Psychological Test and Assessment Modeling, Volume 64, 2022 (3), 253-271

A hierarchical model for data showing an item-position effect

Karl Schweizer[1]*, Andreas Gold*[2]*, and Dorothea Krampen*[3]

Abstract

The hierarchical position-effect model for investigating whether data show an item-position effect is presented. This model includes a hierarchy of latent variables for representing such an effect. Several lower level latent variables associated with subsets of items originating from the segmentation of the item set constitute the first level of the hierarchy while the second level includes the general position-effect latent variable only. This model is proposed for situations where an item-position effect is not exactly monotonically increasing, as the customary position-effect model assumes it. In the application to a real data, model fit varied depending on the specification of the hierarchical model. The comparison with the customary position-effect model yielded similar outcomes, but the best model fit was achieved by the hierarchical position-effect model with linear effect specifications at both levels.

Keywords: item-position effect, hierarchical position-effect model, customary position-effect model, method effect, structural investigation

[1] Karl Schweizer https://orcid.org/0000-0002-3143-2100. *Correspondence concerning this article should be addressed to* Karl Schweizer, Goethe University Frankfurt, Faculty of Psychology and Sports Sciences, Institute of Psychology, Theodor-W.-Adorno-Platz 6, 60323 Frankfurt, Germany. E-mail: K.Schweizer@psych.uni-frankfurt.de

[2] Andreas Gold, Goethe University Frankfurt, Faculty of Psychology and Sports Sciences, Institute of Psychology, Theodor-W.-Adorno-Platz 6, 60323 Frankfurt, Germany.

[3] Dorothea Krampen. Goethe University Frankfurt, Faculty of Psychology and Sports Sciences, Institute of Psychology, Theodor-W.-Adorno-Platz 6, 60323 Frankfurt, Germany.

Investigating an item-position effect in the framework of confirmatory factor analysis (CFA) requires the separation of this effect from the effect due to the latent source of responding that is usually a psychological attribute. So far, this has been achieved by constraints strictly following a mathematical function. An item-position effect may not always exactly unfold corresponding to such a function. For achieving a better degree of compatibility with a variety of possible courses of the item-position effect, a hierarchical model is proposed that allows for adaptation to the actual course of the effect. In the following, the item-position effect is described in some detail and its major explanation is outlined before two position-effect models are described: the customary position-effect model and the hierarchical position-effect model.

An item-position effect is a method effect that is apparent in positive relationships between item characteristics and item positions along the sequence of items that constitute a scale. Item means, item variances and even item reliabilities have turned out to be related to the positions of items (Carlstedt, Gustafsson, & Ullstadius, 2000; Hamilton & Shuminsky, 1990; Hartig, Hölzel, & Moosbrugger, 2007; Knowles, 1988; Knowles & Byers, 1996). This effect has been investigated and confirmed in taking different methodological perspectives. In the framework of item response theory (IRT), it has been demonstrated by using linear logistic test models and multidimensional Rasch models (Debeer & Janssen, 2013; Embretson, 1991; Hohensinn et al., 2008; Kubinger, 2008; Verguts & De Boeck, 2000). Furthermore, CFA models have been employed for examining the item-position effect. In this case, the focus has been on modeling variances and covariances among items (Lozano, 2015; Ren et al., 2012; Ren et al., 2014). The demonstrations of an item-position effect by CFA reveals that this effect is not restricted to mean changes but also extends to changes in individual differences.

A number of different sources that can drive an item-position effect have been considered. The list ranges from learning to impulsivity to fatigue (Carlstedt et al., 2000; Embretson, 1991; Krampen, Gold, & Schweizer, 2020; Kubinger, 2008; Lozano, 2015). The learning hypothesis stating that experiences with previous items influence the responses to subsequent items has received the most support yet (Embretson, 1991; Ren et al., 2012; Ren et al., 2014; Schweizer, Zeller, & Reiß, 2020; Verguts & De Boeck, 2000).

However, the learning hypothesis only provides a vague explanation of the effect because of the manifold of learning theories and the many different ways in which learning can potentially unfold in a sample. For example, the focus can be on the acquisition of rules, as in the seminal study by Carpenter, Just, and Shell (1990) that investigated the detection of the rules underlying the matrix problems of Advanced Progressive Matrices. Associative learning (Kaufman et al., 2009; Williams & Pearlberg, 2006) based on the idea of gradual establishment of links between knowledge items is another possible focus of studying learning. Furthermore, learning may be conceptualized as property of executive processes of working memory that can be improved by extended training of cognitive control in completing cognitive tasks (Kray et al., 2012; Karbach & Verhaeghen, 2014; Stankov & Lee, 2020). Moreover, learning has been associated with a change of the mode of mental information processing. It is assumed that repeated stimulation of cognitive processes can lead to a switch from the conscious and controlled to the unconscious and automatic mode of mental processing (Nordgren, Bos, & Dijksterhuis, 2011; Schweizer et al., 2020; Schweizer et al., 2021). Such a switch can result in improved performance since it means a reduction of the load on working memory and fast processing routines.

Different learning theories give rise to different hypotheses on how learning can be expected to unfold. For example, in learning as rule acquisition (Carpenter et al., 1990) stepwise learning curves are expected whereas associative learning (Kaufman et al., 2009) suggests continuously increasing learning curves. Furthermore, the consequences for individual differences are to be considered. Knowles (1988) reports an increase in item reliability that means an increase in systematic variation of data along a sequence of items. Such an increase can be expected if some participants make great progress in learning while other participants learn nothing or show a small learning rate only. It is even possible that there is an upper limit for learning that is gradually reached by most participants. In such a case, an increase in systematic variation may even be followed by a decrease.

The CFA approach for investigating item-position effects that seeks to reproduce the observed variances and covariances is designed for detecting increase in systematic variation. It makes use of systematically increasing fixed factor loadings. There are reports of investigations with either linearly, logarithmically or quadratically increasing factor loadings (Ren et al., 2012; Schweizer, 2012; Schweizer & Troche, 2018; Sun, Schweizer, & Ren, 2019; Troche et al., 2019; Wang, Zhang, & Schweizer, 2020). A detailed investigation of the course of the item-position effect in data collected in a large sample using Raven's Advanced Progressive Matrices considered different ways of how the effect may unfold. It reports similarity to a quadratically increasing curve but no exact correspondence (Zeller et al., 2017).

