This is the first of three sessions introducing the basics of multilevel modelling. In this session I'll introduce multilevel modelling by providing a general overview of what it is, and what it's for, I'll then discuss the types of data structure that multilevel modelling can be used to capture, starting with a simple two-level nested data structure, and I'll finish by considering some of the benefits of adopting a multilevel approach. In the next session I'll provide a more detailed discussion of the most commonly used multilevel model, the two-level random intercept and random coefficient models.

Multilevel models are a series of models that extend basic regression approaches to better account for the complex forms of grouping structures that are evident throughout society. For example, when thinking about my own research in criminology, I often want to recognize that individual experiences of crime do not occur in a vacuum. Rather they may also be shaped by the social context in which they're experience. Multilevel models give us a way of incorporating this grouping structure, alongside individual differences, as well as a way to explore the link between the grouping structure and in the individual differences. In other words, how does being a member of a particular group influence the experiences of the individuals in it. They also provide us with a way to statistically adjust our estimates for the fact that observations from the same group, may tend to be more similar to one another, on average, than they are to observations from different groups. We can also think of simple multilevel models as a way to re-partition variation in our outcome of interests between groups and individuals.

Perhaps the simplest structure we can deal with is a two-level nested structure. Here for example we've data from respondents showing just the first nine. In multilevel models we typically refer to data measured on individual units as level one. We also know that these respondents are resident in different areas. Here we refer to the next level, where we have information, as level two. In addition, we may have reason to expect that residents of the same area, may have more similar outcomes on average, than with residents of different areas. Perhaps reflecting shared social experiences, visible proximity or neighbourhood networks. We may want to examine differences between the units at level 1, but also consider the possibility of differences across level 2, and how belonging to a particular level 2 unit, may shape the responses at level 1. Take fear of crime for example. Here we can see that scores for fear of crime tend to be higher in area 1, lower in area 2 and more varied in area 3. We may want to know systemically whether there are differences in fear of crime between all neighbourhoods, in addition to any observed differences between residents, and exactly how much of the variation is to do with groups. And we may also want to understand whether individual differences in fear, for example between those who have been a victim of crime and those who haven't, are shaped by features of the area. Level 1 doesn't have to be individuals, rather it's the lowest level that we have data. For example, here we have a two-level model where individuals at the group level, and we have repeated measurements of the same test for each respondent. So, we could explore how improvement differs across students and why this is. Here note that we can include group level data here gender by repeating it across each measurement occasion. Importantly here we only have observations at t1 and t2 for person 2, and t1 and t3 for person 3. The multilevel case is robust to uneven group sizes, which in this context relates to those missing observations.

Although two-level models are still most common, it's then straightforward to generalize to many more levels, with some software no longer restricting the number of levels at all. Of course, the size and complexity of the data structure will be limited by the data availability, and sample size at all levels will come into play with more complex structures. They also become increasingly difficult to make sense of, so it's always advisable to start with much simpler two-level structures first.

Multilevel models have also been developed for a range of more complex non-nested data structures. For example, we might have a non-nested or cross classified structure. Here pupils are grouped hierarchically into neighbourhoods, or schools, but not all pupils who live in the same neighbourhood also attend the same school; and similarly, not all pupils from the same school, also live in the same neighbourhood. We may also have a multiple membership structure where individuals belong to more than one group. Or a more explicitly spatial structure, where we don't just want to account for similarities between people from the same area, but we also want to account for the similarities that exist between areas that are closer to one another, or combinations of all three structures.

So, there are several reasons why we might want to adopt a multilevel approach. Firstly, and as noted, this gives us a comprehensive framework to correctly account for complex data structures. Ignoring this structure when it exists, for example, with a single level regression, can result in underestimated standard errors and a substantive blind spot. It also enables us to incorporate group level information simultaneously with individual information, rather than separate individual and aggregate analyses; and we can then link the context back to the individual to explore more carefully how individual relationships can be moderated by the broader context. Finally, and in a more advanced framework, it also gives us a way to model heterogeneity. This means we can move beyond the typical focus on average relationships, to also explore how relationships vary, and why.