Belgian-Dutch Network for
ESM Research in Mental Health

Heerlen Meetup 3 & 4 October 2024

Hackathon submission form	
Title	Fitting generalized linear mixed-effects models with SPSS: Cause for concern?
Contact person *(NB contact person is not necessarily same person as the organizer)*	*Name* : Wolfgang Viechtbauer *Affiliation* : □ UMCG/RUG Groningen □ **Maastricht/Heerlen/GGzE** □ Leuven □ Tilburg □ Rotterdam □ Other: *Email address:*
Organizer(s) *(=the one who will prepare and facilitate the session)*	*Name* : Wolfgang Viechtbauer *Affiliation* : □ UMCG/RUG Groningen □ Maastricht/Heerlen □ Leuven □ Tilburg □ Rotterdam □ Other: *Email address :* wolfgang.viechtbauer@maastrichtuniversity.nl

Abstract (max 200 words)	Generalized linear mixed-effects models (GLMMs) are a general class of models that are used in the analysis of ESM/EMA data when the response variable of interest cannot be assumed to have a normally distributed error term. The most prominent special case of GLMMs are logistic mixed-effects models, which are typically used when the response variable is dichotomous. Fitting such models is considerably more complex compared to linear mixed-effects models, as the equation for the likelihood involves integrals. Algorithms for fitting such models either involve approximations that avoid solving these integrals (e.g., the Laplace approximation, penalized quasi-likelihood estimation) or use numerical methods (e.g., adaptive Gaussian quadrature, MCMC) to evaluate them. On top of this, numerical optimization methods must be used to maximize the (log) likelihood. Therefore, software packages for fitting such models can differ in the results they provide and their convergence behavior (e.g., Austin, 2010; Bolker at al. 2009; Guo & Zhao, 2000; Kim et al., 2013; Li et al., 2011; Rodriguez & Goldman, 1995; Schoeneberger, 2016; Zhou et al. 1999). An informal comparison among some commonly used software packages (i.e., R package lme4, Stata command melogit, SPSS procedure GENLINMIXED, SPSS procedure PROC GLIMMIX) suggests that especially SPSS gives more noticeably divergent results compared to the other options, which may raise concerns about the results provided by SPSS. In this hackathon, we will discuss and examine this issue in more detail. To study this issue more systematically, we may consider running a simulation study to compare the results across different software packages. Possible end product: this could be written up as a proper article
Relevance for attendees	Although a bit more statistical and technical in nature, this might be interesting and challenging for some of the attendees. I wouldn't mind if the hackathon group is a bit smaller, as the output of this hackathon could easily be written up as a proper article and such things are simpler when the group is not too large.
Other comments	
The number of participants per session will be 25-30 participants. The hackathons will take 60 min on Oct 3th and (max) 120 min on Oct 4th. In addition, a summary of the session will be presented on the 4th (5 min).	

All end products will be shared with the network on Basecamp so they can be used as starting points for follow-up actions/collaborations.