The discussion of the possible ways of how an item-position effect may unfold suggests that there may be variability in the course along a sequence of items depending on the cognitive processes stimulated by the scale items and on the participants' properties. Therefore, the representation of the item-position effect by fixed factor loadings strictly following linearly or quadratically increasing functions may not always be appropriate. In the following, we describe a hierarchical model that allows for more flexibility in the representation of the course of an item-position effect and report the results of its evaluation.

The Position-effect Models

This section presents descriptions of the customary position-effect CFA model that shows the structure of the bifactor model and the hierarchical position-effect structural equation model (SEM). *Note*. A position-effect (also referred to as sequence effect) can be observed when comparing two experimental treatment levels of a repeated measures design. In this case, the focus is on check and avoidance of error regarding the treatment effect in the statistical investigation. In contrast, when investigating the position-effect in a sequence of items, the focus is on representing it appropriately and measuring the amount of latent variance for which it accounts. We make the difference obvious by using the term item-position effect for addressing the effect in a sequence of items. Yet, we stay with the denotation as position-effect when characterizing a latent variable of a measurement model.

The *customary position-effect CFA model* includes two latent variables, $\xi_{\text{Attribute}}$ and $\xi_{\text{Position-effect}}$, that represent the attribute of interest and the item-position effect, respectively. It also includes p centered manifest variables. Figure 1 provides an illustration of this model.

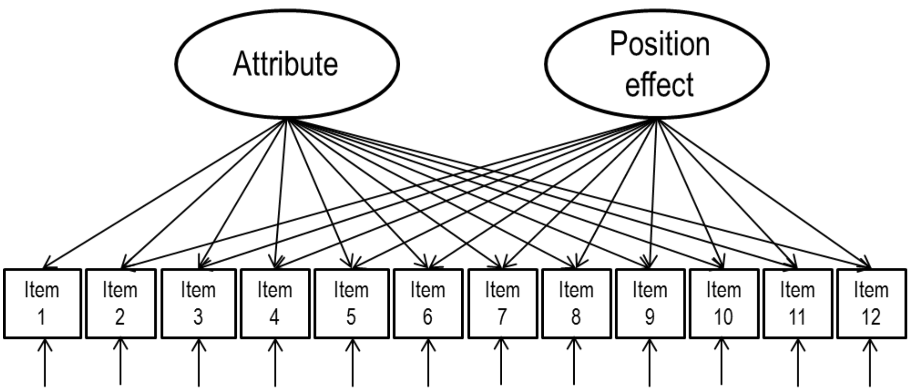

Figure 1.
Position-effect CFA model as bifactor CFA model including the attribute latent variable with factor loadings of all manifest variables and the position-effect latent variable with factor loadings of all but one manifest variables. Arrows with small shafts represent random influences. An example with 12 items.

As is obvious from the arrows linking the latent variables (ellipses) to the manifest variables (rectangles), the attribute latent variable shows factor loadings from all manifest variables whereas the position-effect latent variable has factor loadings from all but the first manifest variables since the first items cannot show such an effect by definition.

The formal model is given by

$$\boldsymbol{x} = \boldsymbol{\lambda}_{\text{Attribute}}\xi_{\text{Attribute}} + \boldsymbol{\lambda}_{\text{Position-effect}}\xi_{\text{Position-effect}} + \boldsymbol{\delta} \quad (1)$$

where $\mathbf{x}$ is the $p \times 1$ vector of centered manifest variables, $\boldsymbol{\lambda}_{\text{Attribute}}$ and $\boldsymbol{\lambda}_{\text{Position-effect}}$ are the $p \times 1$ vectors of factor loadings on latent variables $\xi_{\text{Attribute}}$ and $\xi_{\text{Position-effect}}$ and $\boldsymbol{\delta}$ is the $p \times 1$ vector of idiosyncratic disturbances. It can be perceived as an extension of the congeneric CFA model (Jöreskog, 1971) or alternatively as a version of the bifactor model (Reise, 2012).

In order to assure that the position-effect latent variable accounts for the systematic variation of data due to an item-position effect, the factor loadings are fixed to pre-specified values. In the case of assuming a linear increase, the fixation of the factor loadings of manifest variables x_i on $\xi_{\text{Position-effect}}$ is given by

$$\lambda_{\text{Position}-\text{effect}}(i) = (i-1)^1/(p-1)^1 \quad (2)$$

with ($i = 1,\ldots, p$) and in the case of assuming a quadratic increase by

$$\lambda_{\text{Position}-\text{effect}}(i) = (i-1)^2/(p-1)^2 \ . \quad (3)$$

This kind of fixation means that fitting of model to data cannot be achieved by estimating factor loadings. In each latent variable with fixed factor loadings ξ_j ($j = 1,\ldots, q$) the corresponding variance parameter ϕ_j of the model-implied covariance matrix ($\mathbf{\Sigma}$) is estimated.

When dichotomous data are to be investigated, an additional link transformation is required that can be achieved by multiplication of factor loadings with weight w_i:

$$w_i = \sqrt{Pr(x_i = 1)\,[1 - Pr(x_i = 1)]} \quad (4)$$

where Pr symbolizes probabilities.

The *hierarchical position-effect SEM model* is proposed as another method for investigating data showing an item-position effect. It is expected to show more flexibility for deviations from a strictly linear or quadratic course of effect than the customary model. This flexibility is achieved by segmentation of the sequence of items into neighboring subsets of items together with a hierarchical structure. In this model each subset of items loads on its own latent variable so that there are as many first-order latent variables as there are subsets of items in addition to the first-order attribute latent variable. Furthermore, the model includes a second-order latent variable with a factor loading from each first-order latent variable representing a subset but not from the attribute latent variable. An illustration of such a model is provided by Figure 2.

For making it easy to survey Figure 2, the position-effect latent variables are shifted from the upper half of the figure to the lower half. As with the customary position-effect model, there is no factor loading of the first item on the corresponding latent variable.

