# Latent Transition Analysis: Video 3 - Structural Models

## Transcript

Video: <https://youtu.be/x-wd4RlTipY>

Oliver Perra: Hello, I'm Oliver Perra and this is the third presentation. In the previous presentation, I have illustrated the first stages of latent transition analysis. So we start by investigating the number of classes that optimally explain heterogeneity of behaviour, and we do so separately at each timepoint. If we identify latent classes that appear to be similar across timepoints, we can also test measurement invariance of these classes across time. Once we have selected one or a few optimal measurement models, we can extract the measurement parameters from these models and use the three-step approach to investigate the associations between latent classes across time and also the association between those classes and other variables. In this presentation, we'll talk about these in more details.

 So I will follow the outline of stages of latent transition analysis that I have described in a chapter linked here. And in particular in this presentation, we'll talk about these issues: investigating transitional probabilities; models where I include covariates and distal outcomes; and also models that include moderation effects. And I will mention associative latent transition models. So I will start by talking about investigation of transition probabilities in the models. And, as I illustrated in the previous presentation, once we have chosen optimal measurement models at each timepoint, we can use the latent class memberships estimated to which participants are assigned as imperfect indicators of latent classes, feed into the model the information about the extent of uncertainty in this assignment in the membership of participants into these classes. This information is represented by the classification (inaudible 0:02:17) parameters that I have extracted from the optimal measurement models. In the previous presentation, I had shown how we can calculate these (inaudible 0:02:30), but software like Mplus and others will provide them for use without the need to calculate them. These (inaudible 0:02:38) parameters that represent uncertainty in classification are indicated in this picture here as the parameters besides the arrows in the figure.

 Having defined a fixed measurement model, I can now impose structural relationships between data and classes at each timepoint. Now, the latent classes are considered nominal variables even when we might think of them as being ordered. So the association between latent classes are expressed by multinomial logistic regressions and these are basically regressions, multinomial logistic regression of the latent classes at age 15 on the latent classes at age 14. Since we are using logistic regressions, the associations between latent classes can be expressed by odds ratios. For example, we will have the odds ratios of being an abuser rather than an abstainer at 15 years, if someone was an experimenter rather than an abstainer at 14 years, and so on.

 As well as the odds ratios, the other parameters we obtain from a latent transition analysis output are the transition probabilities. These represent the probabilities of being in a latent class at 15 years conditional on latent class membership at 14 years of age, and these probabilities sum up to one in the rows, as you can see here. So if we have the same or similar classes at different timepoints, we can report what are the probabilities of individuals staying in the same class across time. For example, here we can see that abstainers at age 14 have a 68% probability of remaining abstainers at 15 years. So these probabilities represent continuity in behavioural patterns across time. By considering the number of participants in the study, we can also calculate the total probability of individuals remaining in the same classes over time. This represents continuity as long as the latent classes are the same across time.

 And at this stage, I just wanted to give a summary of the key parameters of latent transition analysis. And here the letter P represents probability and the vertical bar sign means conditional on. So the first two parameters we obtain are those produced by latent class models. So conditional item response probabilities at each timepoint representing the probability of responding to an item in a certain way for people in a certain latent class. And we also obtain the prevalence of latent classes at each timepoint. The first set of parameters are unique to latent transition analysis and they represent the structural model. These are the probabilities of being one latent class at 1.1 age given membership to latent class at the previous timepoint or the previous age. So, for example, the probability of being an abuser at 15 years of age conditional on being an experimenter at 14 years of age.

[0:06:20]

 I will show that it is possible to constrain some of these transition probabilities to test specific hypotheses about developmental patterns. For example, here we see that the probability of a substance use experimenter becoming an abstainer the following year is quite low. We can hypothesise that this reported 5% probability of moving from experimenter to abstainer is due to measurement error and we can test if the model will provide the same fit when we constrain this probability to be zero, that is when we assume that the transitions from experimenter to abstainer classes have just an effect of random variation. So we can estimate a model where this transition probability is constrained to be zero. And since this model is nested within the model where the transition probabilities are freely estimated, we can formally test our hypothesis by calculating the likelihood ratio test between the two models. In the exercises I have attached with these presentations, I provide more guidance on how to conduct this test.

 And in a similar manner, we can formally test other hypotheses on transitions across age. Here I am presenting a fictional example of a stage-like model of key skill mastery. Individuals may be at different levels where we have novices, learners and masters at a particular skillset. Our model may state that we should not observe individuals backsliding, that is moving from being a master to a novice level. Similar transitions will be only observed because of random variation and we can test this hypothesis by comparing a model with transitions freely estimated and a model where we constrain these backsliding probabilities to be equal to zero, like I indicate here. A likelihood ratio test comparing the two models provides a formal test of this hypothesis.

 Having talked about transition probabilities here, I want to emphasise that date and transition analysis is an autoregressive model. So individuals' status at one timepoint is directly related to the individual's status at the previous timepoint. So this also means that, unlike growth models where changes are modelled as a function of time, in latent transition analysis, time is not usually considered as a variable in the model. In this example of three measurement occasions and latent classes measured at time zero, time one and time two, the time variable does not really matter, so time zero may be age 10 years, time one may be age 11 years and time two may be age 18 years. So the ages at which we collect data are not usually part of the model. And while we usually consider that individuals' status at one timepoint is directly related to the status at the previous timepoint, that is first order effects, as described in this model, we can also test hypotheses where we model second order effects. For example, individuals' status at time two may be also related to individuals' status at time zero or else there may be lagged effects of the initial status. And we can test these models using likelihood ratio test, as I mentioned before.

