Associazione Italiana di Psicologia | SEZIONE SPERIMENTALE

XXIX CONGRESSO NAZIONALE

Latent Growth Model in Psychology: What is the risk if the measurement is not (really) on an interval scale?

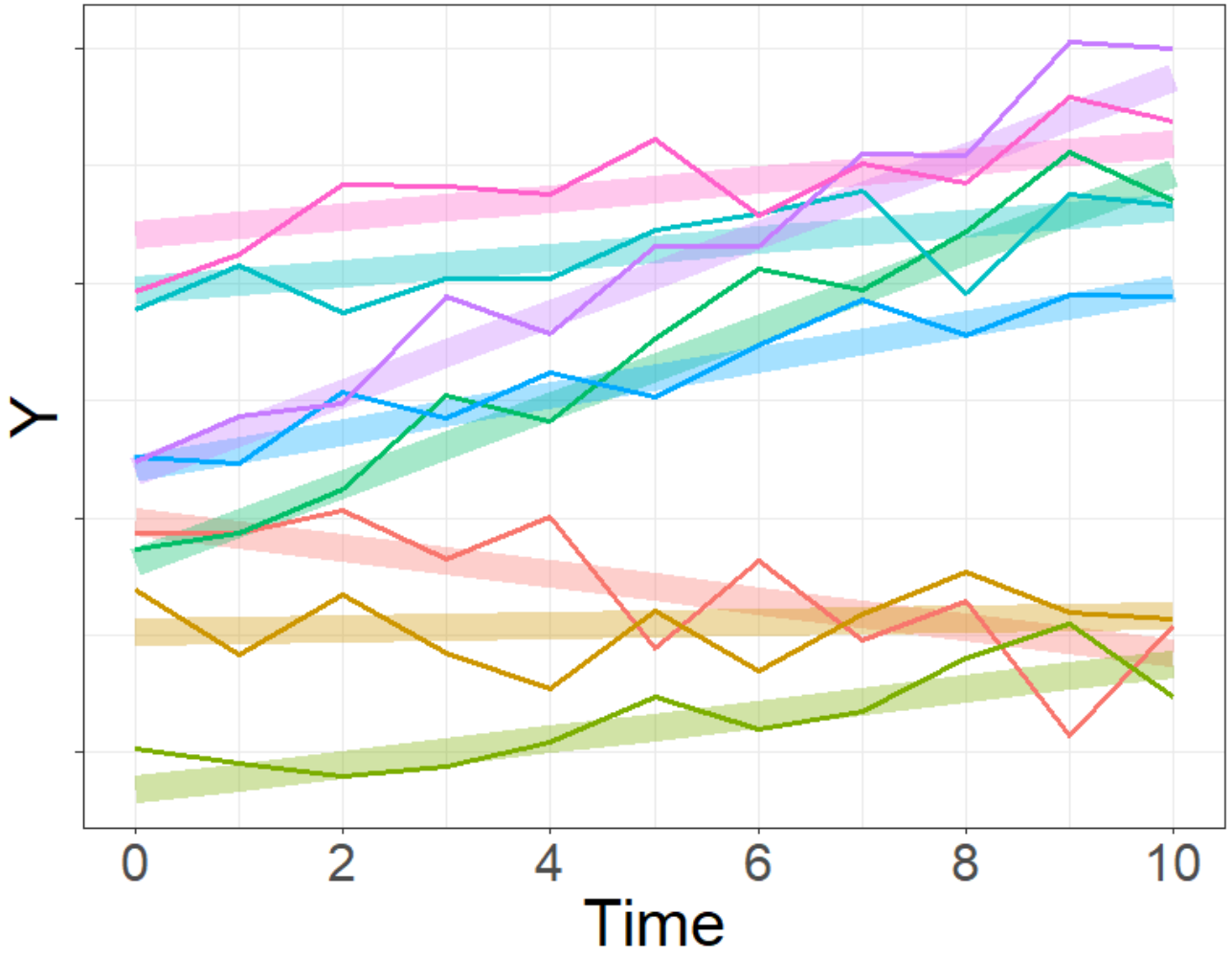
ENRICO TOFFALINI — DPG, Università di Padova
TOMMASO FERACO — DPG, Università di Padova
MASSIMILIANO PASTORE — DPSS, Università di Padova

Session: Methodology 2

LUCCA - Mercoledì, 20 Settembre 2023

Usefulness of latent growth models

Each path has a starting point (intercept) and a trajectory (slope). LGMs help us describe its parameters



