

## **The confirmatory investigation of APM items with loadings as a function of the position and easiness of items: A two-dimensional model of APM**

KARL SCHWEIZER<sup>1</sup>, MICHAEL SCHREINER<sup>2</sup>, & ANDREAS GOLD<sup>2</sup>

### **Abstract**

The structure of APM is investigated by constraining the loadings of confirmatory factor analysis (CFA) according to item position and item easiness. The constraint of loadings according to an increasing function represents the hypothesis that the item position influences performance in completing APM items. Because of the dependency of item variance on item easiness in binary data this dependency is considered additionally. Several models with one or two latent variables associated with constant and increasing constraints that additionally reflect item easiness were applied to three subsets of APM items. A broad range of item easiness characterized two subsets whereas the range of the remaining subset was rather small. As expected, in the subsets with a broad range the model with two latent variables representing the assumed position effect and dependency did considerably better than the standard CFA model. The superiority of this model suggested that the structure underlying APM is two-dimensional.

Key words: APM, position effect, item easiness, latent curve model, confirmatory factor analysis, fixed-links model.

---

<sup>1</sup> Correspondence should be addressed to Karl Schweizer, PhD, Department of Psychology, Goethe University Frankfurt, Mertonstr. 17, 60054 Frankfurt a. M., Germany; email: k.schweizer@psych.uni-frankfurt.de

<sup>2</sup> Goethe University Frankfurt, Germany

This paper serves two purposes. The first is to investigate the structure of Raven's (1962; Raven, Raven, & Court, 1997) Advanced Progressive Matrices (APM) in considering the position effect on error variance since there is evidence of position-dependent changes of error variance (Hartig, Hölzel, & Moosbrugger, 2007). The item ordering according to item easiness seems to be a crucial precondition for the occurrence of such an effect since it creates similarity between neighboring items. In accordance with corresponding observations Knowles (1988) concluded that "measuring changes the measure".

The second purpose is to advance the method for investigating hypotheses with a high precision by means of confirmatory factor analysis (CFA). Whereas the standard model of confirmatory factor analysis enables the decision on the assignment of items to groups of items associated with latent variables, new models provide the opportunity to investigate curves or even patterns of relatedness, as for example growth curves (Duncan & Duncan, 2004; McArdle, 1986, 1988; McArdle & Epstein, 1987; Meredith & Tisak, 1984, 1990; Schweizer, 2006, 2008). A pattern of relatedness is regarded as a whole that integrates assumed degrees of relatedness between the items serving as manifest variables with respect to the corresponding latent variable. Such a pattern may reflect a hypothesis with a very high precision although it must not provide the basis for a smooth curve.

Complying with the two purposes requires the accommodation of confirmatory factor analysis to the demands of investigating the position effect. Such an investigation means the decomposition of variance into a part that is due to the position effect and another part, and it provides the opportunity to search for sources of this effect since the latent variable representing the position effect can be related to various other latent variables. Another advantage of this approach is that the position effect is investigated within the framework, which has frequently been selected for investigating the dimensionality of APM items, so that the position effect can be evaluated in resorting to familiar concepts and methods.

### *The position effect*

As will be shown in this and the following paragraph, the publications on the position effect suggest that the research was conducted as part of two research traditions. There are factor-analytic investigations of the position effect and investigations by means of IRT models. In the factor-analytic tradition the position effect on error variance was detected in research that focused on the so-called *context effect*. There are several studies showing that item reliability increases as a function of item serial position (e.g., Hamilton & Schuminsky, 1990; Knowles, 1988; Knowles & Byers, 1996). The results of these studies indicate that the responses to the items become increasingly consistent. The increased consistency is attributed to the context of the item. An especially sophisticated study by Hartig, Hölzel and Moosbrugger (2007) reveals that the repeatedly observed increase in item reliability is due to the reduction of error variance. Increasing consistency is associated with decreasing error variance and, consequently, the relationship between true variance and error variance is shifted in the direction of improved reliability. These results suggest that the similarity of items is a precondition for the position effect.

In the IRT tradition the development of the so-called "linear logistic test model" (Fischer, 1972, 1973; Scheiblechner, 1972) provided the starting point for investigating the position effect. Kubinger (1979, 1980; Hohensinn, Kubinger, Reiff, Holocher-Ertl, Khor-

ramdel, & Frebort, 2008) applied this model for the investigation of the position effect in problem solving, mathematical reasoning and in the items of a teaching evaluation form. It was even possible to demonstrate a position effect in Raven's Standard Progressive Matrices by means of this model (Kubinger, Formann, & Farkas, 1991). Furthermore, this model allows the adaptation of the representation of the position effect to various sources, as for example learning and fatigue (for an overview see Kubinger, 2008).

Although the position effect originates from the field of personality research, there is reason for expecting it in achievement data too. Learning can be suspected as the source of such an effect. For example, concomitant transformation processes, as they characterize skill acquisition (Anderson, Fincham, & Douglass, 1997), may give rise to a gradual improvement in performance. Furthermore, increasing familiarity with the small set of rules, which guide the responses to reasoning items (Carpenter, Just & Shell, 1990), can be another possible source of the position effect.