 Now I'm going to talk about inclusion of covariates and distal outcomes in the latent transition models. When we use the three-step approach I described in the second presentation and, therefore, we have a fixed measurement model that takes into account uncertainty in class membership, it becomes straightforward to include covariates in the models. And that means we can investigate associations between the latent class statuses at different timepoints and covariates. Here, for example, consider the effects of gender on latent class status at age 14 and latent class membership at age 15. Since latent classes are considered nominal variables, again, we use multinomial logistic regressions where the data and classes are regressed on the covariates. And here I considered a nominal covariate like gender, but we can include all type of covariates - ordinal, continuous or latent variables. So the three-step approach also allows a lot of flexibility in inclusion of different types of covariates in the latent transition models.

 In a similar model described by this picture, we will obtain these parameters. First, we will have the probability of being in one class conditional on gender and the parameters at 15 years of age. However, report the probability of being one latent class conditional on gender and latent class in the previous age point age 14 years. In other words, this parameter represents latent class probabilities adjusted based on the previous class membership. So here I simplified example, including only two classes at each timepoint. And, again, for the sake of example, consider gender as a dichotomous variable where we have females and males. Based on the parameters of the model, we can create probabilities matrices like these, which represent the probabilities of staying in the same class or transitioning to a different one for females and males, respectively.

[0:12:51]

 And we can see, for example, that males have a higher probability, 0.38, of becoming users between timepoint zero and timepoint one compared to females who only have 28% probability of becoming users. And we can use these probabilities, as I illustrate in this slide, to calculate the odds of becoming a user for males and females and use this to calculate the odds ratios of becoming a user. Compared to females, males display a 56% increase in the odds of becoming a user between age 14 and age 15. So the odds ratios indicate that males have a higher likelihood of becoming a user. I will illustrate here that, however, the sign and the strength of the association between latent status at the two timepoints is the same across gender.

 For example, in this graph I report the probabilities of being a user across the gender conditional on previous latent class membership. So the two lines represent males and females and the two values on the X axis represent previous latent class status users at time zero and abstainers at time zero. And you can see that white males have higher probabilities of being users. The lines here are parallel, indicating that the strength of the association between latent variables was similar for males and females. However, we may be interested in testing more complex models where covariates may moderate the associations between the latent classes at different timepoints, and I represented this moderation by gender in this graph here.

 To illustrate what might happen with a moderation, I looked here, the probability matrices for females and males when we model the moderation effect of gender. And in this case, we can see that males have a higher probability of becoming users compared to females. But males also reported lower probabilities of remaining users compared to females. Overall, this moderation effect indicates that the association between latent classes across time is stronger for females and, consequently, males have a higher likelihood to change substance use status between age 14 and age 15. So we can use latent transition analysis to look at these types of moderation effects by gender or by other covariates.

 And, finally, using the three-step approach, we can also include distal outcomes in the models where we have fixed latent class measurement models while also taking into account uncertainty in measurement models. We can, for example, investigate the association between substance use categories at 15 years of age and school attainment at 16 years while controlling for covariates like gender. And we will do this by comparing the adjusted average school attainment across individuals in the three latent classes at age 15. The three-step approach also makes it possible to develop more complex models like this one. This is an example of an associative latent transition analysis where we investigate changes into two different processes, in this example substance use and depression. So we can identify measurement models for those different processes and investigate how people and adolescents move across different categories of substance use between age 14 and age 15 and how they move across different categories of depression from age 14 to age 15.

 And we may then investigate questions that concern how, for example, the depression status may affect substance use across time. For example, we may assume a cross-lagged effect of depression from age 14 to substance use status in age 15. And by adding this regression and controlling for substance use class at early ages, we are effectively testing if depression status at 14 years affects changes in substance use classes at age 15. We can test the cross-lagged model with a model without this effect, and since these models are nested, we can use the likelihood ratio test to formally test if there is a difference in the fit of the two models. So, similarly, we can develop more complex models when we may have more points, more covariates of different nature and distal outcomes.

[0:18:17]

 So to summarise, in this third presentation I have talked about ways in which we can use latent transition analysis to test hypotheses about transitions and orders of development where, for example, we can use latent transition analysis to test whether the probabilities of backsliding or individuals moving from a more advanced stage of development to a less advanced stage of development are only observed because of random variation in the data. We can also include covariates in latent transition analysis, and the three-step approach I discussed makes it very straightforward. So we can use latent transition analysis and the three-step approach to test different hypotheses about the effects of covariates, including more complex moderation effects of covariates, so testing whether some covariates may moderate the association between data and statuses at different timepoints.

 And we can also easily include distal outcomes to test whether membership in some latent classes at some timepoint affect other outcomes later on. And, finally, we can use these methods to develop more complex models where we can also look at associative latent transition analysis. And this you can add how person-centred changes across one process, for example changes in mental health statuses from one age to another, may affect changes in another process, for example changes in substance use patterns from one change to another. So I have provided more material and exercises together with these presentations, so I hope this is useful and thank you very much for your attention. Bye.

National Centre for Research Methods (NCRM)
Social Sciences
Murray Building (Bldg 58)
University of Southampton
Southampton SO17 1BJ
United Kingdom

**Web** www.ncrm.ac.uk
**Email** info@ncrm.ac.uk
**Tel** +44 23 8059 4539
**Twitter** @NCRMUK