The first-order structure of the hierarchical position-effect model is given by

$$\boldsymbol{y} = \boldsymbol{\lambda}_{\text{Attribute}}\eta_{\text{Attribute}} + \sum_i^q \boldsymbol{\lambda}_{\text{Position}-\text{effect\_}i}\eta_{\text{Position}-\text{effect\_}i} + \boldsymbol{\varepsilon} = \boldsymbol{\Lambda\eta} + \boldsymbol{\varepsilon} \quad (5)$$

where $\mathbf{y}$ is the $p \times 1$ vector of manifest variables, $\boldsymbol{\lambda}_{\text{Attribute}}$ and $\boldsymbol{\lambda}_{\text{Position-effect\_}i}$ are the $p \times 1$ vectors of factor loadings on latent variables $\eta_{\text{Attribute}}$ and $\eta_{\text{Position-effect\_}i}$ and $\boldsymbol{\varepsilon}$ is the p

× 1 vector of idiosyncratic disturbances. Furthermore, $\mathbf{\Lambda}$ is the $p \times (q + 1)$ matrix of factor loadings on latent variables and $\boldsymbol{\eta}$ the $(q + 1) \times 1$ vector of latent variables.

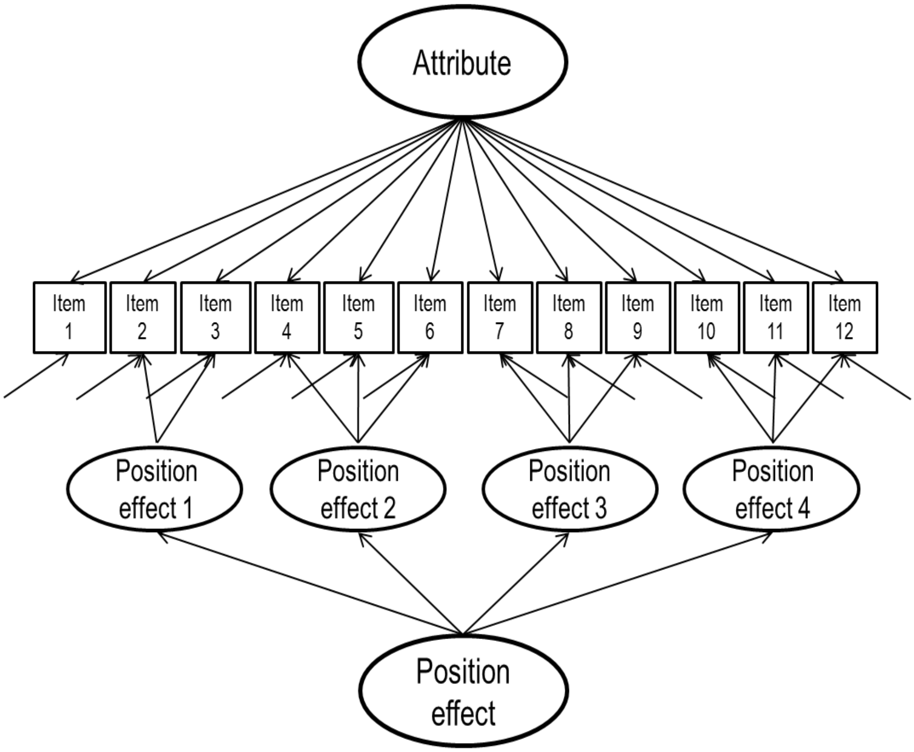

Figure 2.
Hierarchical position-effect CFA model with one attribute latent variable and four specific position-effect latent variables as first-order latent variables and a second-order position effect latent variable. Arrows with small shafts represent random influences. An example with 12 items, four first-order position-effect latent variables and one second-order position-effect latent variable.

A structural model that turns the hierarchical model into a SEM model is additionally necessary. The structural model relating first-order and second-order latent variables is given by

$$\boldsymbol{\eta} = \boldsymbol{\gamma}\xi_{\text{Position}-\text{effect}} + \boldsymbol{\zeta} \quad (6)$$

where $\boldsymbol{\gamma}$ is the $(q + 1) \times 1$ vector of loadings of first-order latent variables on second-order latent variable $\xi_{\text{Position-effect}}$ and $\boldsymbol{\zeta}$ the $(q + 1) \times 1$ vector of residuals. Since $\eta_{\text{Attribute}}$ does not load on ξ, $\gamma_1 = 0$ and $\zeta_1 = 0$. Combining Equations 5 and 6 gives the final hierarchical SEM model:

$$\boldsymbol{y} = \boldsymbol{\Lambda\gamma}\xi_{\text{Position-effect}} + \boldsymbol{\Lambda}\zeta + \boldsymbol{\varepsilon} \quad (7)$$

(for more information on hierarchical models see Schweizer, Moosbrugger, and Schermelleh-Engel, 2003).

As in the case of the customary position-effect CFA model, investigating an item-position effect requires to assure that the position-effect latent variables account for the systematic variation of data due to the item-position effect. This involves the fixation of factor loadings according to pre-specified values while the variances of the latent variables are set free for estimation. This can be accomplished according to Equations 2 or 3. In the case of the hierarchical position-effect SEM model it is also useful to fix the γ coefficients accordingly. The flexibility of this model regarding deviations from a strictly linear or quadratic course of increase is achieved by the estimation of the latent variances, the second-order fixations and the possibility to choose between different segmentations.

Application to a Real Data

The study served the comparison of the customary and hierarchical CFA models in an application to data showing an item-position effect. A real data was used for the investigation. The data were collected by Advanced Progressive Matrices (APM) (Raven, Raven, & Court, 1997) in a sample of 104 university students. The responses were coded as either correct ($x = 1$) or incorrect ($x = 0$).

Method

Customary and hierarchical CFA models were prepared according to Equations 1 and 7 in considering linear and quadratic increases (Equations 2 and 3). Additional information regarding the hierarchical CFA models is that not only the variance parameters of the first-order latent variables were set free for estimation but also the covariances between neighboring latent variables whereas the variance parameter of the second-order latent variable was fixed to one. This is necessary because of unexplained correlation between neighboring items that leads to bad model fit if it is ignored. The fixations for the links relating first-order and second-order latent variables were varied. They are reported in combination with fit results (see corresponding tables). Additionally, a simple one-factor CFA model was estimated in order to demonstrate the advantage of considering the item-position effect. In this case, the latent variable was specified as attribute latent variable, i.e., all factor loadings were constrained to one.