Consequently, it is worthwhile to investigate whether performance in completing APM items shows the position effect. In order to have a formal description of the assumed position effect with a high precision, the definition of a continuous function  $f$  that assigns numbers to the possible positions is proposed. In the absence of special requirements that need to be integrated it is appropriate to define  $f$  as linear function  $f_{\text{linear}}$  such that

$$f_{\text{linear}}(i) = \frac{i}{p} \quad (1)$$

where  $i$  represents the position of the individual items in the ordered sequence of  $p$  item positions. Furthermore, because of the suspected influence of learning on the position effect a modification that predicts acceleration is reasonable. Whereas a very small increase may characterize the first few positions, a steep slope may be achieved in the end. Such a non-linear accelerating function  $f_{\text{accelerating}}$  is given by

$$f_{\text{accelerating}}(i) = \left(\frac{i}{p}\right)^2 \quad (2)$$

### *The investigation of hypotheses with a high precision*

The investigation of hypotheses with a high precision in the framework of confirmatory factor analysis requires the replacement of the standard model, which is known as the congeneric model of measurement (Jöreskog, 1971) and is characteristic of congeneric test theory (Jöreskog, 1971; Lucke, 2005; Raykov, 1997, 2001). Suitability for the investigation of such hypotheses can be achieved by replacing the standard model by the fixed-links model (Schweizer, 2006, 2008, 2009) that is characterized by loadings constrained according to a theory-based pattern of relatedness. This model is especially well suited for this purpose since it is not restricted to the representation of smooth curves.

It is convenient to present the formal model of the covariance matrix as outset of the further considerations since the model of the covariance matrix is basic to all confirmatory models. Assume the  $p \times p$  covariance matrix  $\Sigma$ , the  $p \times q$  matrix of loadings  $\Lambda$ , the  $q \times q$

covariance matrix of latent random variables  $\Phi$ , and the  $p \times p$  diagonal matrix of error variances  $\Theta$ . Based on these definitions the model of the covariance matrix is given by

$$\Sigma = \Lambda\Phi\Lambda' + \Theta \quad (3)$$

Confirmatory factor analysis according to the standard model means that at least a subset of the elements of  $\Lambda$  is estimated whereas the diagonal elements of  $\Phi$  are fixed to one or are otherwise scaled. In contrast, in the fixed-links model all the elements of  $\Lambda$  are constrained whereas the diagonal elements of  $\Phi$  are estimated. The selection of constraints according to the relevant function or the relevant patterns of relatedness is the precondition for investigating a hypothesis by means of the fixed-links model. The estimate of the variance of the corresponding latent variable signifies whether the data provide evidence in favor of the hypothesis. An insignificant variance suggests the rejection of the hypothesis.

The original version of the confirmatory model is restricted to one latent variable (Jöreskog, 1971) so that it is reasonable to concentrate on the case of one latent variable. In selecting the perspective of the  $i$ th manifest random variable ( $i=1, \dots, p$ ) equation (3) reduces to

$$\sigma_i = \lambda_i\phi\lambda_i' + \vartheta_i \quad (4)$$

where  $\sigma_i$  is the variance of the  $i$ th manifest variable,  $\lambda_i$  the loading of the  $i$ th manifest variable on the latent variable,  $\phi$  the variance of the latent variable and  $\vartheta_i$  the error variance of the  $i$ th manifest variable.

Unfortunately, there is a complication of the situation that must be taken into consideration in order to achieve a good model fit: APM items are binary variables that usually show a special characteristic. The variance of items showing a low or high degree of easiness is rather small whereas in the middle of the scale the variance is large. Obviously, there is a specific relationship between easiness and variance in binary data, and it is useful to consider this dependency of variance and easiness in representing hypotheses. The constraints of loadings must reflect both item position and easiness.

The following paragraphs serve the achievement of appropriate constraints. The models including such constraints enable the investigation of the position effect in APM alone and in combination with the dependency of variance on easiness. In the empirical section these models are applied to three subsets of the 36 APM items. Since only two subsets show a broad range according to item easiness, there are different properties that can be compared among each other: the position effect needs to be detected in the subsets with broad ranges but not in the remaining subset with a small range.

### *The representation of the dependency of variance on easiness*

Item easiness is usually considered as the true component of measurement (Carmines & McIver, 1981; Ferrando & Lorenzo-Seva, 2005; Jöreskog, 1971; McDonald, 1999). It characterizes the item and is invariant for all the individuals of the population. The measurement

according to the truncated congeneric model excludes item easiness. Nevertheless, item easiness is important for representing the systematic change of the variance of items.

In order to achieve a representation of the variance according to equation (4) including item easiness, it is assumed that the loadings of the  $i$ th manifest variable ( $i=1, \dots, p$ )  $\lambda_i$  is describable by a continuous function  $s$  of the corresponding item easiness  $v_i$  such that

$$\lambda_i = s(v_i) . \quad (5)$$

Because of the concentration of this paper on APM items that are binary variables (coded as “0” if incorrect and “1” otherwise) it is convenient to define easiness with respect to the binary variables  $y_i$  as the probability of a correct response:

$$v_i = p(y_i = 1) . \quad (6)$$

Furthermore, in binary data the variance of  $y_i$  is given by

$$\text{var}(y_i) = p(y_i = 1)p(y_i = 0) . \quad (7)$$

In order to have a first definition of the relationship of easiness and loading, the function  $s$  is set equal to the standard deviation of the binary variable so that

$$s(v_i) = \sqrt{v_i(1-v_i)} \quad (8)$$

where the probabilities of equation (7) are replaced according to equation (6). The reasonableness of the previously presented assumptions and definitions becomes obvious in considering the limits for true and error variances at the end of the next paragraph.

