasing	Item type	Yes (%)
1	Before the end of high school, did you more than once swipe things from stores or from other children, or steal from your parents or from anyone else?	Child	18
2	Before the end of high school, did you more than once get into fights that you started?	Child	3
3	Before the end of high school, were you ever involved with Social Services (Department of Youth Protection), in trouble with the police or arrested BECAUSE OF YOUR OWN MISBEHAVIOUR?	Child	4
4	Before the end of your high school, did you ever skip school at least twice in one year?	Child	48
5	Before the end of your high school, did you ever run away from home overnight?	Child	9
6	Since finishing school, have you been FIRED from your job (do not take account layoffs resulting from lack of work)?	Adult	9
7	Since finishing school, have you ever been arrested for anything OTHER than traffic violations?	Adult	1
8	Since finishing school, did you ever hit or throw things at your spouse (or partner that you were living with)?	Adult	10
9	Since finishing school, have you ever been in trouble at work, with the police or with your family, or had a car accident BECAUSE OF DRUGS OR ALCOHOL?	Adult	9

from a population-based longitudinal study of the development of children of the province of Quebec, Canada (Jette, Desrosiers, Tremblay & Thibault, 2000; Zoccolillo, 2000).

In the top panel of Table 2, three fit measures for exploratory analyses with two, three, and four classes are displayed. Hoijtink (1998, 2001) developed these fit measures to be able to select the best of a number of *exploratory* and *confirmatory* (see below) latent class models. The first measure is −2 log of the marginal likelihood (Kass & Raftery, 1995). This measure can be seen as the Bayesian counterpart to information criteria like AIC, CAIC and BIC (Kass & Raftery, 1995). It is a relative (with respect to the corresponding numbers for other models) fit measure: the smaller the number the better the model. If, a priori, each model is considered to be equally good, −2 log of the marginal likelihood can be transformed to posterior probabilities (Kass & Raftery, 1995), which are easier to interpret. The third measure is an absolute (how good is the model at hand) fit measure. It is a modification of the likelihood ratio goodness of fit test (see, for example, Formann, 1985) that is traditionally used in latent class analysis. For this test, *p*-values smaller than .05 indicate that the model at hand is not able to reproduce adequately the frequencies with which each response pattern in the data matrix is observed.

As can be seen in Table 2, the exploratory model with three classes appears to be the best model. Note that this does not exclude the possibility that in the population the women are grouped in two or four classes. In exploratory latent class analysis the number of classes can only approximately be determined using fit measures (see, for example, Everitt, 1988; Hoijtink, 1998, 2001; Lin & Dayton, 1997). In Table 3, estimates of the parameters can be found of the three-class exploratory model. These estimates are obtained using Bayesian computational statistics; the interested reader is referred to Hoijtink and Molenaar (1997) and Hoijtink (1998). Each class gets a weight, that is the proportion of persons belonging

TABLE 2
Evaluation of exploratory and confirmatory models for the antisocial behaviour data

Model	−2 log marginal likelihood	Post. prob.	p-values
Exploratory two classes	9602	.00	.12
Exploratory three classes	9564	.91	.23
Exploratory four classes	9570	.09	.22
One-dimensional	9545	.08	.33
Two-dimensional	9540	.92	.38
Three-class	9553	.00	.29

TABLE 3
Class-specific probabilities for the three class exploratory model

Item type	j	Class 1	Class 2	Class 3
Child	1	.07	.74	.54
	2	.01	.60	.09
	3	.01	.80	.12
	4	.37	.80	.81
	5	.02	.74	.30
Adult	6	.07	.47	.16
	7	.01	.46	.04
	8	.05	.37	.26
	9	.01	.42	.03
		$\omega_1 = .77$	$\omega_2 = .01$	$\omega_3 = .23$

to that class, and each class has its own class-specific probabilities, that is the probability of answering "yes" to each of the items. Note that j will be used to indicate item numbers, ω_1 denotes the proportion of persons in class 1, and π_{2j} denotes the probability of responding "yes" to item j in class 2.