Usefulness of latent growth models

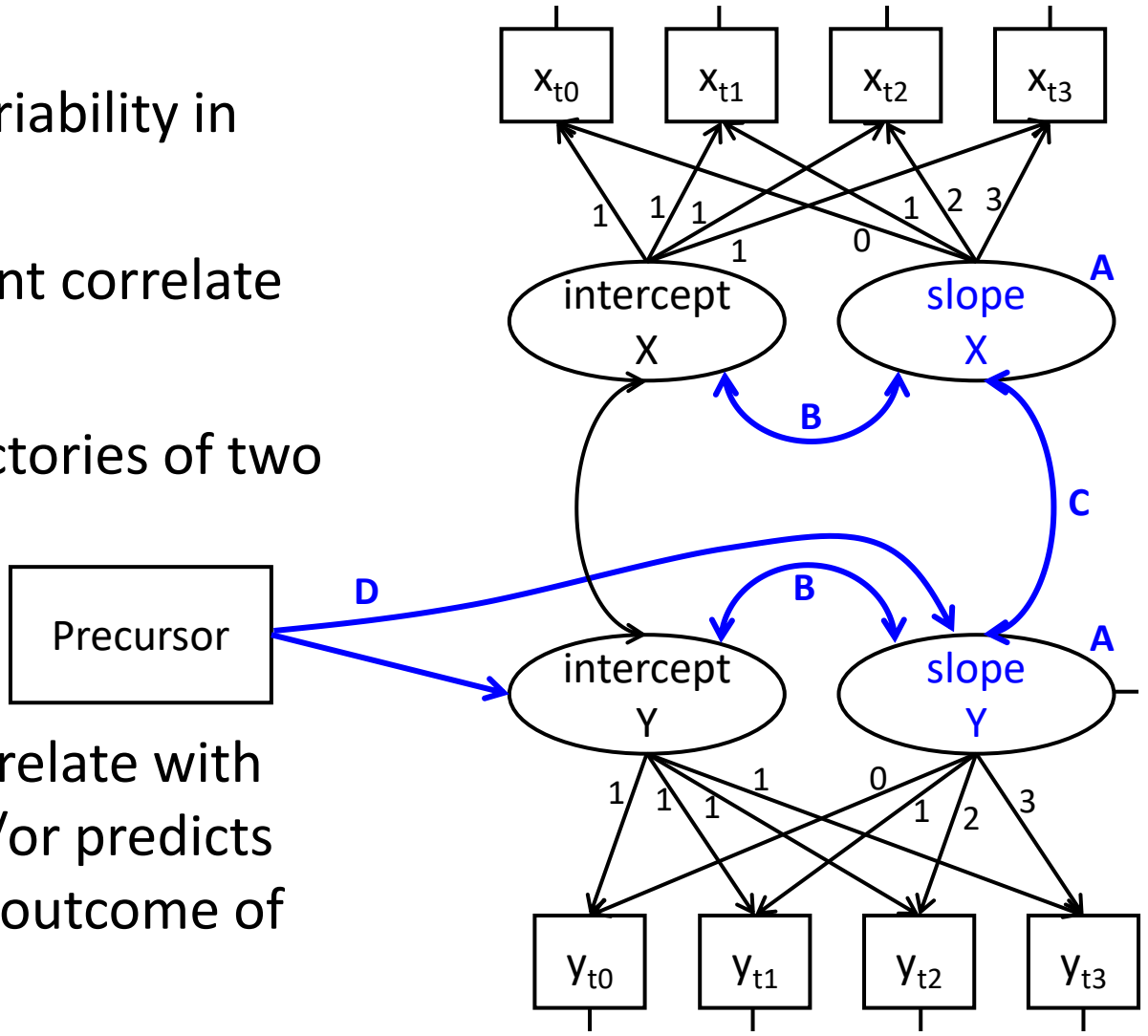
What might we ask?

A Is there individual variability in trajectories?

B Does the starting point correlate with the slope?

C Are the growth trajectories of two variables correlated?

D Does a precursor correlate with the starting point and/or predicts the future slope of an outcome of interest?



Usefulness of latent growth models

These phenomena are of scientific interest... and are sometimes given curious evangelical names...

"(…) Growth and stability can also be distinguished from spread (see Figure 1), which relates to whether the distribution of scores over time stays the same, decreases, indicating a **compensatory pattern**, or increases, such that the ***"rich get richer and the poor get poorer"*** (so called **Matthew effects**; Pfost, Hattie, Dörfler, & Artelt, 2014; Stanovich, 1986). **Matthew effects** have been reported for reading comprehension between 7 and 10 years (Quinn et al., 2015), and for vocabulary knowledge between childhood and 16 years (Duff et al., 2015). However, a compensatory pattern has also been reported for a reading composite between 6 and 12 years (Shaywitz et al., 1995). A recent metaanalysis shows inconsistency across studies but suggests that Matthew effects were more likely for sensitive and reliable reading measures (Pfost et al., 2014)"

Ricketts et al. (2020)

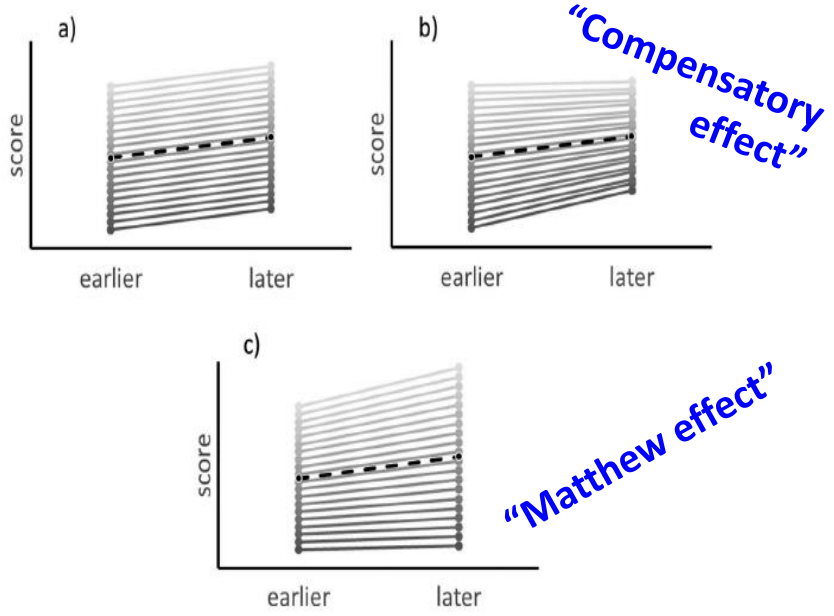


Figure 1. Hypothetical illustration of how spread can be consistent (a), decrease, indicating a compensatory pattern (b), or increase, indicating a Matthew effect (c).