The replacement of the loadings in equation (4) leads to

$$\sigma_i = s(v_i)\phi s(v_i)' + \vartheta_i \quad (9)$$

in the first step (see equation (5)) and to

$$\sigma_i = \sqrt{v_i(1-v_i)}\phi\sqrt{v_i(1-v_i)}' + \vartheta_i \quad (10)$$

in the second step (see equation (8)). It is apparent that equation (10) can be true if  $\phi \leq 1$ . Otherwise the variance of the manifest variable is overestimated. For  $\vartheta_i=0$  and  $\phi=1$  it is true. Furthermore, it is important that  $\phi$  is appropriate for all the items ( $i=1, \dots, p$ ).

### *The representation of the position effect*

At this point the discussion concerning the representation of the position effect can be resumed. According to the arguments of the previous section it may not be sufficient to integrate the functions according to equations (1) and (2) into fixed-links models without any further modification. Since there is reason to assume that valid results can only be expected when the dependence of variance on easiness is considered additionally, a multiplicative relationship (●) is installed between the functions reflecting position effect and dependency. The constraints of the loadings are defined as

$$\lambda_i = f_{\text{linear}}(i) \bullet s(v_i) \quad (11)$$

and

$$\lambda_i = f_{\text{accelerating}}(i) \bullet s(v_i) . \quad (12)$$

The functions  $f_{\text{linear}}$ ,  $f_{\text{accelerating}}$  and  $s$  vary between 0 and 1. The corresponding equations for the variance of the manifest variable  $y_i$  are given by

$$\sigma_i = \frac{i}{p} \sqrt{v_i(1-v_i)} \phi \left( \frac{i}{p} \sqrt{v_i(1-v_i)} \right) + \vartheta_i \quad (13)$$

and

$$\sigma_i = \left( \frac{i}{p} \right)^2 \sqrt{v_i(1-v_i)} \left[ \left( \frac{i}{p} \right)^2 \sqrt{v_i(1-v_i)} \right] + \vartheta_i \quad (14)$$

in corresponding order.

### *The separated representation of effects due to item position and standard processing*

The formulas of the previous section for predicting the variances of the manifest variables show a crucial disadvantage: they predict very small variances for the first items of a multi-item measure like APM although even the responses to the first items are the result of complex information processing. Furthermore, it is unrealistic to assume that the position effect applies to the whole of information processing. Consequently, it is reasonable to assume some kind of standard processing that is independent of the position effect and, therefore, needs a separated representation.

For achieving appropriate representations of both the position effect and the effect of standard processing it is necessary to consider an additional latent variable. However, the original version of the congeneric model (Jöreskog, 1971) includes one latent variable only. It is the item factor analysis model (e.g., Bock & Aitkin, 1981; Ferrando, 2005; Gibbons, Bock, Hedeker, Weiss, Segawa, Bhaumik, Kupfer, Frank, Grochocinski, & Stover, 2007) that provides an extension of the original versions of the congeneric model for two latent

variables. Equation (15) gives a formal model of the  $i$ th manifest random variable  $y_i$  ( $i=1, \dots, p$ ) including two latent variables:

$$y_i = \lambda_{i1}\xi_1 + \lambda_{i2}\xi_2 + \delta_i \quad (15)$$

where  $\lambda_{i1}\xi_1$  and  $\lambda_{i2}\xi_2$  represent the true components of measurement and  $\delta_i$  the error component. Each true component is defined as the product of a loading  $\lambda$  and a latent score with respect to the attribute of interest  $\xi$ .

The separated representation of the effects of item position and standard processing in accordance with the formal structure of equation (15) can be achieved by specifying the loading of the first summand according to equations (11) and (12) and the loading of the second summand as weighted constant. The constant representing standard processing needs a weight that reflects variation because of the dependency of variance on easiness. The following equations give the adaptations of the formal model for representing both effects:

$$y_i = [f_{\text{linear}}(i) \bullet s(v_i)] \xi_{\text{position}} + s(v_i) \xi_{\text{constant}} + \delta_i \quad (16)$$

and

$$y_i = [f_{\text{accelerating}}(i) \bullet s(v_i)] \xi_{\text{position}} + s(v_i) \xi_{\text{constant}} + \delta_i \quad (17)$$

where the latent scores are identified by subscripts that relates them to the position effect (“position”) and the constant of standard processing (“constant”).