The class-specific probabilities can be used to interpret the nature of the class. Looking at Table 3, class one will probably be labelled "school-skippers" because the probability of skipping school at least twice in one year (see Table 1 for the item labels) is rather large (37%). However, class 1 might also be labelled "normal children" because skipping school is not really unusual behaviour, even for "normal children". Class 2 will probably be labelled as persons who have a high probability of some form of antisocial behaviour, either in childhood or adulthood. The interpretation of class 3 is less straightforward. Different researchers may attach different labels to class 3. This post-hoc labelling of the latent classes is one of the drawbacks of exploratory latent class analysis. Even if a researcher has a number of theories with respect to the nature of the classes in the population of women, an exploratory analysis may not help to decide which theory is the best, because the number of classes chosen and the labelling attached may not straightforwardly indicate which theory is the best.

A researcher using exploratory analysis behaves as if his research field has not yet been explored very thoroughly, and theories do not exist. This, however, is not always an accurate reflection of the true state of affairs. ELCA has been done in areas that have been thoroughly explored, and where theories are well developed (Boom, Hoijtink, & Kunnen, 2001; Hoben & Hettmansperger, 2001; Jansen & van der Maas, 1997; Pascual-Leone & Baillargeon, 1994). The plausibility of real scientific progress is larger if the current state of affairs (existing knowledge and theories) is properly accounted for in the statistical models used for the analysis.

In this paper a specific form of confirmatory latent class analysis (CLCA) will be proposed (Hoijtink, 1998, 2001; Hoijtink & Molenaar, 1997). It will be shown how theories can be translated into a CLCA using inequality constraints among the parameters of the model. Subsequently, the best model, that is, the best theory, can be selected.

In the next section, the approach proposed will be illustrated continuing the example concerning parental antisocial behaviour (Zoccolillo, Pickles, Quinton, & Rutter, 1992; Zoccolillo, Price, Ji, & Hwu, 1999). In the third section, some of the hypotheses and theories described in Pascual-Leone and Baillargeon (1994) concerning the development of mental attention will investigated using CLCA. The paper will be concluded with a discussion.

THE DIMENSIONS OF ANTISOCIAL BEHAVIOUR

In this section, the data with respect to antisocial behaviour will be analyzed using confirmatory latent class analysis (CLCA) (Hoijtink, 1998, 2001; Hoijtink & Molenaar, 1997). This approach consists of four steps. In the first step a researcher has to make an inventory of existing theories with respect to the research domain and research questions of interest. In the second step each theory has to be formalized. As will be illustrated below, this will be done using inequality constraints among the parameters of latent class models. In the third step the fit measures introduced in the previous section will be used to determine which constrained latent class model (and thus which theory) is supported most by the data. In a possible fourth step the parameter estimates for the best model are inspected. Note that for each model the inequality constraints are active both when fit measures are computed and when parameters are estimated.

For these data, three theories exist. The first model reflects the theory that antisocial behaviour is a invariant characteristic of human personality, affecting some persons more than others. The corresponding inequality constraints for this theory restrict the class-specific probabilities such that these increase with the number of the latent class, independent of whether it concerns child or adult items. This is illustrated in Table 4, where it can be seen that with increasing class number, both the child and adult items have higher class-specific probabilities. Note that these constraints imply a one-dimensional structure for the data; low-class numbers indicate a low degree of antisocial behaviour, high-class numbers indicate a high degree.

The second model reflects the theory that child and adult antisocial behaviour are separate concepts. Being antisocial in childhood does not imply being antisocial in adulthood and visa versa. This is illustrated in Table 5. The inequality constraints are implicit, a minus sign indicates that the corresponding class-specific probability is smaller than the probabilities associated with a plus sign. Persons in the first and second class are

constrained to have lower probabilities of responding positively to the items indicative of childhood antisocial behaviour than persons in the third and fourth class. Persons in the first and third class are constrained to have lower probabilities of responding positively to the items indicative of adulthood antisocial behaviour than persons in the second and fourth class. This leads to a two-dimensional structure, i.e., the four latent classes separate in not being antisocial in both childhood and adulthood, being antisocial in one of these, and being antisocial in both of them. Note that in Table 5, the constraints apply to each row, e.g., the class-specific probabilities for item one in classes one and two are restricted to be smaller than the class-specific probabilities for item one in classes three and four.