Figure 3 provides an illustration of an example of a hierarchical position-effect CFA model for APM data.

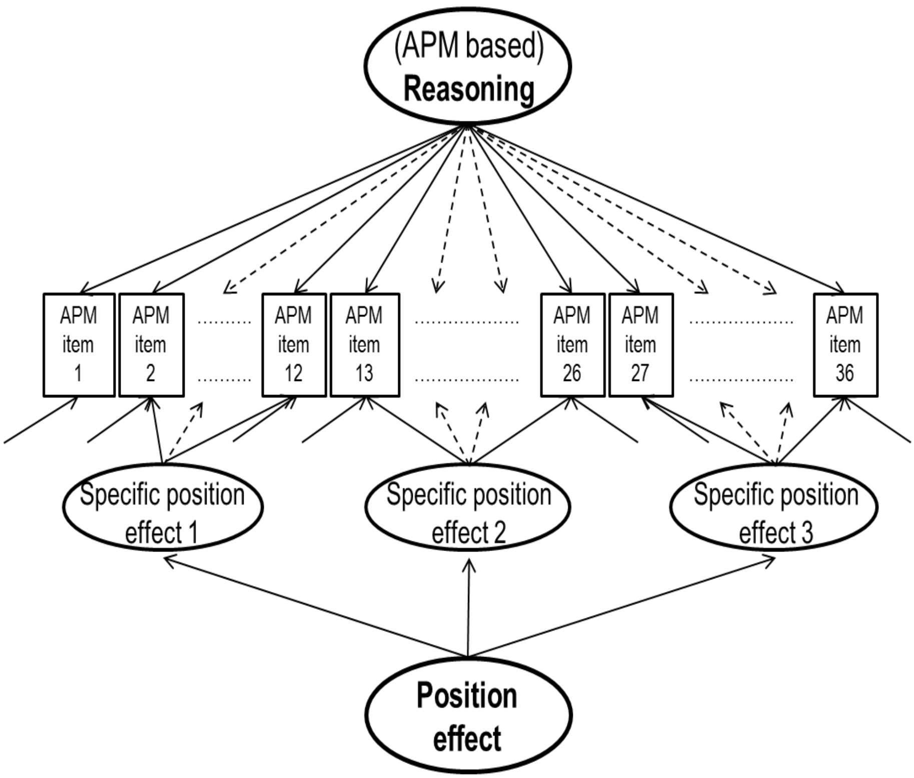

Figure 3.
Hierarchical position-effect CFA model for APM data (36 items) with reasoning ability latent variable and three specific position-effect latent variables and a second-order position-effect latent variable (dotted and dashed lines indicate omitted parts because of the large number of items). Arrows with small shafts represent random influences.

The subdivision of the set of 36 APM items into three even subsets giving rise to three first-order position-effect latent variables was selected for this example. Because of the large number of items it was not possible to represent all of them in this figure. These omissions were symbolized by dotted and dashed lines. Furthermore, since

APM represents reasoning the label *attribute* of Figures 1 and 2 was replaced be *reasoning*.

The models were estimated using LISREL software (Jöreskog & Sörbom, 2006). Probability-based covariances served as input. The focus of the investigation was on model fit. The following fit indices are considered in the evaluation of results: RMSEA, SRMR, NNFI, CFI and AIC. Furthermore, in some investigations latent variances were estimated in order to provide information on the size of the item-position effect. For this purpose the variance parameter was scaled to correspond to the sum of squared factor loadings (Schweizer, 2011; Schweizer & Troche, 2019; Schweizer, Troche, & DiStefano, 2019).

Results

Results for customary models. The investigation of the data by customary CFA models led to the fit statistics provided in Table 1.

The results presented in the first row were obtained by the one-factor CFA model that did not include a position-effect latent variable. RMSEA indicated acceptable model fit whereas all other fit indices with a cutoff (SRMR, NNFI and CFI) (see Distefano, 2016) suggested model misfit. The results obtained by the position-effect CFA model with linearly increasing fixations (second row) showed fit improvement with the exception of SRMR although there was no change in quality level. Allowing the latent variables to correlate with each other (third row) further improved model fit. RMSEA signified good model fit, SRMR acceptable model fit. NNFI and CFI were still indicating model misfit according to cutoffs. Replacing the linear increase by quadratic increase led to small numerical improvements in model fit in the version with no correlation among the latent variables whereas in the other version there was virtually no improvement (see fourth and fifth rows).

Table 2 includes the scaled estimates of latent variances.

This table comprises four columns with results on the right side. The first and second columns include the variance estimates for the attribute latent variable and the position-effect latent variable in corresponding order. Including the position-effect latent variable into the model increased the variance estimate of the attribute latent variable from 0.82 to an average of 1.34. The variance size of the position-effect latent variables was on average 58 percent of the variance size of the attribute latent variable. All covariances were negative. The fourth column includes the variance sums that were obtained by adding up the two variance estimates and combining them with two times the covariance. The variance sums for the two models with correlated latent variables virtually corresponded to the variance estimate for the one-factor model.

Table 1
Fit Results Observed in Investigating Ability Data with CFA Models Including an Ability Latent Variable Alone or in Combination with a Position-Effect Latent Variable (N = 104)

Characteristics of model	χ^2	*df*	RMSEA	SRMR	NNFI	CFI	AIC	*r*
One factor (ability)	902.7	629	0.065	0.110	0.76	0.76	976.7	-
Two factors (ability and *linear* PE)	801.8	628	0.052	0.146	0.78	0.78	877.8	-
Two *correlated* factors (ability and *linear* PE)	745.4	627	0.043	0.099	0.82	0.82	823.4	-0.77
Two factors (ability and *quadratic* PE)	776.6	628	0.048	0.130	0.80	0.80	852.6	-
Two *correlated* factors (ability and *quadratic* PE)	745.2	627	0.043	0.099	0.82	0.82	823.2	-0.64

Note. PE: position effect, χ^2: chi-square, *df*: degree of freedom, RMSEA: root mean square error of approximation, SRMR: standardized root mean square residual, NNFI: nonnormed fit index, CFI: comparative fit index, AIC: Akaike information criterion, *r*: (standardized) correlation.