The corresponding equations for the variance of the manifest variable are

$$\sigma_i = \frac{i}{p} \sqrt{v_i(1-v_i)} \phi_{\text{position}} \left( \frac{i}{p} \sqrt{v_i(1-v_i)} \right)' + \sqrt{v_i(1-v_i)} \phi_{\text{constant}} \sqrt{v_i(1-v_i)}' + \vartheta_i \quad (18)$$

and

$$\sigma_i = \left( \frac{i}{p} \right)^2 \sqrt{v_i(1-v_i)} \phi_{\text{position}} \left[ \left( \frac{i}{p} \right)^2 \sqrt{v_i(1-v_i)} \right]' + \sqrt{v_i(1-v_i)} \phi_{\text{constant}} \sqrt{v_i(1-v_i)}' + \vartheta_i \quad (19)$$

Since  $i$  and  $p$  are normally given and an estimate of  $v_i$  can be obtained independent of the other parameters, only  $\phi_{\text{position}}$  and  $\phi_{\text{constant}}$  need to be estimated in addition to  $\vartheta_i$  ( $i=1, \dots, p$ ).

### The application to APM items

The items of APM (Raven, 1962; Raven, Raven, & Court, 1997) are especially well suited for an investigation of the suspected position effect and the dependence of variance on easiness in the framework of confirmatory factor analysis since APM items show a very broad range of item easiness. Furthermore, there is the characteristic arrangement of items

that can be expected to favor the occurrence of the position effect: the APM items are arranged in such a way that the degree of item easiness decreases from the first to last items.

*The models for item analysis*

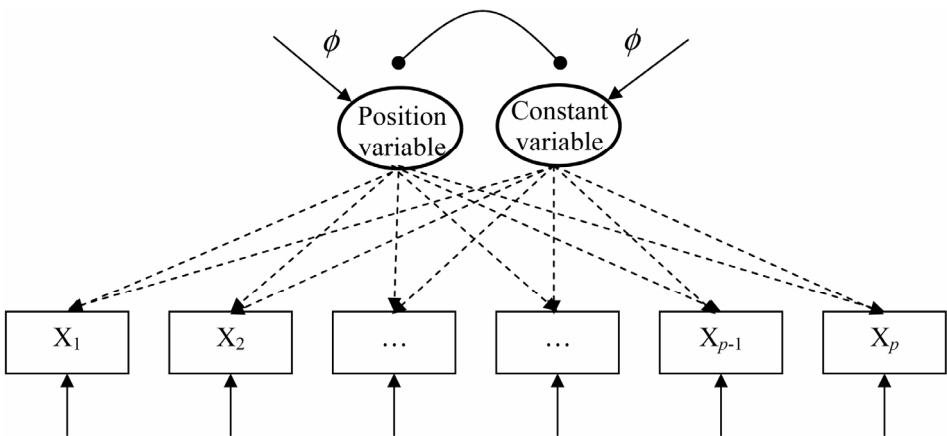
The following models were considered: (1) the *standard model*. This model included one latent variable and  $p$  manifest variables. All the loadings were estimated whereas the variance of the latent variable was set to one. (2) The *position model*. Each one of the two versions of this model included one latent variable and  $p$  manifest variables. The constraints for the loadings of this model were constructed according to equations (11) and (12):

$$\lambda_i = \frac{i}{p} \sqrt{v_i(1-v_i)} \quad \text{resp.} \quad \lambda_i = \left(\frac{i}{p}\right)^2 \sqrt{v_i(1-v_i)} .$$

(3) The *constant model*. One latent variable and  $p$  manifest variables were included in this model. The constraints for the loadings of this model were taken from equation (10):

$$\lambda_i = \sqrt{v_i(1-v_i)} . \tag{20}$$

(4) The *position-constant model*. Two uncorrelated latent variables and  $p$  manifest variables characterized this model. There were two versions of the position-constant model because of the two functions suggested for the position effect. Figure 1 provides an illustration of the position-constant model.



**Figure 1:**

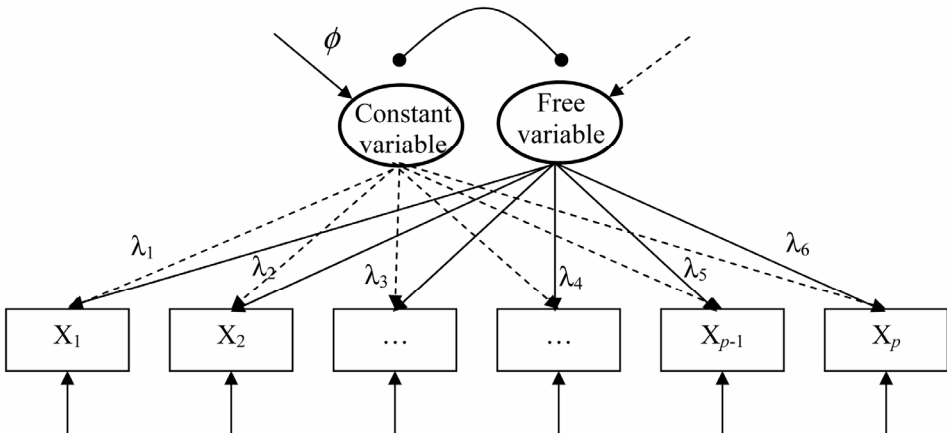
Illustration of the fixed-links model with two latent variables representing effects due to item position and standard processing



There are two ellipses representing the latent variables associated with the position effect and the constant of standard processing. The arrows with dashed shafts represent constrained parameters whereas solid shafts signify free parameters. The bent line with two small discs as ends indicates that the correlation between the latent variables is set to zero.