   Longitudinal and epidemiologic studies have found that there are broadly three groups of women: those without significant antisocial behaviour in childhood or adulthood; those with significant antisocial behaviour in childhood but not adulthood; and those with significant antisocial behaviour in childhood and adulthood (Zoccolillo et al., 1999).Women with significant antisocial behaviour in adulthood but not childhood are rare. The third model reflects this theory that many antisocial children do

TABLE 4
Inequality constraints for the one-dimensional theory

Item type:	j	Class 1		Class 2		Class 3		Class 4
Child	1	π_{11}	<	π_{21}	<	π_{31}	<	π_{41}
	...	π_{1j}	<	π_{2j}	<	π_{3j}	<	π_{4j}
	5	π_{15}	<	π_{25}	<	π_{35}	<	π_{45}
Adult	6	π_{16}	<	π_{26}	<	π_{36}	<	π_{46}
	...	π_{1j}	<	π_{2j}	<	π_{3j}	<	π_{4j}
	9	π_{19}	<	π_{29}	<	π_{39}	<	π_{49}

TABLE 5
Inequality constraints for the two-dimensional theory. Inequality constrains are implicit for each row of the table: $- < +$

Item type:	j	Class 1	Class 2	Class 3	Class 4
Child	1	−	−	+	+
	...	−	−	+	+
	5	−	−	+	+
Adult	6	−	+	−	+
	...	−	+	−	+
	9	−	+	−	+

not become antisocial adults, but that most antisocial adults were antisocial children. The inequality constraints for this theory are equal to the two-dimensional structure without class two.

In the bottom panel of Table 2, it can be seen that the two-dimensional model receives the most support from the data. The *p*-value of the likelihood ratio test is highest for the two-dimensional model. This indicates (.38 is larger than .05) that this model adequately reproduces the frequencies with which the response vectors are observed in the data matrix. The two-dimensional model also has the smallest −2 log marginal likelihood. Assuming that, a priori, each model is equally likely, this implies that after observing the data, the two-dimensional model has a probability of 92% of being the best of the three models under investigation.

The confirmatory approach does not have the drawbacks of the exploratory approach. The number of classes and the labelling of these classes have been specified *before* the analysis, using existing theories and knowledge. Subsequently, the data are used only to decide which of these specifications is the best. This leaves no room for over-interpretation, or interpretations of results (see the previous section) that differ among researchers. This does not imply that the confirmatory approach is always preferred over the exploratory approach. However, if the existing theories and knowledge can be formalized in a number of competing models, it is an excellent tool to select the best model. This includes the situation where an exploratory analysis on a first data set is used to generate theories, and a confirmatory analysis of a second data set is used to select the best of these theories. An elaborate example of the latter will be given in the next section.

In Table 6, the estimates of the class-specific probabilities and class weights are displayed for the two-dimensional model. Note that these estimates are obtained accounting for the inequality constraints imposed on the class-specific probabilities. Class two accounts for the persons who have low probabilities of child antisocial behaviour, and high probabilities of adult antisocial behaviour. Since this class accounts for 17.4% of all persons, it is clear why the three-class model, where this class was left out, does not have a good fit. Class four contains 1% of the persons. This is realistic, because this class contains the persons that are quite antisocial, both in childhood and adulthood.

THE COGNITIVE DEVELOPMENT OF FIVE-YEAR-OLD CHILDREN

Introduction

In this section models concerning the relation between mental attentional resources and performance will be tested using the figural intersection task

TABLE 6
Class-specific probabilities and class weights for the two-dimensional model

Item type	j	Class 1 Low child Low adult	Class 2 Low child High adult	Class 3 High child Low adult	Class 4 High child High adult
Child	1	.05	.41	.64	.76
	2	.01	.05	.21	.61
	3	.00	.06	.36	.82
	4	.35	.65	.87	.87
	5	.01	.15	.55	.75
Adult	6	.06	.23	.10	.55
	7	.00	.09	.02	.52
	8	.03	.32	.13	.46
	9	.00	.12	.01	.48
		$\omega_1 = .71$	$\omega_2 = .17$	$\omega_3 = .10$	$\omega_4 = .01$

(FIT) (Pascual-Leone & Baillargeon, 1994). According to these authors, mental attention, a test's mental demand and performance have very specific relationships, which can be formalized into five theoretical structural predictions (TSP). In their paper they attempt to verify these five predictions with the help of latent class analyses. Although their predictions are presented and described using inequality constraints, the methodology and software necessary to translate their theories into CLCA of the type discussed in this paper was not available at that time. They had to resort to various other types of LCA. In this section part of their theories will be translated into CLCA using inequality constraints. The parts of their theory used in this paper are summarized below. For the complete theory, interested readers are referred to Pascual-Leone and Baillargeon (1994). Subsequently, new data will be used to determine which theory is the best.