Table 2
Variance Estimates of Latent Variables of One-factor and Two-factor CFA Models with and without Factor Correlations (N = 104)

Characteristics of model	$Var_{Ability}$	Var_{PE}	Cov	Σ_{var}
One factor (ability)	0.82	-	-	0.82
Two factors (ability and *linear* PE)	1.07	0.62	-	1.69
Two *correlated* factors (ability and *linear* PE)	1.92	1.30	-1.21	0.80
Two factors (ability and *quadratic* PE)	1.01	0.50	-	1.51
Two *correlated* factors (ability and *quadratic* PE)	1.36	0.73	-0.64	0.81

Note. PE: position effect, $Var_{Ability}$: variance of ability latent variable, Var_{PE}: variance of the position-effect latent variable, Cov: covariance, Σ_{var}: $Var_{Ability}$ plus Var_{PE} plus 2 × Cov.

Results for hierarchical models. The results reported in the following paragraphs were obtained by means of hierarchical CFA models. For investigating the data, the sequence of items was subdivided in subsets of either four, five or six items. The most promising results were observed for subsets of five items. These results are included in Table 3 whereas the results based on the other subsets can be found as Appendices 1 and 2.

Table 3

Fit Results Observed by Hierarchical Model with Eight First-order Factors and One Second-order Factor (Item Segmentation into Five Item Subsets) (N = 104)

Second-order fixations	χ^2	*df*	RMSEA	SRMR	NNFI	CFI	AIC	*r*
Hierarchical linear-linear/quadratic multi-factor model								
1 1 1 1 1 1 1[1]	724.6	616	0.041	0.277	0.77	0.77	824.6	-
1 1 1 1 1 1 1[2]	715.0	615	0.040	0.102	0.82	0.82	817.0	-0.85
.1 .1 .1 .1 .1 .1 .1[1]	742.4	616	0.045	0.109	0.80	0.81	842.4	-
.1 .1 .1 .1 .1 .1 .1[2]	738.2	615	0.044	0.103	0.80	0.81	840.2	-1.52[3]
.1 .2 .3 .4 .5 .6 .7[1]	728.2	616	0.042	0.149	0.80	0.80	828.2	
.1 .2 .3 .4 .5 .6 .7[2]	714.7	615	0.040	0.098	0.82	0.83	816.7	-0.62
.02 .08 .18 .32 .51 .73 1[1]	727.5	616	0.042	0.162	0.80	0.79	827.5	-
.02 .08 .18 .32 .51 .73 1[2]	719.8	615	0.041	0.104	0.81	0.81	821.8	-0.45
Hierarchical quadratic-linear/quadratic multi-factor model								
1 1 1 1 1 1 1[1]	727.0	616	0.042	0.187	0.79	0.79	827.0	-
1 1 1 1 1 1 1[2]	719.4	615	0.041	0.101	0.81	0.82	821.4	-0.65
.1 .1 .1 .1 .1 .1 .1[1]	747.5	616	0.046	0.114	0.79	0.80	847.5	-
.1 .1 .1 .1 .1 .1 .1[2]	740.0	615	0.044	0.103	0.80	0.81	842.0	-1.91[3]
.1 .2 .3 .4 .5 .6 .7[1]	731.6	616	0.043	0.126	0.80	0.81	831.6	-
.1 .2 .3 .4 .5 .6 .7[2]	721.1	615	0.041	0.101	0.82	0.82	823.1	-0.46
.02 .08 .18 .32 .51 .73 1[1]	728.3	616	0.042	0.134	0.80	0.81	828.3	-
.02 .08 .18 .32 .51 .73 1[2]	721.1	615	0.041	0.105	0.81	0.82	823.1	-0.31

Note. χ^2: chi-square, *df*: degree of freedom, RMSEA: root mean square error of approximation, SRMR: standardized root mean square residual, NNFI: nonnormed fit index, CFI: comparative fit index, AIC: Akaike information criterion, *r*: (standardized) correlation.

[1] The attribute and second-order position-effect latent variables are uncorrelated.

[2] The attribute and second-order position-effect latent variables are correlated.

[3] Unlike Pearson correlations estimated correlation can be larger than one.

Table 3 is subdivided into two parts. Results obtained by models with fixations of first-order factor loadings showing a linear increase are presented in the first part while in the case of a quadratic increase they are presented in the second part. The first column of this table informs about the fixations used for relating first-order and second-order latent variables to each other. The results reported in the first part of Table 3 compare with the results reported in the second and third rows of Table 1 and the results of the second part of Table 3 with what is reported in the fourth and fifth rows of Table 1. For making it easy to capture the main message of Table 3, we computed means for the results obtained by linear models with and without correlated latent variables and also by quadratic models with and without correlated latent variables.

The comparison of the linear customary and hierarchical models without correlated latent variables revealed the numerically better results for the hierarchical model regarding χ^2, RMSEA, NNFI, CFI and AIC. After allowing the latent variables to correlate, there were only two better results for the hierarchical model regarding χ^2 and RMSEA. The numerically best model fit was observed for the linear hierarchical model when there were correlated latent variables and the following set of numbers was used for the fixation of the first-order to second-order relationships: .1 .2 .3 .4 .5 .6 .7.

When comparing the quadratic customary and hierarchical models without correlated latent variables of the second part of Table 3 with the corresponding part of Table 1, numerically better results characterized the hierarchical model regarding χ^2, RMSEA, NNFI, CFI and AIC. Inserting correlations between the latent variables reduced the numerical advantage for the hierarchical model to the χ^2 and RMSEA results. The numerically best model fit was observed for the quadratic hierarchical model when the latent variables were correlated and the following set of numbers was used for the fixation of the first-order to second-order relationships: 1 1 1 1 1 1 1.

The overall best fitting model was a hierarchical position-effect model with linear increases of the constraints for the first-order latent variable and also for the second-order latent variable. This model showed fit improvement over the best fitting customary model regarding χ^2, RMSEA, SRMR, CFI and AIC.

In sum, the investigations by the hierarchical position-effect model revealed considerable variation due to different fixations of the first-order to second-order relationships and yielded the configuration of fixations giving rise to the overall best fitting model.