Note. Growth, stability and spread are statistically independent from each other. In this example, there is there is consistent mean growth (black dashed line) and perfect stability (gray solid lines), with earlier rank in distribution identical to later rank.

SCIENTIFIC STUDIES OF READING
2020, VOL. 24, NO. 5, 380–396
https://doi.org/10.1080/10888438.2019.1689244

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Reading and Oral Vocabulary Development in Early Adolescence

Jessie Ricketts [a], Arne Lervåg [b], Nicola Dawson [a], Lucy A. Taylor [a], and Charles Hulme [c]

[a]Department of Psychology, Royal Holloway, University of London, London, UK; [b]Department of Education, University of Oslo, Oslo, Norway; [c]Department of Education, University of Oxford, Oxford, UK

Phenomena and explanations

Theoretical explanations? E.g. possible Matthew effect in reading skills (Pfos et al. 2014) as a virtuous circle: I read well → I enjoy it → I read more → I read better and better. Or: SES affecting both starting point and trajectory. Compensatory effect, however, sometimes interpreted as a temporary delay in development which is then recovered (Pfos et al., 2014)

But before offering sophisticated psychological explanations, simple psychometric explanations must be excluded

Trivial artifacts to guard against? If I only have 2 time points could I correlate X0 with (X1-X0)? NO. the correlation will be negative ("compensatory effect"?!) due to trivial regression towards the mean. To model the latent variables of the LGMs, at least 3 time points are needed

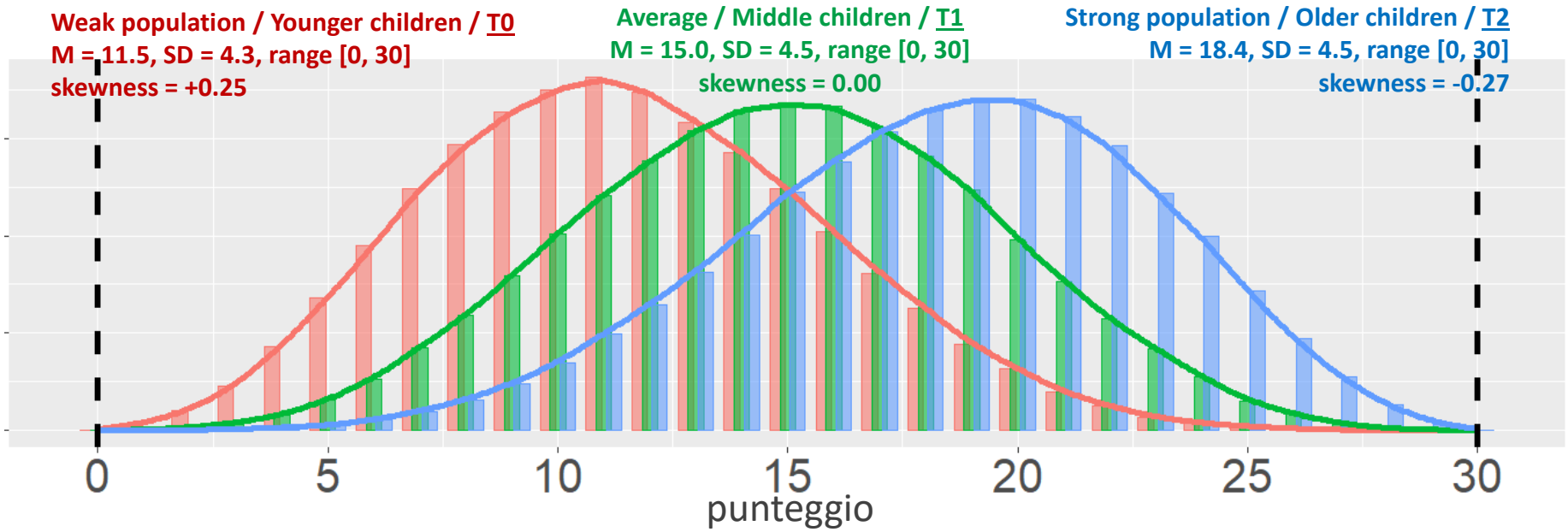
A less obvious problem, even when choosing the right model, concerns data distributions...

Data distributions

In psychology we often have bounded (min-max) scores that are not normal, but apparently not too skewed, e.g.

- Sum of Likert scale item responses to questionnaire
- Sum correct answers to tests (e.g. reading comprehension, mathematics, reasoning)