The figural intersection task was administered to 106 children and 96 of them completed all test items. These subjects took part in the pilot survey of the Québec longitudinal survey of child development (QLSCD). The total population for the QLSCD's pilot survey were babies of French- or English-speaking mothers residing in seven administrative regions of the province of Québec, Canada. At wave one, babies were between 58 and 61 months of age.

The theory

The idea that mental attentional "energy" or capacity (nowadays often thought of as "working memory") is essential in cognitive development might go back to Binet (Pascual-Leone & Baillargeon, 1994). In this

research the mental capacity is referred to as M-capacity, which is the number of schemes that a subject can simultaneously boost into activation. The M-capacity is a limited resource. A related property of a test item is M-demand, the minimum M-capacity that a respondent needs to solve an item using a stipulated solution strategy.

Every item of the FIT presents on the right-hand side of the page a number of geometric shapes separated from each other (see Figure 1 for an example). Shapes on the right-hand side are the task's relevant figures. Every item has two subtasks. The first is to place a dot inside each figure found on the right-hand side of the page to ensure proper exploration of all relevant figures. The second is to place on the left-hand side a single dot that is inside all relevant figures at the same time, i.e., to find the total intersection. The placement of the final dot determines whether the response is judged to be correct or incorrect.

For 5-year-old children, the appropriate number of figures in each item is two or three. On the left hand there is a figural compound with all relevant (i.e., right-hand side) figures. In other items of the task this figural compound may contain one irrelevant figure, not found on the right-hand side. When present, the irrelevant figure is a distracter that must be ignored.

The version of the FIT used here consists of 8 items, two items per item type: two relevant figures; two relevant figures and one irrelevant figure; three relevant figures; and three relevant figure and one irrelevant figure.

Children could solve an item by trying all combinations of the intersections one by one to find the common intersection. This strategy does not lead a correct estimate of a child's M-capacity, because the different figures or "schemes" are not boosted at once into the memory. To discourage this partial sectioning strategy children have to practice the FIT, using items that contain figures that are irrelevant, i.e., that do not have a common intersection with the other figures. Thus, in the training the child learns that the partial sectioning strategy is not a good one, because it will not lead to the correct solution.

The theoretical structural predictions and an alternative model are displayed in Tables 7 and 8. The point of departure is a latent class model,
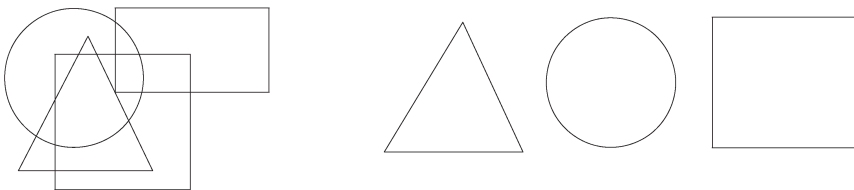
Figure 1. An item from the figural intersection task.

TABLE 7
Inequality constraints for Pascual-Leone and Baillargeon, TSP 123 and TSP 12344*

j	Class 1		Class 2		Class 3		Class 4		Class 5
1,2	L	<	H		H		H		H
		↖	∨	↘		↘		↘	
3,4	L		L	<	H		H		H
		↖		↖	∨	↘		↘	
5,6	L		L		L	<	H		H
		↖		↖		↖	∨	↘	
7,8	L		L		L		L	<	H