In order to evaluate the results, a *mixed model* was considered additionally. The mixed model included one latent variable with loadings constrained according to the constant of standard processing that was represented according to equation (20). Additionally, the mixed model comprised a latent variable with free loadings. The free loadings were of special interest since in the case of a position effect the free loadings should reflect this effect. An illustration of the mixed model is given by Figure 2.

The latent variable according to the standard model is identified as “free variable” and the latent variable according to the fixed-links model as “constant variable”. The difference between the estimated and constrained loadings is indicated by solid and dashed shafts of arrows in corresponding order.



**Figure 2:**

Illustration of the mixed model with a standard part and a fixed part (The fixed part represents the effect due to standard processing)

### Items and data

Special sets of items were constructed for the investigation. The items originate from Raven’s (Raven, Raven, & Court, 1997) Advanced Progressive Matrices (APM) that are considered as a measure of fluid intelligence, of reasoning and, as it was proposed recently, of fluid reasoning (McGrew, 2005). The reasons for constructing sets of items were the establishment of specific properties and the achievement of manageable set sizes.

The 36 items of Set II of APM were subdivided into three subsets of 12 items each after applying the items to the participants in keeping to the original ordering. Strong position effects were expected because of the original ordering since the distances between the item

positions were large. In contrast, the distances between the items of a subset would be rather small if the subsets of items were administered separately. The items of the *first* subset were selected with respect to the following properties: a high degree of homogeneity and the coverage of a broad range of item easiness. It is denoted *consistent-complete* item set in the results section. The *second* subset of items was constructed for having an item set with a low degree of homogeneity and a good coverage of the range of item easiness. This subset is denoted *inconsistent-complete* item set in the following section. The remaining items were included into the *third* subset of items. Since there was no further opportunity for eliminating items, it was not possible to assign a favorable property to this item set. Instead, it could be expected to show unfavorable properties only and is denoted *inconsistent-incomplete* item set.

Data were available from 324 university students. In collecting the data the items of Set I of APM were used for giving instructions and making the nature of APM items obvious. Set II of APM was administered according to the instructions of the manual. Correct responses were coded as ones and incorrect responses as zeros. Missing responses were considered as incorrect responses and coded accordingly.

## Results

Arithmetic means and standard deviations were computed for all the items of Set II. Table 1 provides the results of this investigation separately for the three item sets.

**Table 1:**

Arithmetic Means and Standard Deviations for the Items of the Three Item Sets (Consistent-complete Item Set, Inconsistent-complete Item Set, Inconsistent-incomplete Item Set) (N=324)

Consistent-complete Set			Inconsistent-complete Set			Inconsistent-incomplete Set		
No. of item	Mean	SD	No. of item	Mean	SD	No. of item	Mean	SD
3	.95	.22	7	.95	.21	1	.95	.21
18	.76	.43	13	.72	.45	2	.95	.22
22	.70	.46	14	.88	.33	4	.94	.24
23	.69	.46	17	.77	.42	5	.91	.28
24	.58	.49	19	.73	.45	6	.96	.20
26	.60	.49	20	.69	.46	7	.93	.26
28	.41	.49	21	.78	.41	9	.99	.10
30	.40	.49	25	.56	.50	10	.92	.28
32	.33	.47	27	.57	.50	11	.96	.19
33	.37	.48	29	.34	.47	12	.92	.27
35	.33	.47	31	.54	.50	15	.89	.31
36	.10	.29	34	.41	.49	16	.92	.27

This Table includes three parts. The first part is arranged on the left, the second part in the middle and the third part on the right. The first part includes the results for the items of the consistent-complete item set. The arithmetic means of the items varied between .10 and .95 and the standard deviations between .22 and .49. A Cronbach's alpha of .69 was observed for this item set. The results for the items of the inconsistent-complete item set are presented in the middle part. The arithmetic means of the items varied between .34 and .95. For this item set a Cronbach's alpha of .59 was found. The third part provides the results for the set of inconsistent-incomplete items. The arithmetic means were between .89 and .99 and the Cronbach's alpha .45. Apparently, the consistent-complete item set showed the broadest range of arithmetic means and the most favorable consistency according to Cronbach's alpha. The consistency was impaired in the inconsistent-complete item set, and the coverage of the possible range was not as good as the coverage by the consistent-complete item set. The least favorable characteristics were observed for the inconsistent-incomplete item set. The adjustment of the consistency coefficients according to Spearman-Brown led to .87, .81 and .71 for the consistent-complete, the inconsistent-complete and the inconsistent-incomplete item sets in corresponding order. Since a consistency coefficient of about .80 is usually reported for this measure, only the inconsistent-incomplete item set showed a real impairment.

*The results achieved for the consistent-complete item set*

The investigations of the three item sets were performed by means of LISREL (Jöreskog & Sörbom, 2001). The fit results obtained for the consistent-complete item set with respect to  $\chi^2$ ,  $\chi^2/df$ , RMSEA, GFI, CFI, NNFI and AIC are included in Table 2.

