The unexplained variance at level 2 can increase when you add in an explanatory variable- it doesn't mean something's wrong with your model. I find this easiest to understand using Kelvyn Jones' house price example:

Imagine you have a dataset which has information on house size and house price for 4 areas in a city. (Obviously you wouldn't usually fit a multilevel model when you only have 4 level 2 units, but the logic is the same and it's easier to get your head round if you imagine just 4 areas, I think).

You start by fitting a variance components model with level 1 = house, level 2 = area, and just a fixed intercept, a level 2 random intercept, and a level 1 residual term; the response is price

$$y_{ij} = \beta_0 + u_j + e_{ij}$$

This model does not take into account the information you have about the house sizes. Actually, within each area there is a positive relationship between house size and house price: larger houses are more expensive. But when you look at the areas, you find that the more expensive areas, which are in the city centre, have smaller houses (because there is not room for large houses in the centre and who could afford them if there was?), while the cheaper areas in the suburbs have larger houses. So the average price of a house in the most expensive area, not adjusting for house size but just averaging the observed prices, will be more expensive than the average, but it will not be as expensive as it would be if that area had average size houses: the small houses in that area bring its average price down. Similarly, the average price of a house in the cheapest area, while still cheaper than the average, is pushed up because the houses in that area are larger than the average house size.

Now we fit a model that includes house size:

$$y_{ij} = \beta_0 + \beta_1 \, size_{ij} + u_j + e_{ij}$$

Because this model includes information on house sizes, the predicted value from the fixed part for the most expensive area, which is based on house size, will be cheaper than the average, since the most expensive area has smaller houses than average. So the level 2 random intercept for this area will be larger than in the previous model, where the predicted value was the overall mean price for the dataset. Similarly, the predicted value for the cheapest area will be more expensive than average, based on its larger than average houses, and the random intercept for this area will also be larger (in magnitude: more negative) than in the previous model. The same will be true generally of all the areas, and this gives rise to the larger level 2 variance compared to the previous model.

****

Thanks,

Rebecca Pillinger