TSP 1:	∨
TSP 2:	<
TSP 3:	L,H
TSP 4:	↘
TSP 4*:	↖

TABLE 8
Inequality constraints for Pascual-Leone, double monotone model

j	Class 1		Class 2		Class 3		Class 4		Class 5
1,2	π_{1j}	<	π_{2j}	<	π_{3j}	<	π_{4j}	<	π_{5j}
	∨		∨		∨		∨		∨
3,4	π_{1j}	<	π_{2j}	<	π_{3j}	<	π_{4j}	<	π_{5j}
	∨		∨		∨		∨		∨
5,6	π_{1j}	<	π_{2j}	<	π_{3j}	<	π_{4j}	<	π_{5j}
	∨		∨		∨		∨		∨
7,8	π_{1j}	<	π_{2j}	<	π_{3j}	<	π_{4j}	<	π_{5j}

where it is assumed that there are four M-demand levels for the items (the most easy are the items with two figures, the most difficult items have three figures and one irrelevant figure). With four different M-demand levels, there can be five M-capacity classes: the first class where children fail any item; the second where the children pass the first item type and fail the others; up to the fifth class where the children pass all items. From this TSP 1 and 2 it is possible to derive: whether an item has a higher M-demand than the child's M-capacity, the probability of a correct response is smaller than if the item has a smaller M-demand than the child's M-capacity; and, if a child has a higher M-capacity than the item's M-demand, the probability of a correct response is larger than if the child has a smaller M-capacity than the item's M-demand. This implies a boundary as indicated by the ∨ and < in

Table 7. The TSP 3 in the original theory implies that probabilities below the boundary are equal, and that probabilities above the boundary are equal. This reflects the idea of Pascual-Leone and Baillargeon that the only relevant aspect of M-capacity is whether or not it is bigger than the M-demand. We relaxed this aspect of their theory such that the probabilities (denoted by H[igh]) are restricted to be larger than the probabilities (denoted by L[ow]). These three prediction are joined in one set and will be called model TSP 123.

Pascual-Leone and Baillargeon also present TSP 4 and 4*. These predictions reflect the theory that if the M-capacity is larger than the M-demand, it is easier to master items with a low M-demand, while having a low M-capacity (for example, responding correctly to items 1 and 2 while in class 2) than to master items with a higher M-demand, while having a higher M-capacity (for example, responding correctly to items 5 and 6 while in class 4). If the M-capacity is smaller than the M-demand, this relation is reversed: it is easier to master items with a low M-demand while having a low M-capacity (for example, responding correctly to items 1 and 2 while in class 1) than to master items with a higher M-demand while having a higher M-capacity (for example, responding correctly to items 5 and 6 while in class 3). TSP 4 and 4* reflect the M-capacity/M-demand difference effect (Pascual-Leone & Baillargeon, 1994): whenever one simultaneously varies M-capacity and M-demand the greater the M-capacity/M-demand ratio the greater the probability of success. This can be formalized by restricting the upper triangle of Table 7, such that the probability of correctly responding to an item gets smaller in the down diagonal direction. This model will be called TSP 1234. In the lower triangle this relation is reversed: the probability of correctly responding to an item increases in the down diagonal direction. This model will be called TSP 1234*. The total set of the predictions will be called model TSP 12344*.

Hoijtink and Molenaar (1997) present the latent class equivalent of a non-parametric item response model that might be appropriate for the data at hand. The five-class version of this model is presented in Table 8. As can be seen, the main difference with respect to model TSP 123 is the absence of a diagonal boundary. This, so called, double monotonous model assumes an ordering of both the items and the classes, i.e., M-capacity increases with class number, M-demand increases with item complexity. As can be seen in Table 8, this model implies inequality constraints in two directions. In the horizontal directions the probability of a correct response is restricted to increase with class number (that is M-capacity), in the vertical direction the probability of a correct response is restricted to decrease with item complexity.

This double monotone model contains two simpler models. One with row ordering, i.e., M-demand increases with item complexity, and one with column ordering, i.e., M-capacity increases with class number. Both the

double monotone and the two simpler models will be included in the set of models that will be investigated.

Results

Table 9 shows that TSP 1234 is the best model. The *p*-value of the likelihood ratio test indicates that this model adequately predicts the frequencies with which each response vector is observed. Furthermore, assuming that, a priori, each model is equally likely, the posterior probability is 71%, indicating that in comparison with the other models, this model is the best model. Note that TSP 1234 contains TSP 123, which has a posterior probability of 27%. Stated another way, the support for TSP 123 is huge (27% plus 71%) and the support for TSP 1234 substantial (71%). Estimates of the class-specific probabilities and class weights for the TSP 1234 model can be found in Table 10. The 95% central credibility interval (the Bayesian counterpart of a confidence interval) can be found within brackets. These give the precision with which each of the numbers is estimated. As can be seen, the credibility intervals (also computed to account for the constraints) have an average range of about .30 (distance between lower and upper bound). This is caused by the relatively small sample of children used for the analysis. Nevertheless, the approach proposed is able to select the best of the seven theories (see Table 9) under investigation.