The fit results indicated a good or acceptable fit for all the models with one exception. The best model fit characterized the "accelerating" version of the position-constant model with two latent variables. Each one of the variances of the latent variables of this model

**Table 2:**

Fit Statistics of the Models Obtained for the Consistent-complete Item Set Derived from Advanced Progressive Matrices (N=324)

Type of model	$\chi^2$	df	$\chi^2/df$	RMSEA	GFI	CFI	NNFI	AIC
Standard	96.32	54	1.78	.049	.95	.91	.90	144.32
Position								
- linear	133.27	65	2.05	.057	.94	.83	.83	159.27
- accelerating	238.12	65	3.66	.091	.89	.61	.61	264.12
Constant	90.55	65	1.39	.035	.96	.93	.93	116.55
Position-constant								
- linear	66.15	64	1.03	.010	.97	1.00	1.00	94.15
- accelerating	62.73	64	0.98	.000	.97	1.00	1.01	90.73

reached the level of significance (position:  $\sigma=0.22$ ,  $t=3.97$ ,  $p<.05$ ; constant:  $\sigma=0.13$ ,  $t=7.15$ ,  $p<.05$ ). The results for the “linear” version of the position-constant model differed only slightly from the results for the “accelerating” version. The  $\chi^2$ s and AICs observed for these models were considerably lower than the  $\chi^2$ s and AICs observed for all the other models. It was very interesting to find more favorable results for the position-constant model than for the standard model. The worst results were found for the versions of the position model. The results for the “accelerating” version of the position model were not even acceptable. All in all, the consideration of item position, item easiness and standard processing yielded very favorable results whereas the concentration on either item position or standard processing proved considerably less appropriate.

Support for the conclusion of the previous paragraph was additionally provided by the results achieved for the mixed model. This model showed a good degree of fit ( $\chi^2(53)=52.40$ ,  $\chi^2/df=0.99$ ,  $RMSEA=.000$ ,  $GFI=.97$ ,  $CFI=1.00$ ,  $NNFI =1.01$  and  $AIC=102.40$ ). The model fit for this model is only slightly inferior to the model fit of the best model. The first to third columns of Table 3 include the standardized constraints, standardized loadings and standardized error variances.

Whereas the original constraints differed considerably from each other, they showed hardly any variation after standardization. The estimated loadings increased in absolute size from top to bottom with a few exceptions. Only the lower half of loadings reached the level of significance. Apparently, the increase of the standardized loadings was similar to the increase expected because of the position effect.

**Table 3:**

Completely Standardized Constrained and Free Loadings and Error Variances for the Mixed Model Obtained for the Consistent-complete Item Set (N=324)

No. of item	Constrained loading on “constant” latent variable	Free loading on “free” latent variable	Error variance
3	0.37	0.19	0.83
18	0.38	0.11	0.85
22	0.38	-0.03	0.85
23	0.38	0.01	0.86
24	0.38	0.01	0.86
26	0.37	-0.08	0.85
28	0.38	-0.18*	0.82
30	0.37	-0.19*	0.82
32	0.38	-0.36*	0.73
33	0.37	-0.32*	0.76
35	0.37	-0.41*	0.69
36	0.37	-0.28*	0.79

\* Significant at the 5%-level

*The results achieved for the inconsistent-complete item set*

The investigation of the inconsistent-incomplete item set with the models adapted to the characteristics of this item set gave rise to the results reported in Table 4.

**Table 4:**  
Fit Statistics of the Models Obtained for the Inconsistent-complete Item Set Derived from  
Advanced Progressive Matrices (N=324)

Type of model	$\chi^2$	df	$\chi^2/df$	RMSEA	GFI	CFI	NNFI	AIC
Standard	82.22	54	1.52	.040	.96	.88	.85	130.52
Position								
- linear	97.24	65	1.50	.039	.95	.86	.86	123.24
- accelerating	131.23	65	2.02	.056	.94	.72	.72	157.23
Constant	104.51	65	1.61	.043	.95	.82	.82	130.51
Position-constant								
- linear	74.36	64	1.16	.022	.96	.95	.95	102.36
- accelerating	65.03	64	1.01	.007	.97	.99	.99	93.03

One or a few good or acceptable results were achieved for each one of the models. However, several models showed an unacceptable degree of fit according to CFI and NNFI: the standard model, both versions of the position model and the constant model. Only, the versions of the position-constant model yielded an overall good model fit. The results of the “accelerating” version were more favorable than the results of the “linear” version. Each one of the variances of the latent variable of the best model reached the level of significance (position:  $\sigma=0.25$ ,  $t=4.61$ ,  $p<.05$ ; constant:  $\sigma=0.08$ ,  $t=5.33$ ,  $p<.05$ ). Furthermore, it needs to be emphasized that also in the inconsistent-complete item set considerably better results were observed for the versions of the position-constant model than the standard model.

Further support for this model was provided by the results achieved for the mixed model. This model showed a good degree of fit ( $\chi^2(53)=49.76$ ,  $\chi^2/df=0.94$ ,  $RMSEA=.000$ ,  $GFI=.97$ ,  $CFI=1.00$ ,  $NNFI =1.01$  and  $AIC=99.76$ ). It was marginally inferior to the “accelerating” version of the position-constant model according to AIC and slightly better than the “linear” version. Table 5 provides the standardized constraints, standardized loadings and standardized error variances.