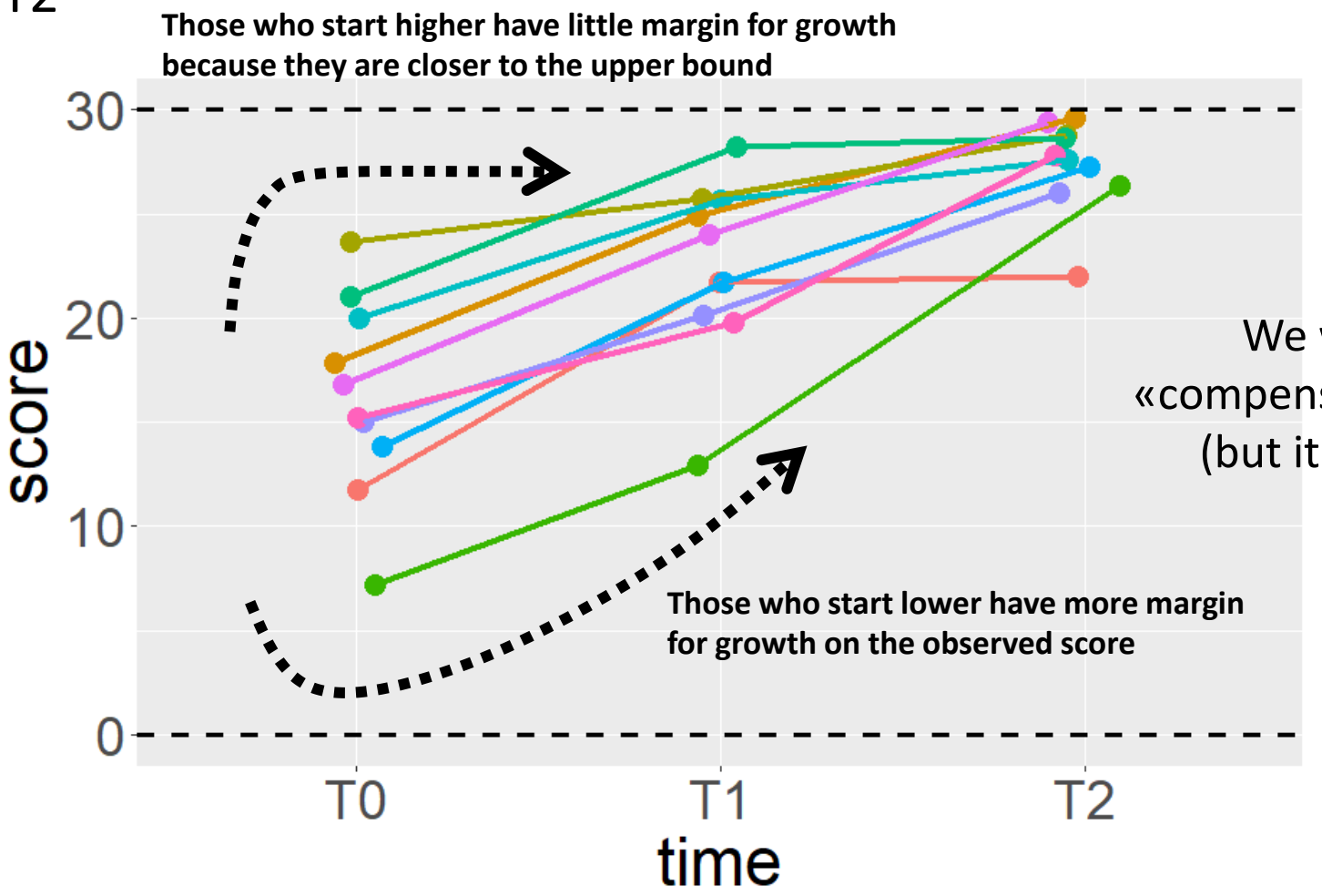
EXAMPLE – Data from a comprehension test with 30 response items…



Let's proceed with a similar case, **simulating data**...

Data distributions

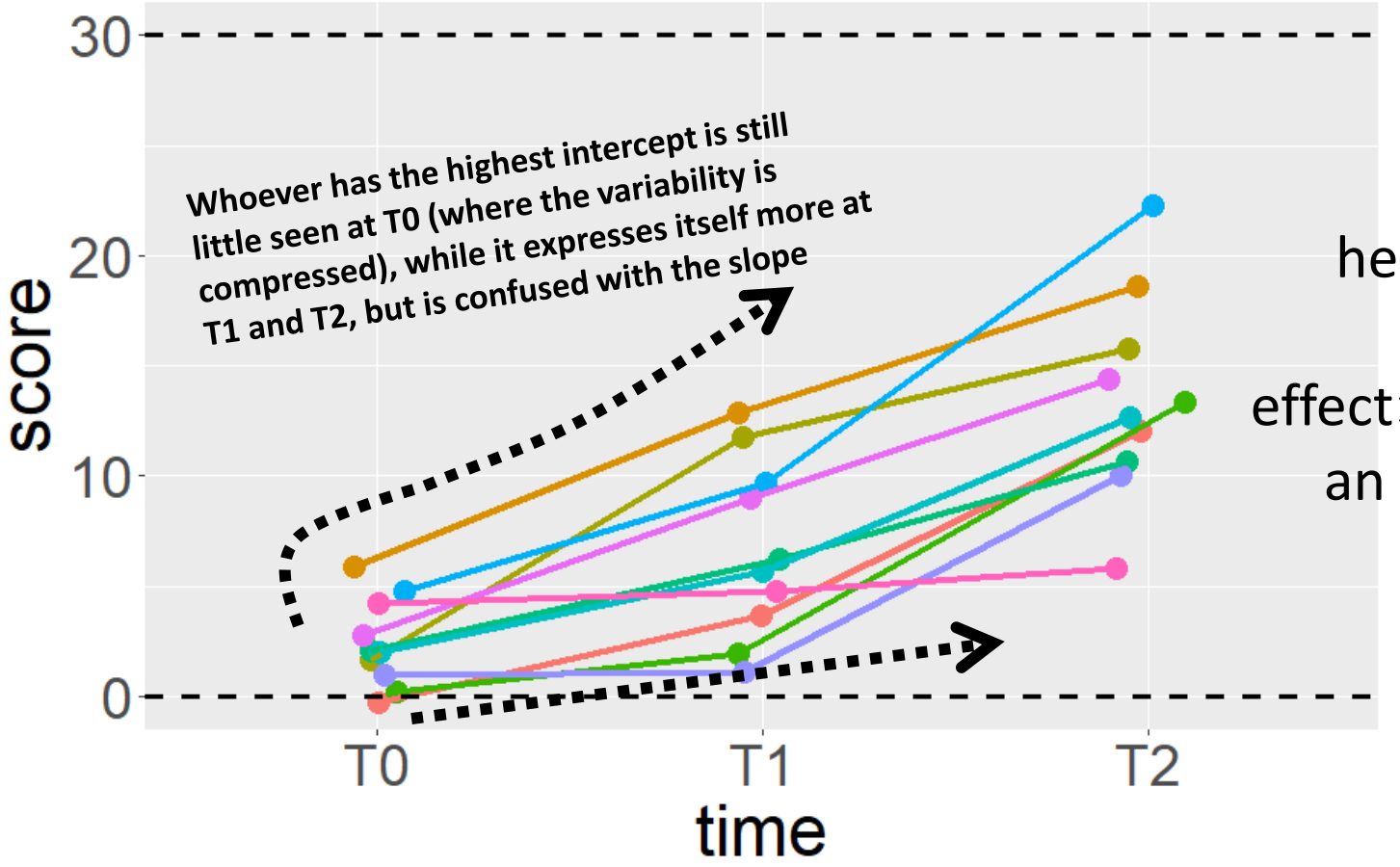
This is where troubles begin: let's take the case of a sample taken at an intermediate score at T0 and followed up at T1 and T2