Discussion

The research reported in this essay focuses on a method effect that may be considered a negligible nuisance. This characterization mirrors a position that is favored because it liberates applied research from the necessity to check data for the presence of such an effect. Yet, negligence of a method effect may not be the best research strategy. Research based on the multitrait-multimethod design has shown that taking method effects into account increases the validity of results (Byrne, 2016). Method effects have been found to suggest convergent validity where there is no convergent validity and vice versa (Campbell & Fiske, 1959). An item-position effect is another method effect that can influence the validity of a scale. In the present study it accounted for about a third of the latent variance of the scale, which is remarkable. Von Gugelberg, Schweizer, and Troche (2021) even report an experimental condition where the item-position effect surmounted the effect due to the construct that was reasoning.

The present study included the application to a real data. The usefulness of such a study can be questioned since the true underlying structure of the data at hand is not known. This means it is not a demonstration of the efficiency of the method in recovering the given latent structure used in data generation. Regarding this point, we like to argue that simulation studies demonstrating the efficiency of the basic method for investigating the structure of data with an item-position effect are available (e.g., Schweizer, 2012; Schweizer & Troche, 2018). Given the similarity of non-hierarchical and hierarchical methods it is unlikely that the hierarchical method captures data variation that is unrelated to the data variation captured by the non-hierarchical method. Furthermore, we like to point out that the insight provided by simulation studies can be very limited. There are many simulation studies that assume conditions which we do not observe in real data. For example, there are studies assuming factor loading sizes which we have never observed when investigating real data. The relevance of an effect can only be shown by using real data. Therefore, demonstrating an effect in a real data is as important as demonstrating efficiency in using simulated data. Regarding the present research, the next step has to be a simulation study.

Regarding the present study, we would like to point out that the results provide further support for the hypothesis that performance in completing the items of a reasoning scale, like APM, shows an item-position effect. The inclusion of a position-effect latent variable into the one-factor CFA model improved model fit substantially, as is indicated by a substantial chi-square difference test result and CFI as well as RMSEA differences in model fit larger than cutoffs derived from Cheung and Rensvold's (2002) work. Support for the assumption of an item-position effect was provided by customary position-effect models and hierarchical position-effect models.

The outcomes for the hierarchical position-effect models showed some variability but rather no major differences. Some models differed substantially from each other according to the cutoff for CFI differences while no such differences between models were observed regarding the cutoff for RMSEA differences. This means that the search for the specification of the model that especially well reflects how an item-position effect unfolds implies a small though substantial advantage regarding model fit. The search for further improvement in model fit requires the consideration of other possible method effects as, for example, the difficulty effect underlying a difficulty factor (Bandalos & Gerstner, 2016) or omissions due to the lack of enough processing speed (Schweizer, Gold, & Krampen, 2020).

It was interesting to observe that the overall best model fit was found for a model assuming a linearly increasing item-position effect when APM data were investigated. This observation contradicts the results by Zeller et al. (2017) who observed a quadratically increasing effect in APM data. One possible reason explaining the difference is the type of model since Zeller et al. (2017) only employed customary models but no hierarchical models. Another possible reason could be that the samples differ from each other. There is the possibility that different percentages of high ability and low ability participants lead to differences in how the item-position effect unfolds. Both these options will need to be further explored in the future. In addition, it would be interesting to examine item-position effects with other tests of fluid intelligence (e.g., inductive reasoning tests based on numbers or letters from the alphabet). The domain can also be expanded to include measures of other broad factors of the Cattell-Horn-Carroll (CHC) theory of intelligence (see Schneider & McGrew, 2018).

Finally, we would like to point out that the present study shows limitations. The selection of a real data limited the number of manifest variables to the number of items of the scale with consequences for the possible subsets of items that provide the basis for the first-order latent variables. Furthermore, there is a restriction to the course of increase of the effect that is typical to APM. Other types of increases could not be taken into account.

Overall, we think that the availability of the hierarchical position-effect model considerably enlarges the variety of alternative representations of an item-position effect. This means that the chance of detecting such an effect is increased.

References

Bandalos, D. L., & Gerstner, J. J. (2016). Using factor analysis in test construction. In K. Schweizer, & C. DiStefano (Eds.), *Principles and methods of test construction. Standards and recent advances* (pp. 26–51). Göttingen: Hogrefe.

Byrne, B. M. (2016). Using multitrait-multimethod analysis in testing for evidence of construct validity. In K. Schweizer, & C. DiStefano (Eds.), *Principles and methods of test construction. Standards and recent advances* (pp. 288–307). Göttingen: Hogrefe.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. https://doi.org/10.1037/h0046016

Carlstedt, B., Gustafsson, J.-E., & Ullstadius, E. (2000). Item sequencing effects on the measurement of fluid intelligence. *Intelligence, 28*(2), 145–160. https://doi.org/10.1016/S0160-2896(00)00034-9

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*(3), 404–431. https://dx.doi.org/10.1037/0033-295X.97.3.404

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Debeer, D., & Janssen, R. (2013). Modeling item-position effects within an IRT framework. *Journal of Educational Measurement, 50*(2), 164–185. https://doi.org/10.1111/jedm.12009

DiStefano, C. (2016). Examining fit with structural equation models. In K. Schweizer, & C. DiStefano (Eds.), *Principles and methods of test construction. Standards and recent advances* (pp. 166–193). Göttingen: Hogrefe.

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 56*(3), 495–515. https://doi.org/10.1007/BF02294487

Hamilton, J. C., & Shuminsky, T. R. (1990). Self-awareness mediates the relationship between serial position and item reliability. *Journal of Personality and Social Psychology, 59*(6), 1301–1307. https://doi.org/10.1037/0022-3514.59.6.1301

Hartig, J., Hölzel, B., & Moosbrugger, H. (2007). A confirmatory analysis of item reliability trends (CAIRT): differentiating true score and error variance in the analysis of item context effects. *Multivariate Behavioral Research, 42*(1), 157–183. https://doi.org/10.1080/00273170701341266

Hohensinn, C., Kubinger, K. D., Reif, M., Holocher-Ertl, S., Khorramdel, L., & Frebort, M. (2008). Examining item-position effects in large-scale assessment using the Linear Logistic Test Model. *Psychology Science Quarterly, 50*(3), 391–402.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 36*, 109–133. https://doi.org/10.1007/BF02291393

Jöreskog, K. G., & Sörbom, D. (2006). *LISREL 8.80*. Lincolnwood, IL: Scientific Software International Inc.