The standardized constraints of the first column show a minimum of variation only whereas the original constraints differed considerably from each other. The estimated loadings of the second column showed an increase in absolute size from top to bottom with a few exceptions. The loadings for the 7th, 13th and 14th items deviated from the general trend. However, only the first loading reached the level of significance besides the loadings of 27th, 29th, 31th and 34th items. The latter loadings suggested that there was again the expected increase from top to bottom due to the position effect.

**Table 5:**  
Completely Standardized Constrained and Free Loadings and Error Variances for the Mixed Model Obtained for the Inconsistent-complete Item Set (N=324)

No. of item	Constrained loading on "constant" latent variable	Free loading on "free" latent variable	Error variance
7	0.32	0.25*	0.83
13	0.32	0.13	0.88
14	0.32	0.15	0.87
17	0.32	0.05	0.89
19	0.32	0.02	0.89
20	0.32	0.08	0.89
21	0.33	0.02	0.89
25	0.32	0.09	0.89
27	0.33	0.21*	0.85
29	0.32	0.32*	0.79
31	0.32	0.46*	0.69
34	0.32	0.42*	0.72

\* Significant at the 5%-level

*The results achieved for the inconsistent-incomplete item set*

The models adapted to the characteristics of the inconsistent-incomplete item set yielded the results of Table 6.

**Table 6:**  
Fit Statistics of the Models Obtained for the Inconsistent-incomplete Item Set Derived from Advanced Progressive Matrices (N=324)

Type of model	$\chi^2$	df	$\chi^2/df$	RMSEA	GFI	CFI	NNFI	AIC
Standard	92.71	54	1.72	.047	.95	.69	.62	140.71
Position								
- linear	125.83	65	1.93	.054	.95	1.00	1.42	135.83
- accelerating	157.48	65	2.42	.066	.93	1.00	1.42	183.48
Constant	109.98	65	1.69	.046	.96	1.00	1.42	135.98
Position-constant								
- linear	107.83	64	1.68	.046	.96	1.00	1.42	135.83
- accelerating	109.11	64	1.70	.047	.96	1.00	1.42	137.11

Good or acceptable fit results were achieved for  $\chi^2/df$ , RMSEA and GFI whereas all the NNFI results were unacceptable. In the standard model the CFI result was additionally bad. The overall comparison of the fit results indicated the best outcomes for the constant and position-constant models. Since the worst results were observed for the position model, the further investigation of the item set by means of the mixed model was omitted. Because of the minor differences between the standard and position-constant models the results could not really be considered as evidence in favor of the position effect.

## Conclusions

The most spectacular finding of this study is that APM shows to be two-dimensional instead of one-dimensional. There is the dimension of standard processing that may reflect reasoning, as it was investigated in the seminal study by Carpenter, Just and Sell (1990). Furthermore, there is the position dimension that may represent learning. The ascription to learning is supported by the observation that in both the consistent-complete and inconsistent-complete item sets the “acceleration” version does a bit better than the “linear” version. These results provide an account for the observation that APM performance is a predictor of school achievement (Brody, 1997).

A position effect was also detected in Raven’s SPM (Kubinger, Formann, & Farkas, 1991). This time the linear logistic test model of the IRT framework enabled the detection of the effect. Therefore, the question arises, what is the reason for the similarity of results. We argue that it is the similarity of the basic models. Both types of models, the congeneric model of confirmatory factor analysis and the linear logistic test model, are members of the family of generalized linear item response models (Bartholomew, 2002; Bartholomew & Knott, 1999; Mellenbergh, 1994; Moustaki & Knott, 2000). Furthermore, both basic models are linear and are composed of several components/parameters. This similarity of the basic models probably outweighs the difference according to the estimation methods.

After reading the results section it should also be obvious to the reader that investigating hypotheses with a high degree of precision, as it is achieved by constraining all the loadings, does not necessarily mean a low degree of model fit. Precise hypotheses can lead to good model fit in spite of the many constraints. The hypotheses investigated in this paper are transformed into functions that assure a high precision. The functions representing the position effect give rise to a curve that fits into the framework of the latent curve approach (Duncan & Duncan, 2004; Meredith & Tisak, 1990). However, this function is not really appropriate for investigating variances and covariances since variances and covariances show dependency on item easiness. The modification of the original functions because of the consideration of item easiness leads to patterns of relatedness, as they can be investigated by fixed-links models (Schweizer, 2006, 2007, 2008, 2009; Stankov & Schweizer, 2007). Although such a pattern may be describable by a product of functions, it does normally not give rise to a smooth curve. Patterns of relatedness must simply be reasonable from the perspective of the hypothesis of interest.