We will observe a «compensatory effect» (but it is an artifact)

Distribuzioni dei dati

But the opposite can also happen if we take a sample with a low score at T0, then followed up at T1 and T2

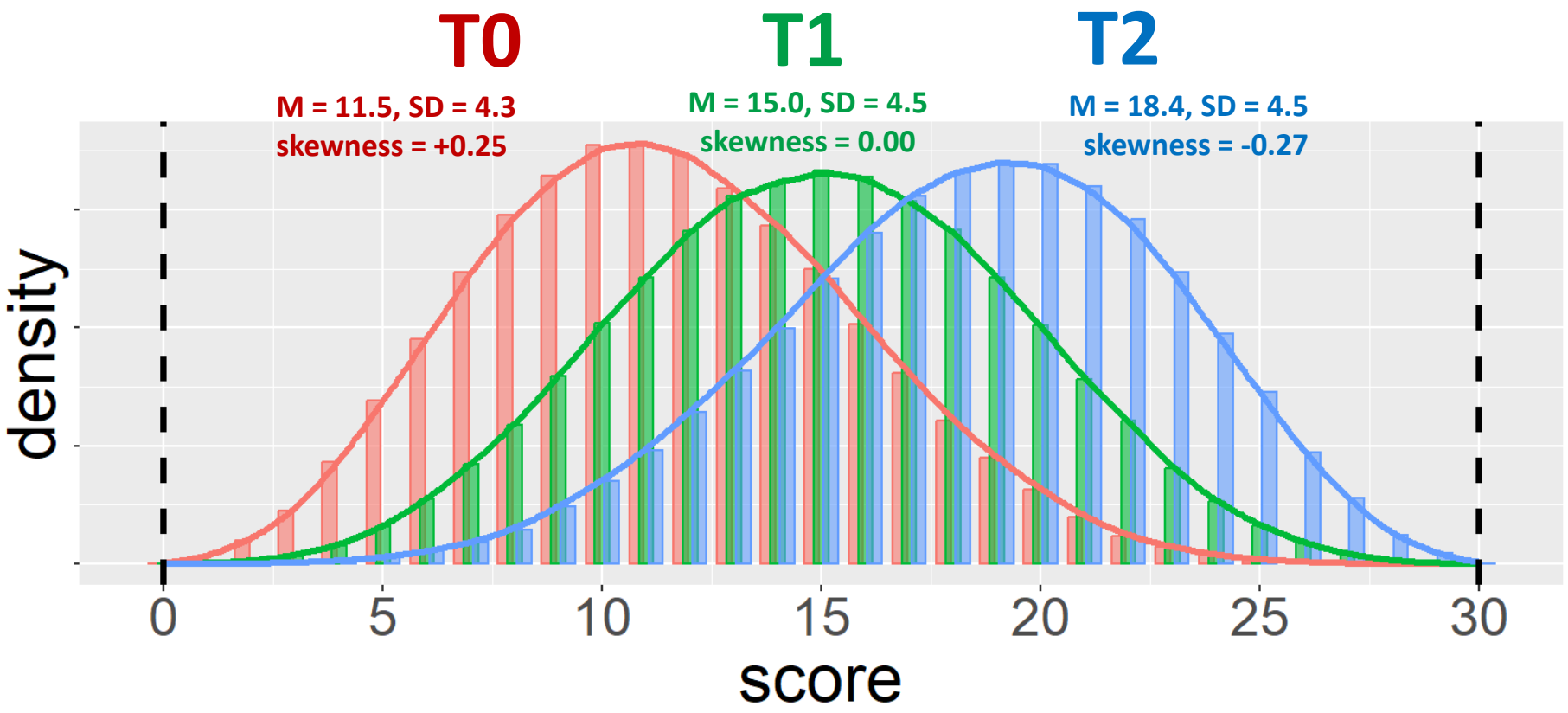


here it is, the «Matthew effect»! (but it is an artifact) ☺

Data distributions

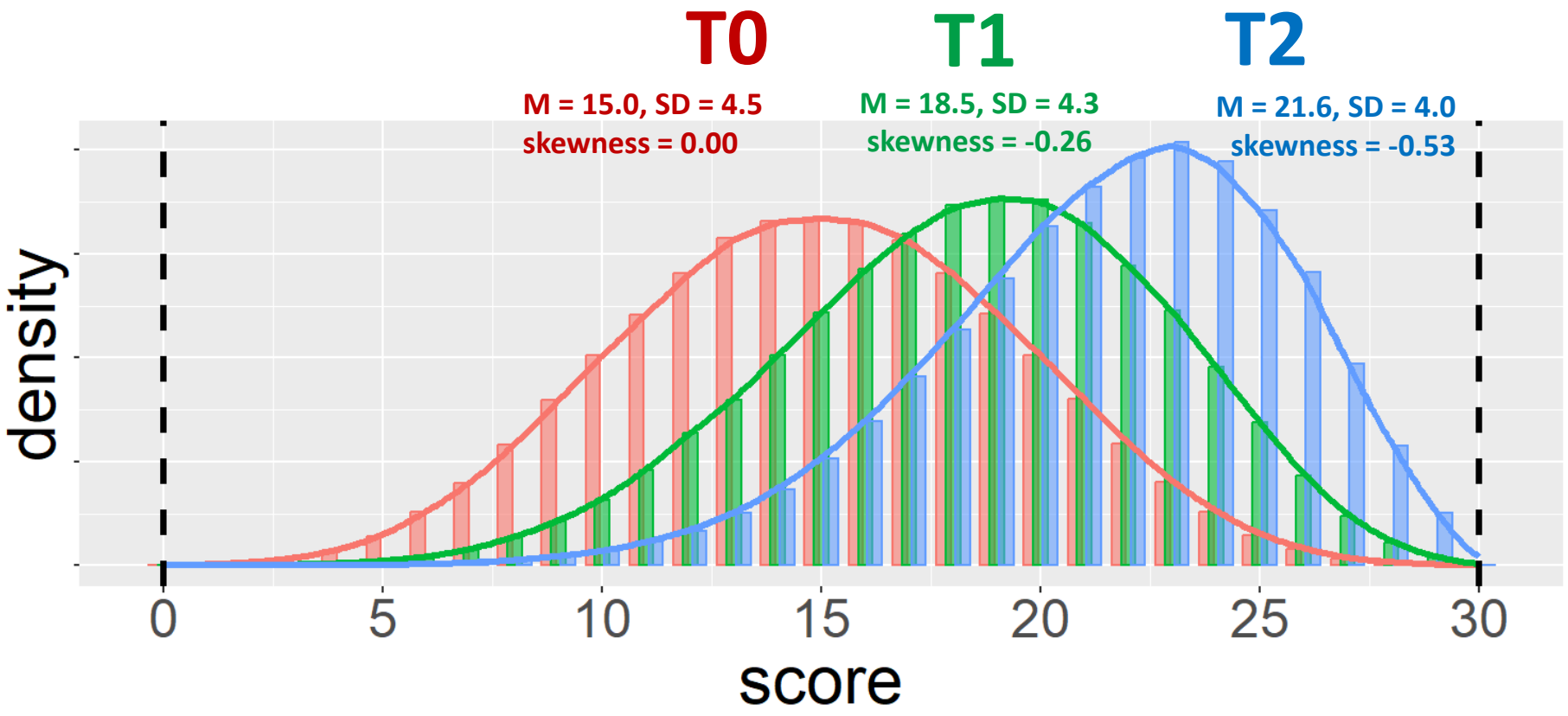
The previous examples are extreme (ceiling and floor effects evident), but what would happen with a case like the one seen before?

In the case below we can have an **intercept-slope correlation of -0.19**
(thus a «compensatory effect» which is *NOT* in the data generating process) ↓



Data distributions

In this other case, actually a little more "suspicious“, but not even too much for usual rules of thumbs, **the intercept-slope correlation actually becomes -0.46**

