A new way to present factor structure: The Factor Based Radex Visualization

Michael Dantlgraber, Stefan Stieger, & Ulf-Dietrich Reips

University of Konstanz

Author Note

Michael Dantlgraber, Department of Psychology, University of Konstanz; Stefan Stieger, Department of Psychology, University of Konstanz; Ulf-Dietrich Reips, Department of Psychology, University of Konstanz.

Correspondence concerning this article should be addressed to Michael Dantlgraber, University of Konstanz, Department of Psychology, Box 31, 78457 Konstanz, Germany
E-mail: michael.dantlgraber@uni-konstanz.de

Abstract

The visualization of latent structures plays an important role in many areas of social sciences. We present a new way of visualizing factor structure. Our visualization uses circles of different size to illustrate higher order and lower order factors. Because of its implementation of circles the model appears similar to Venn diagrams. However, in Venn diagrams manifest equality (R^2 = 1) is visualized with completely covered circles having the same size. In our visualization, latent equality of factors is visualized with circles having the same center but possibly having different sizes depending on their lower order factors. The focus on the center of the circles is based on Guttmans original Radex Model. Therefore, our proposed visualization is called Factor Based Radex Visualization.

Factor Based Radex Visualization was primary developed to locate task groups of intelligence tests and to illustrate their factor structure. In this article we first describe this original implementation. We then use Factor Based Radex Visualization to compare three different scales that measure positive self-concepts (self-esteem and mental toughness) with the aid of a large sample. We show how to build a network of several tests and constructs and how analyzing Factor Based Radex Visualization facilitates appropriate test selection in psychological research and practice.

Keywords: data visualization, structural equation model, network analysis, bifactor model, radex model

Introduction

In applied settings psychologists often face the challenge of selecting the right test suited to their specific questions. For example, if they want to use a personality questionnaire or an intelligence test, they have to choose between several well constructed tests that all deliver reliable results. They may, however, focus on different aspects of the same psychological construct.

Usually factor analysis and structural equation modeling (SEM; Wright, 1921; Kaplan, 2000) are used to compare various tests. Factor analysis explains test results by identifying a structure of underlying, i.e. latent, dimensions. By using SEM the structure of a single test can be modeled and compared to the strucuture of other tests. However, results and visualizations of SEM are complex and often difficult to interpret.

In this article we want to discuss the advantages of a new way to visualize factor structur. The new visualization was primary developed to locate tasks groups of intelligence tests and to illustrate their factor structure. For this reason we describe its original deriviation first. Then, we apply it to a large dataset containing three different questionnaires that measure positive-self concepts (self-esteem and mental toughness).

Introducing the new visualization

Figure 1 shows a traditional visualization of a test comparison. The factor structure of the Modular Short Intelligence Kit (M-KIT: Dantlgraber, 2015) is compared with the factor structure of the Intelligence Structure Test 2000 R (I-S-T 2000 R: Liepmann, Beauducel, Brocke, & Amthauer, 2007).