## References

- Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 932-945.
- Bartholomew, D. J. (2002). Old and new approaches to latent variable modeling. In G. A. Marcoulides & I. Moustaki (Eds.), *Latent variable and latent structure models* (pp. 1-13). Mahwah, NJ: Erlbaum.
- Bartholomew, D. J., & Knott, M. (1999). *Latent variable models and factor analysis* (2nd ed.). London: Arnold.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: application of the EM algorithm. *Psychometrika*, *46*, 443-459.
- Brody, N. (1997). Intelligence, schooling and society. *American Psychologist*, *52*, 1046-1050.
- Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobservable variables: Analysis of covariance structures. In G. W. Bohrnstedt & E. F. Borgatta (Eds.), *Social measurement* (pp. 65-115). Beverly Hills, CA: Sage.
- 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*, 404-431.
- Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, *35*, 333-363.
- Ferrando, P. J. (2005). Factor analytic procedures for assessing social desirability in binary items. *Multivariate Behavioral Research*, *40*, 331-349.
- Ferrando, P. J., & Lorenzo-Seva, U. (2005). IRT-related factor analytic procedures for testing the equivalence of paper-and-pencil and internet-administered questionnaires. *Psychological Methods*, *10*, 193-205.
- Fischer, G. H. (1972). Conditional maximum likelihood estimations of item parameters for a linear logistic test model. *Research Bulletin*, *9*, Psychological Institute University of Vienna, Vienna.
- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, *37*, 359-374.
- Gibbons, D. G., Bock, D., Hedeker, D., Weiss, D. J., Segawa, E., Bhaumik, D. K., Kupfer, D. J., Frank, E., Grochocinski, V. J., & Stover, A. (2007). Full-information item bifactor analysis of graded response data. *Applied Psychological Measurement*, *31*, 4-19.
- 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*, 1301-1307.
- 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*, 157-183.
- Hohensinn, C., Kubinger, K. D., Reiff, M., Holoher-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*, 391-402.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, *36*, 109-133.
- Jöreskog, K.G. & Sörbom, D. (2001). *LISREL* (Version 8.5). Computer Software. Chicago: Scientific Software International.



- Knowles, E. S. (1988). Item context effects on personality scales: measuring changes the measure. *Journal of Personality and Social Psychology*, *55*, 312-320.
- 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*, 1080-1090.
- Kubinger, K. D. (1979). Das Problemlöseverhalten bei der statistischen Auswertung psychologischer Experimente. Ein Beispiel hochschuldidaktischer Forschung [Problem-solving behavior in the case of statistical analyses of psychological experiments. An example of research on universities didactics]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *26*, 467-495.
- Kubinger, K. D. Neue Ergebnisse [The evaluation of effectiveness of university lecturing with the help of the linear logistic test model by Fischer. New results]. *Archiv für Psychologie*, *133*, 69-79.
- 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*, 311-327.
- Kubinger, K. D., Formann, A. K., & Farkas, M. G. (1991). Psychometric shortcomings of Raven's Standard Progressive Matrices SPM in particular for computerized testing. *European Review of Applied Psychology*, *41*, 295-300.
- Lucke, J. F., (2005). The alpha and the omega of congeneric test theory: an extension of reliability and internal consistency of heterogeneous tests. *Applied Psychological Measurement*, *29*, 65-81.
- McArdle, J. J., (1986). Latent variable growth within behavior genetic models. *Behavior Genetics*, *16*, 163-200.
- McArdle, J. J., (1988). Growth curve analysis in contemporary psychological research. In J. R. Nesselrode & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed.) (pp. 561-614). New York: Plenum.
- McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, *58*, 110-133.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll Theory of cognitive abilities: past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: theories, tests, and issues* (pp. 136-181). New York, NY: Guilford Press.
- Mellenbergh, G. J. (1994). Generalized linear item response theory. *Psychological Bulletin*, *115*, 300-307.
- Meredith, W., & Tisak, J. (1984, June). *Tuckerizing curves*. Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107-122.
- Moustaki, I., & Knott, M. (2000). Generalized latent trait models. *Psychometrika*, *65*, 391-411.
- Raven, J. C. (1962). *Advanced progressive matrices*. London: Lewis & Co. Ltd.
- Raven, J. C., Raven, J., & Court, J. H. (1997). *Raven's Progressive Matrices and Vocabulary Scales*. Edinburgh: J. C. Raven Ltd.
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with congeneric components. *Multivariate Behavioral Research*, *32*, 329-353.
- Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*, *54*, 69-76.

- Scheiblechner, H. (1972). Das Lernen und Lösen komplexer Denkaufgaben [Learning and solving complex cognitive problems]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *19*, 476-505.
- Schweizer, K. (2006). The fixed-links model in combination with the polynomial function as tool for investigating choice reaction time data. *Structural Equation Modeling*, *13*, 403-419.
- Schweizer, K. (2007). Investigating the Relationship of Working Memory Tasks and Fluid Intelligence Tests by Means of the Fixed-Links Model in Considering the Impurity Problem. *Intelligence*, *35*, 591-604.
- Schweizer, K. (2008). Investigating experimental effects within the framework of structural equation modeling: an example with effects on both error scores and reaction times. *Structural Equation Modeling*, *15*, 327-345.
- Schweizer, K. (2009). Fixed-links models for investigating experimental effects combined with processing strategies in repeated measures designs: a cognitive task as example. *British Journal of Mathematical and Statistical Psychology*, DOI:10.1348/000711007X268558.
- Stankov, L., & Schweizer, K. (2007). Raven's Progressive Matrices, manipulations of complexity and measures of accuracy, speed and confidence: An integrative re-analysis. *Psychology Science*, *49*, 326-342.