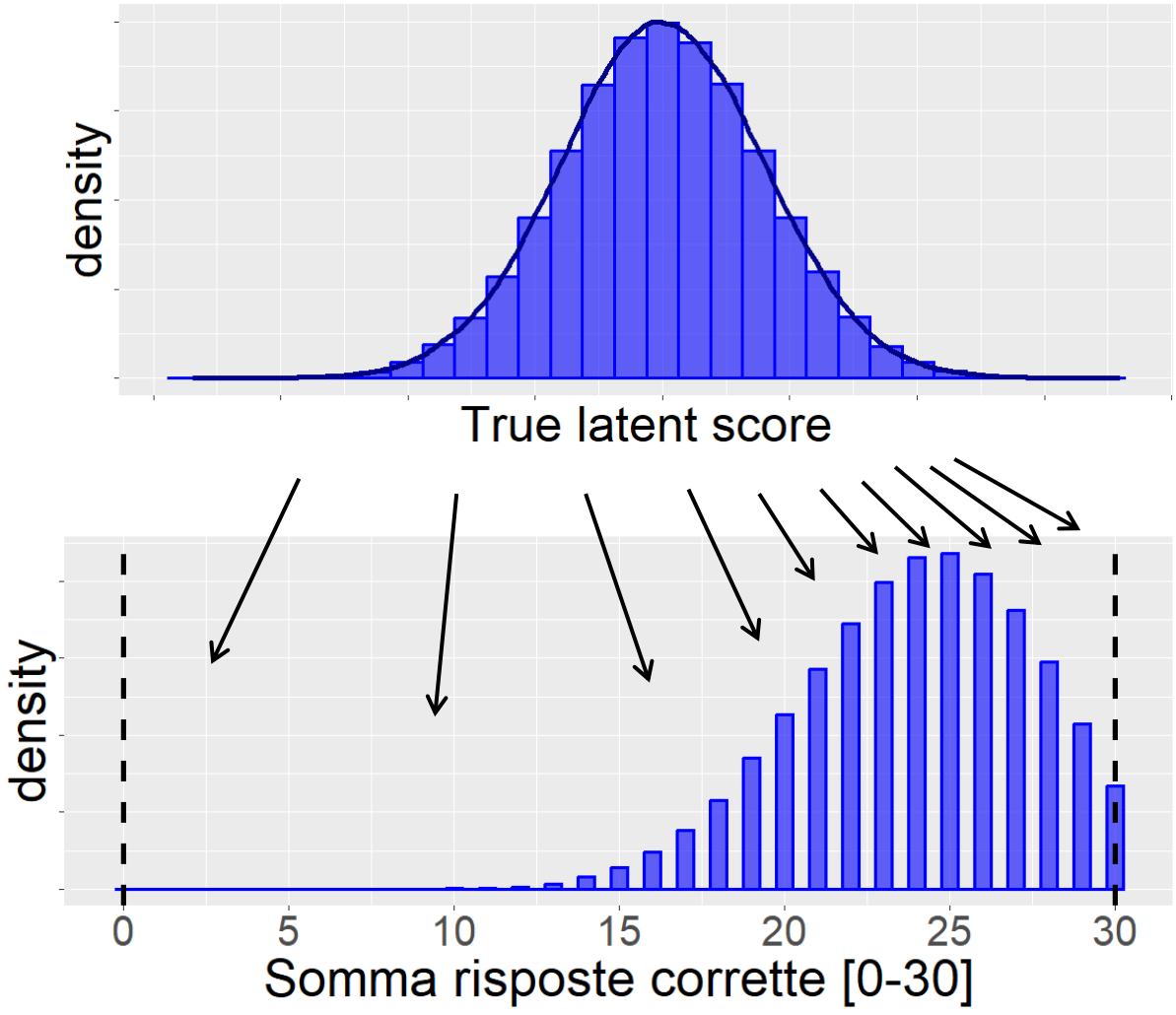


Data distributions

'TRUE' GENERATIVE PROCESS: the latent variables are hypothesized to exist on a linear continuum and normally distributed

GENERATIVE PROCESS OF MEASUREMENT: for practical reasons we are forced to measure them through responses on binomial (right-wrong) or ordinal scales (Likert scales) which, if not treated as such, cannot cover and reflect the entire latent continuum

<u>THE PROBLEM IS THAT EQUAL INTERVALS ON THE OBSERVED SCORE DO NOT CORRESPOND TO EQUAL INTERVALS ON THE LATENT CONSTRUCT</u>



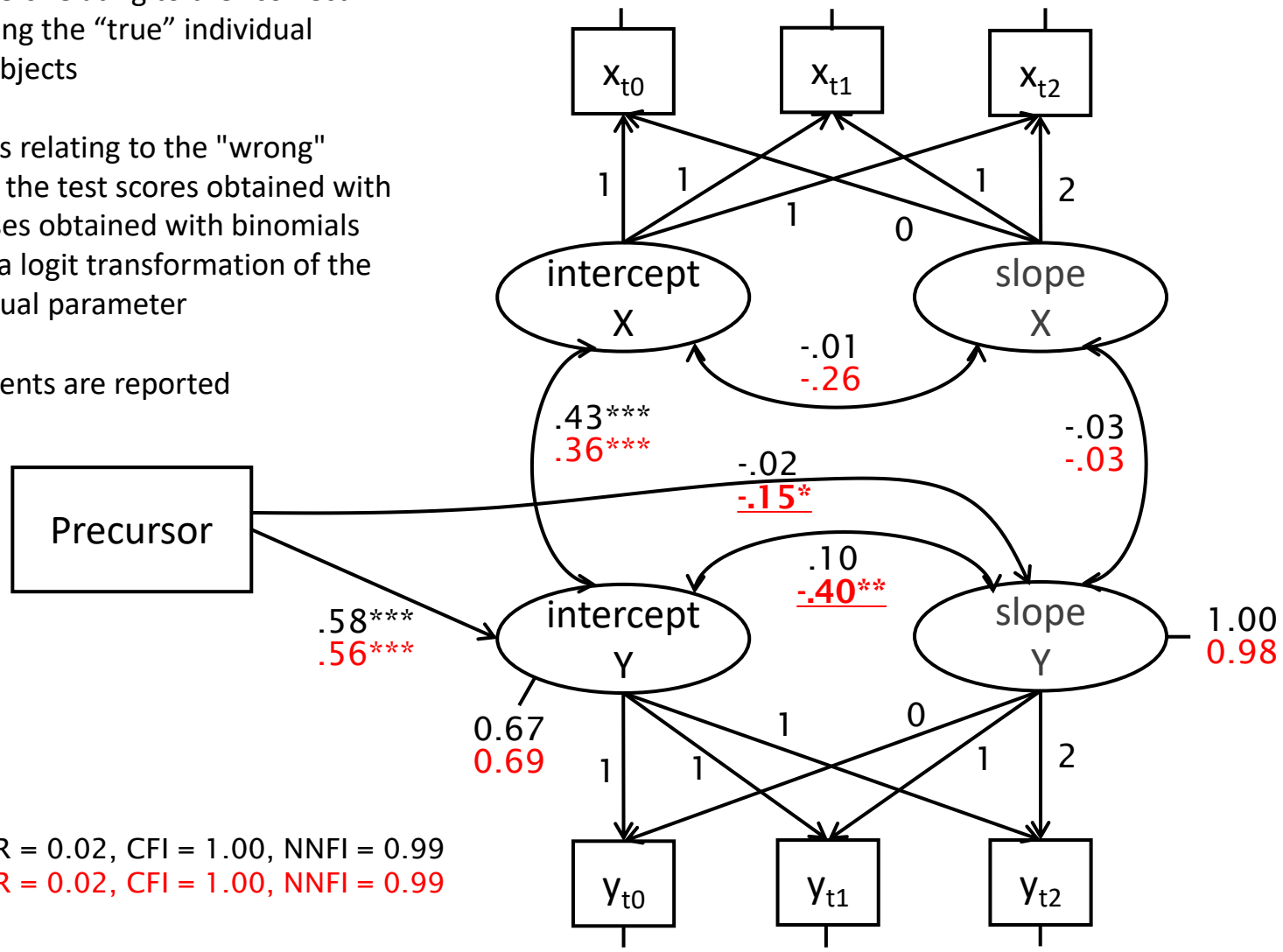
Example of "wrong" estimates under plausible conditions