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: a meta-analysis of executive-control and working memory training in older adults. *Psychological Science, 25*(11), 2027–2037. https://doi.org/10.1177/0956797614548725

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence, 37*(4), 374–382. https: //doi.org/10.1016/j.intell.2009.03.004

Knowles, E. S. (1988). Item context effects on personality scales: measuring changes the measure. *Journal of Personality and Social Psychology, 55*(2), 312–320. https://doi:10.1037/0022-3514.55.2.312

Knowles, E. S., & Byers, B. (1996). Reliability shifts in measurement reactivity: driven by content engagement or self-engagement? *Journal of Personality and Social Psychology, 70*(5), 1080–1090. https://doi:10.1037//0022-3514.70.5.1080

Krampen, D., Gold, A., & Schweizer, K. (2020). Does impulsivity contribute to the item-position effect? *Psychological Test and Assessment Modeling, 64*(3), 375–385.

Kray, J., Karbach, J., Haenig, S., & Freitag, C. (2012). Can task-switching training enhance executive control functioning in children with attention deficit/-hyperactivity disorder? *Frontiers in Human Neuroscience, 5*, 1–9. https://doi.org/10.3389/fnhum.2011.00180

Kubinger, K. D. (2008). On the revival of the Rasch model-based LLTM: from constructing tests using item generating rules to measuring item administration effects. *Psychology Science Quarterly, 50*(3), 311–327.

Lozano, J. H. (2015). Are impulsivity and intelligence truly related constructs? Evidence based on the fixed-links model. *Personality and Individual Differences, 85*, 192–198. https://doi.org/10.1016/j.paid.2015.04.049

Nordgren, L. F., Bos, M. W., & Dijksterhuis, A. (2011). The best of both worlds: integrating conscious and unconscious thought best solves complex decisions. *Journal of Experimental Social Psychology, 47*(2), 509–511. https://doi.org/10.1016/j.jesp.2010.12.007

Raven, J. C., Raven, J., & Court, J. H. (1997). *Raven's progressive matrices and vocabulary scales*. Edinburgh: J. C. Raven Ltd.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Ren, X., Goldhammer, F., Moosbrugger, H., & Schweizer, K. (2012). How does attention relate to the ability-specific and position-specific components of reasoning measured by APM? *Learning and Individual Differences, 22*(1), 1–7. https://doi.org/10.1016/j.lindif.2011.09.009

Ren, X., Wang, T., Altmeyer, M., & Schweizer, K. (2014). A learning-based account of fluid intelligence from the perspective of the position effect. *Learning and Individual Differences, 31*, 30–35. https://doi.org/10.1016/j.lindif.2014.01.002

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan, & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 73–163). New York: The Guilford Press.

Schweizer, K. (2011). Scaling variances of latent variables by standardizing loadings: applications to working memory and the position effect. *Multivariate Behavioral Research, 46*(6), 938–955. https://doi.org/10.1080/00273171.2011.625312

Schweizer, K. (2012). The position effect in reasoning items considered from the CFA perspective. *International Journal of Educational and Psychological Assessment, 11*(2), 44–58.

Schweizer, K., Gold, A., & Krampen, D. (2020). A semi-hierarchical confirmatory factor model for speeded data. *Structural Equation Modeling, 27*(5), 773–780. https://doi: 10.1080/10705511.2019.1707083

Schweizer, K., Moosbrugger, H., & Schermelleh-Engel, K. (2003). Models for investigating hierarchical structures in differential psychology. *Methods of Psychological Research Online, 8*, 159–180.

Schweizer, K., & Troche, S. (2018). Is the factor observed in investigations of the item-position effect actually the difficulty factor? *Educational and Psychological Measurement, 78*(1), 46–69. https://doi.org/10.1177/0013164416670711

Schweizer, K., & Troche, S. (2019). The EV scaling method for variances of latent variables. *Methodology, 15*(4), 175–184. https://doi.org/10.1027/1614-2241/a000179

Schweizer, K., Troche, S., & DiStefano, C. (2019). Scaling the variance of a latent variable while assuring constancy of the model. *Frontiers in Psychology* (Section Quantitative Psychology and Measurement), *10,* ArtID 887. https://doi.org/10.3389/fpsyg.2019.00887

Schweizer, K., Troche, S., Rammsayer, T., & Zeller, F. (2021). Inductive reasoning and its underlying structure: support for difficulty and item-position effects. *Advances in Cognitive Psychology, 17*(4), 274–283. https://doi.org/10.5709/acp-0336-5

Schweizer, K., Zeller, F., & Reiß, S. (2020). Higher-order processing and change-to-automaticity as explanations of the item-position effect in reasoning tests. *Acta Psychologica*. https://doi.org/10.1016/j.actpsy.2019.102991

Stankov, L., & Lee, J. (2020). We can boost IQ: revisiting Kvashchev's experiment. *Journal of Intelligence, 8*(4), 41, 1–12. https://doi.org/10.3390/jintelligence8040041

Sun, S., Schweizer, K., & Ren, X. (2019). Item-position effect in Raven's Matrices: a developmental perspective. *Journal of Cognition and Development, 20*(3), 370–379. https://doi.org/10.1080/15248372.2019.1581205

Troche, S., Wagner, F. L., Schweizer, K., & Rammsayer, T. H. (2019). The structural validity of the Cultural Fair Test under consideration of the item-position effect. *European Journal of Psychological Assessment, 35*(2), 182–189. https://doi.org/10.1027/1015-5759/a000384

Verguts, T., & De Boeck, P. (2000). A rasch model for detecting learning while solving an intelligence test. *Applied Psychological Measurement, 24*(2), 151–162. https://doi.org/10.1177/01466210022031589

von Gugelberg, H. M., Schweizer, K., & Troche, S. J. (2021). The dual mechanisms of cognitive control and their relation to reasoning and the item-position effect. *Acta Psychologica, 221*. http://doi.org/10.1016/j.actpsy.2021.103448

Wang, T., Zhang, Q., & Schweizer, K. (2020). Investigating the item-position effect in a longitudinal data with special emphasis on the information provided by the variance parameter. *Psychological Test and Assessment Modeling, 64,* 404–417.