TSP 1234 is obtained by adding TSP 4 to TSP 123. Adding TSP 4* to either the TSP 123 or the TSP 1234 model yields a substantially worse-fitting model. In fact, neither the TSP 1234* nor TSP 12344* model adequately predicts the observed frequencies of each response pattern (*p*-values smaller than .05). Therefore these results suggest that TSP 1, TSP 2, TSP 3 and TSP 4 adequately describe children's performance on the FIT, but TSP 4* does not. TSP 1234 implies a misleading effect caused by the context when one simultaneously varies M-capacity and M-demand, and when M-capacity is larger than M-

TABLE 9
Evaluation of different models for the FIT data

Model	− 2 log marginal likelihood	Post. prob.	p-value
TSP 123	689	.27	.21
TSP 1234	687	.71	.25
TSP 1234*	699	.00	.03
TSP 12344*	699	.00	.01
Double monotone	702	.00	.20
Row ordering	699	.00	.22
Column ordering	694	.02	.24

TABLE 10
Parameter estimates of the TSP 1234, between () central credibility intervals

j	Class 1	Class 2	Class 3	Class 4	Class 5
1	.23 (.05–.44)	.96 (.87–.99)	.96 (.90–.99)	.89 (.73–.99)	.56 (.17–.97)
2	.06 (.00–.21)	.95 (.85–.99)	.97 (.92–.99)	.94 (.80–.99)	.47 (.06–.94)
3	.05 (.00–.18)	.34 (.04–.62)	.88 (.74–.97)	.90 (.77–.97)	.60 (.25–.90)
4	.04 (.00–.17)	.16 (.01–.43)	.86 (.71–.96)	.88 (.73–.97)	.44 (.09–.85)
5	.06 (.00–.21)	.14 (.01–.33)	.19 (.00–.47)	.74 (.52–.90)	.61 (.24–.89)
6	.10 (.00–.29)	.36 (.13–.59)	.45 (.11–.70)	.78 (.62–.91)	.72 (.44–.90)
7	.10 (.00–.26)	.08 (.00–.26)	.14 (.01–.39)	.38 (.15–.66)	.57 (.34–.80)
8	.04 (.00–.18)	.06 (.00–.19)	.05 (.00–.18)	.18 (.03–.43)	.39 (.13–.73)
	$\omega_1 = .18$ (.11–.27)	$\omega_2 = .23$ (.10–.38)	$\omega_3 = .29$ (.09–.49)	$\omega_4 = .25$ (.06–.44)	$\omega_5 = .03$ (.00–.11)

demand: the greater the number of partial intersections present in the compound—which increases exponentially with the number of figures—the smaller the probability of a correct answer. As can be seen in Table 10, TSP 4* is not visible. Looking at the probabilities in the lower left-hand triangle, it is clear that these do not agree with the constraints implied by TSP 4*.

Furthermore, these results confirm the all-or-nothing character of children's performance on the FIT. That is, no matter the distance between a child's M-capacity and a test item's M-demand, as long as the former is greater than the latter, a child will tend to succeed; otherwise, he or she will tend to fail and again the likelihood of doing so is unaffected by the distance between M-capacity and M-demand. As can be seen in Table 9, the double monotone, row ordering and column ordering models (which are more subtle than the all-or-none implications of TSP 123) can predict the observed frequencies of the response patterns, however, compared to TSP 123 and TSP 1234, these models have very small posterior probabilities. This is supported by the numbers in Table 10, where it is immediately clear that both column-wise and row-wise ordering of the class-specific probabilities are rather unlikely.

Finally, as can be seen in Table 10, the majority (23% + 29% + 25% + 3%) of 58 to 61-month-old children can simultaneously co-ordinate at least two schemes in a goal-directed activity. Only 18% of the children do not have this capacity.