in **black** the parameters relating to the “correct" model, estimated using the “true” individual parameters of the subjects

in **red** the parameters relating to the "wrong" model, estimated on the test scores obtained with a sum of 0/1 responses obtained with binomials whose probability is a logit transformation of the subject's true individual parameter

Standardized coefficients are reported

N = 1,000



Fit indices

RMSEA = 0.03, SRMR = 0.02, CFI = 1.00, NNFI = 0.99

RMSEA = 0.02, SRMR = 0.02, CFI = 1.00, NNFI = 0.99

Conclusions and suggestions

- Use of bounded [min-max] measures given by sum scores:
  - generates intercept-slope correlations in LGMs that are artifactual;
  - where an external predictor is associated with the intercept, it ends up being predictive of the slope as well

- The problem emerges even with levels of skewness that would normally be considered more than acceptable ([0.20, 0.40]; but a warning bell may be the variation in skewness between times)

- What to do? It is necessary to ensure that the observed measurement is on an interval scale, for example by obtaining it as an estimate of individual parameters from an IRT model, or by working on SEMs starting from the observed responses to the questionnaires...

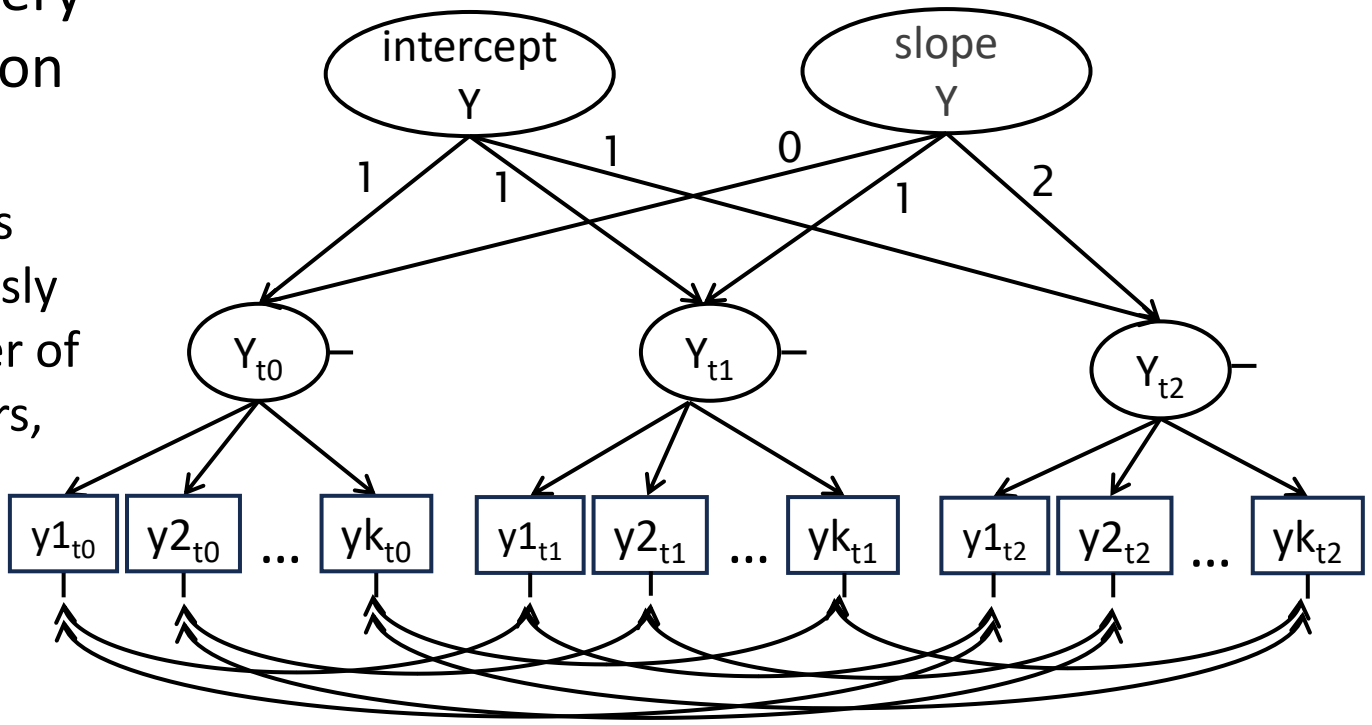
Conclusions and suggestions

...

This would be a very reasonable solution

BUT

it greatly complicates the model, enormously increases the number of estimated parameters, generally imposing larger sample sizes



NB. since the scores y1, y2, … yk are on a Likert/ordinal scale or even dichotomous, remember to specify ordered=TRUE in lavaan

GRAZIE PER L'ATTENZIONE

ENRICO TOFFALINI — DPG, Università di Padova
TOMMASO FERACO — DPG, Università di Padova
MASSIMILIANO PASTORE — DPSS, Università di Padova