Williams, B. A., & Pearlberg, S. L. (2006). Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence, 34*(2), 177–191. https://doi.org/10.1016/j.intell.2005.03.007

Zeller, F., Krampen, D., Reiß, S., & Schweizer, K. (2017). Do adaptive representations of the item-position effect in APM improve model fit? A simulation study. *Educational and Psychological Measurement, 77*(5), 743–765. https://doi.org/10.1177/0013164416654946

Appendix 1

Table: Fit Results Observed by Hierarchical Model with Ten First-order Factors and One Second-order Factor (Item Segmentation into Four Item Subsets) (N = 104)

Second-order fixations	χ^2	*df*	RMSEA	SRMR	NNFI	CFI	AIC	*r*
Hierarchical linear-linear/quadratic multi-factor model								
1 1 1 1 1 1 1 1 1	737.7	612	0.045	0.282	0.75	0.76	845.7	-
1 1 1 1 1 1 1 1 1	717.0	611	0.041	0.104	0.81	0.82	827.0	-0.86
.1 .1 .1 .1 .1 .1 .1 .1 .1	769.1	612	0.050	0.108	0.79	0.79	877.1	-
.1 .1 .1 .1 .1 .1 .1 .1 .1	762.4	611	0.049	0.105	0.79	0.80	872.4	-1.34[1]
.1 .2 .3 .4 .5 .6 .7 .8 .9	902.0	612	0.068	0.304	0.62	0.63	1010	-
.1 .2 .3 .4 .5 .6 .7 .8 .9	721.1	611	0.042	0.100	0.81	0.81	831.1	-0.71
.01234 … (quad) … 1	937.6	612	0.072	0.153	0.72	0.73	1045	-
.01234 … (quad) … 1	724.3	611	0.042	0.104	0.80	0.81	834.3	-0.53
Hierarchical quadratic-linear/quadratic multi-factor model								
1 1 1 1 1 1 1 1 1	733.8	612	0.044	0.189	0.77	0.78	841.8	-
1 1 1 1 1 1 1 1 1	723.6	611	0.042	0.104	0.80	0.81	833.6	-0.72
.1 .1 .1 .1 .1 .1 .1 .1 .1	775.7	612	0.051	0.109	0.78	0.79	883.7	-
.1 .1 .1 .1 .1 .1 .1 .1 .1	765.1	611	0.049	0.104	0.79	0.79	875.1	-1.61[1]
.1 .2 .3 .4 .5 .6 .7 .8 .9	735.6	612	0.044	0.143	0.79	0.79	843.6	-
.1 .2 .3 .4 .5 .6 .7 .8 .9	726.8	611	0.043	0.101	0.80	0.81	836.8	-0.50
. 01234 … (quad) … 1	735.3	612	0.044	0.134	0.79	0.79	843.3	-
. 01234 … (quad) … 1	729.8	611	0.043	0.105	0.80	0.81	839.8	-0.35

Note. χ^2: chi-square, *df*: degree of freedom, RMSEA: root mean square error of approximation, SRMR: standardized root mean square residual, NNFI: nonnormed fit index, CFI: comparative fit index, AIC: Akaike information criterion, *r*: (standardized) correlation.

[1] Unlike Pearson correlations estimated correlation can be larger than one.

Appendix 2

Table: Fit Results Observed by Hierarchical Model with Seven First-order Factors and One Second-order Factor (Item Segmentation into Six Item Subsets) (N = 104)

Second-order fixations	χ^2	*df*	RMSEA	SRMR	NNFI	CFI	AIC	$r_{A\text{-}PE}$
Hierarchical linear-linear/quadratic multi-factor model								
1 1 1 1 1 1	799.6	618	0.053	0.284	0.66	0.67	895.6	-
1 1 1 1 1 1	715.5	617	0.039	0.103	0.82	0.82	813.5	-0.84
.1 .1 .1 .1 .1 .1	722.8	618	0.046	0.111	0.80	0.80	848.8	-
.1 .1 .1 .1 .1 .1	743.2	617	0.045	0.104	0.81	0.81	841.2	-1.73[1]
.1 .2 .3 .4 .5 .6	742.3	618	0.044	0.135	0.80	0.80	838.3	-
.1 .2 .3 .4 .5 .6	724.7	617	0.041	0.101	0.82	0.82	822.7	-0.59
.027 .111 .25 .444 .694 1	737.6	618	0.043	0.159	0.79	0.80	833.6	-
. 027 .111 .25 .444 .694 1	724.9	617	0.041	0.104	0.81	0.81	822.9	-0.42
Hierarchical quadratic-linear/quadratic multi-factor model								
1 1 1 1 1 1	730.8	618	0.042	0.185	0.79	0.79	826.8	-
1 1 1 1 1 1	718.5	617	0.040	0.101	0.81	0.82	816.5	-0.61
.1 .1 .1 .1 .1 .1	779.7	618	0.050	0.149	0.77	0.78	875.7	-
.1 .1 .1 .1 .1 .1	741.4	617	0.044	0.104	0.80	0.81	839.4	-2.02[1]
.1 .2 .3 .4 .5 .6	741.2	618	0.044	0.122	0.80	0.81	837.2	-
.1 .2 .3 .4 .5 .6	721.3	617	0.041	0.103	0.82	0.82	827.3	-0.47
.027 .111 .25 .444 .694 1	735.6	618	0.043	0.132	0.80	0.81	831.6	-
.027 .111 .25 .444 .694 1	725.7	617	0.041	0.106	0.81	0.82	823.7	-0.30

Note. χ^2: chi-square, *df*: degree of freedom, RMSEA: root mean square error of approximation, SRMR: standardized root mean square residual, NNFI: nonnormed fit index, CFI: comparative fit index, AIC: Akaike information criterion, $r_{A\text{-}PE}$: (standardized) correlation.

[1] Unlike Pearson correlations estimated correlation can be larger than one.