CONCLUSION

This paper proposed confirmatory latent class analysis as an alternative for exploratory latent class analysis. Confirmatory latent class analysis is especially appropriate if competing theories can be derived from an existing body of theories, research and knowledge. As exemplified using two data sets and corresponding theories from the domain of developmental psychology, inequality constraints can be used to superimpose theories on to a set of competing latent-class models. Subsequently, fit measures can be used to select the best model and, thus, the best theory. Finally, estimates of the model parameters, obtained properly accounting for the constraints, can be used to determine the proportion of persons in each class and the value of the class-specific probabilities.

Currently, user-friendly software containing an implementation of the proposed approach is being developed. Readers interested in this software can send an email to the first author. The email should include a description of the research project at hand, the data and the theories involved.

REFERENCES

Boom, J., Hoijtink, H., & Kunnen, S. (2001). Rules in the balance: Classes, strategies or rules for the balance scale task. *Cognitive Development*, *16*, 717–735.

Clogg, C. C. (1981). New developments in latent structure analysis. In D. J. Jackson & E. F. Borgatta (Eds.), *Factor analysis and measurement in sociological research* (pp. 215–246). Beverly Hills, CA: Sage.

Everitt, B. S. (1988). A Monte Carlo investigation of the likelihood ratio test for number of classes in latent class analysis. *Multivariate Behavioral Research*, *23*, 531–538.

Formann, A. K. (1985). Constrained latent class models: Theory and applications. *British Journal of Mathematical Statistical Psychology*, *38*, 38–111.

Goodman, L. A. (1974). The analysis of systems of qualitative variables when some of the variables are unobservable. Part I: A modified latent structure approach. *American Journal of Sociology*, *79*, 1179–1295.

Haberman, S. J. (1988). *Analysis of quantitative data: Vol. 2. New developments*. New York: Academic Press.

Hoben, T., & Hettmansperger, T. P. (2001). Modelling change in cognitive understanding with finite mixtures. *Applied Statistics*, *4*, 435–448.

Hoijtink, H. (1998). Constrained latent class analysis using the Gibbs sampler and posterior predictive *p*-values: Applications to educational testing. *Statistica Sinica*, *8*, 691–711.

Hoijtink, H. (2001). Confirmatory latent class analysis: Model selection using Bayes factors and (pseudo) likelihood ratio statistics. *Multivariate Behavioral Research*, *16*, 717–735.

Hoijtink, H., & Molenaar, I. W. (1997). A multidimensional item response model: Constrained latent class analysis using the Gibbs sampler and posterior predictive checks. *Psychometrika*, *62*, 171–180.

Jansen, B. R. J., & van der Maas, H. L. J. (1997). Statistical test of rule assessment methodology by latent class analysis. *Developmental Review*, *17*(3), 321–357.

Jette, M., Desrosiers, H., Tremblay, R. E., & Thibault, J. (2000). *Longitudinal study of child development in Québec (LDEQ)*, *1*(2). Québec: Institut de la Statistique du Québec.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

Lin, T. H., & Dayton, C. M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics*, *22*, 249–264.

Pascual-Leone, J., & Baillargeon, R. (1994). Developmental measurement of metal attention. *International Journal of Behavioral Development*, *17*, 161–200.

Vermunt, J. K. (1996). *Log-linear event history analysis. A general approach with missing data, latent variables, and unobserved heterogeneity*. Tilburg: Tilburg University Press.

von Eye, A., & Clogg, C. C. (1994). *Latent variable analysis. applications for developmental research*. London: Sage.

Zoccolillo, M. (2000). Parent's health and social adjustment. Part II, Social adjustment. *Longitudinal Study of Child Development in Québec (LDEQ)*, *1*(2), 9. Québec: Institut de la Statistique du Québec.

Zoccolillo, M., Pickles, A., Quinton, D., & Rutter, M. (1992). The outcome of conduct disorder: Implications for defining conduct disorder and adult personality disorder. *Psychological Medicine*, *22*, 971–986.

Zoccolillo, M., Price, R. K., Ji, T., & Hwu, H.-G. (1999). Antisocial personality disorder: Comparisons of prevalence, symptoms, and correlates in four countries. In P. Cohen, C. Slomkowski, & L. L. L. Robins (Eds.), *Historical and geographic effects on psychopathology*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.