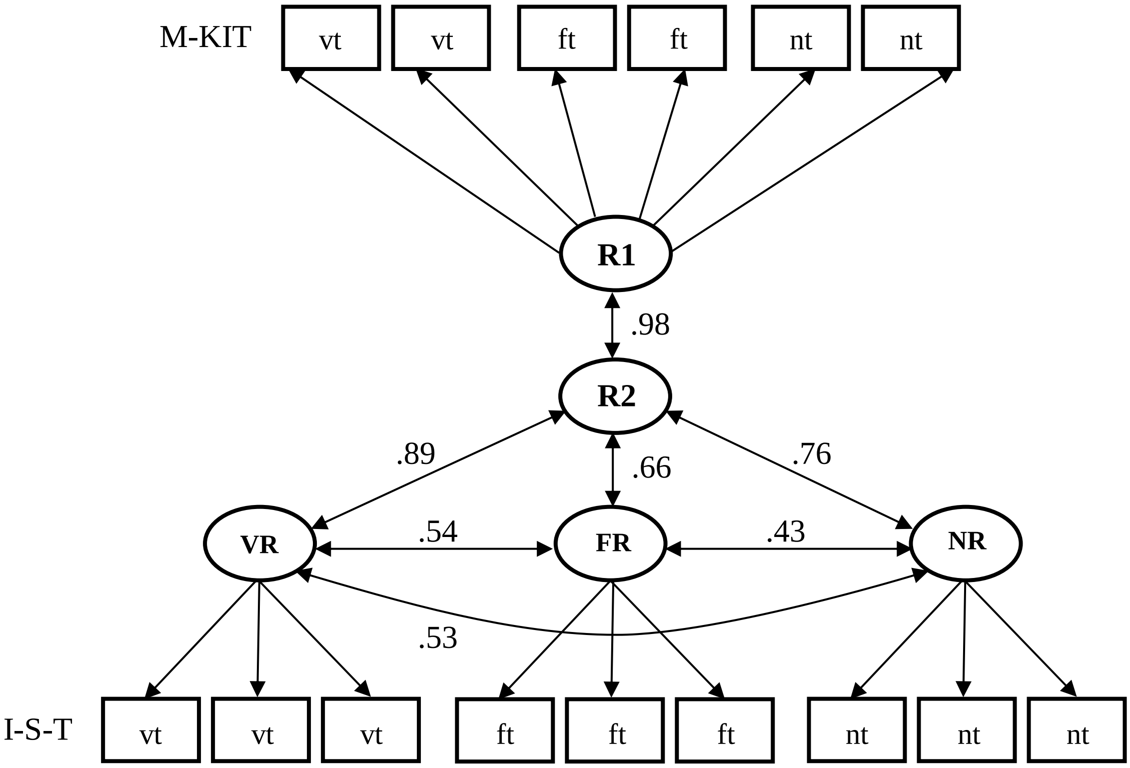

Figure 1. Traditional Visualization of factor structure. I-S-T = Intelligence Structure Test 2000 R; M-KIT = Modular Short Intelligence Kit; vt = verbal task group, ft = figurative task group, nt = numerical task group. R = fluid reasoning, VR = verbal reasoning, FR = figurative reasoning, NR = numerical reasoning.

Both tests consist of verbal reasoning tasks, figurative reasoning tasks and numerical reasoning tasks. That means that all tasks require *reasoning ability*, and additionally some specific verbal, figurative or numerical abilities. As indicated by the authors the I-S-T 2000 R is modeled with three correlated factors (dimensions) representing *verbal reasoning* (VR), *figurative reasoning* (FR) and *numerical reasoning* (NR). The M-KIT also consists of verbal, figurative and numerical tasks, but the reasoning part of these tasks is substantially higher than the reasoning part of the I-S-T 2000 R tasks. For this reason the M-KIT is indicated to be modeled with only one reasoning factor (R1). As expected, R1 as measured by the M-KIT is almost identical with the reasoning (higher order) factor R2 (r = 0.98) estimated by the all I-S-T 2000 R tasks. That means that the main source of variance of both test results can be explained by reasoning ability.

If two tests share the same main source of variance they are often viewed as two tests measuring the same construct, i. e. both tests are called *reasoning tests*. In practice the use of this common terminology has some problematic implications. Psychologists may think that there is no content difference between tests that are viewed to measure the same construct (sharing the same main source of variance). However, each test defines the measured construct with its specific operational definition, i.e. a specific weighting or interaction of main and additional sources of variance. If this is ignored and reliability is the sole consideration in the test selection process, this results in a misdirected use of psychological tests . We propose a more clearly arranged visualization that emphasizes the commonalities and differences between tests in a better way than traditional methods. In addition, the new visualization can be extended more easily and also be used for task group locations.

Figure 2 shows the same structure presented in Figure 1 illustrated with the new visualization. It can be interpreted as the traditional visualization (see Figure 1) without showing the tasks groups and viewed from above, i.e., looking down the vertical dimension from the point of the higher order factor R1.

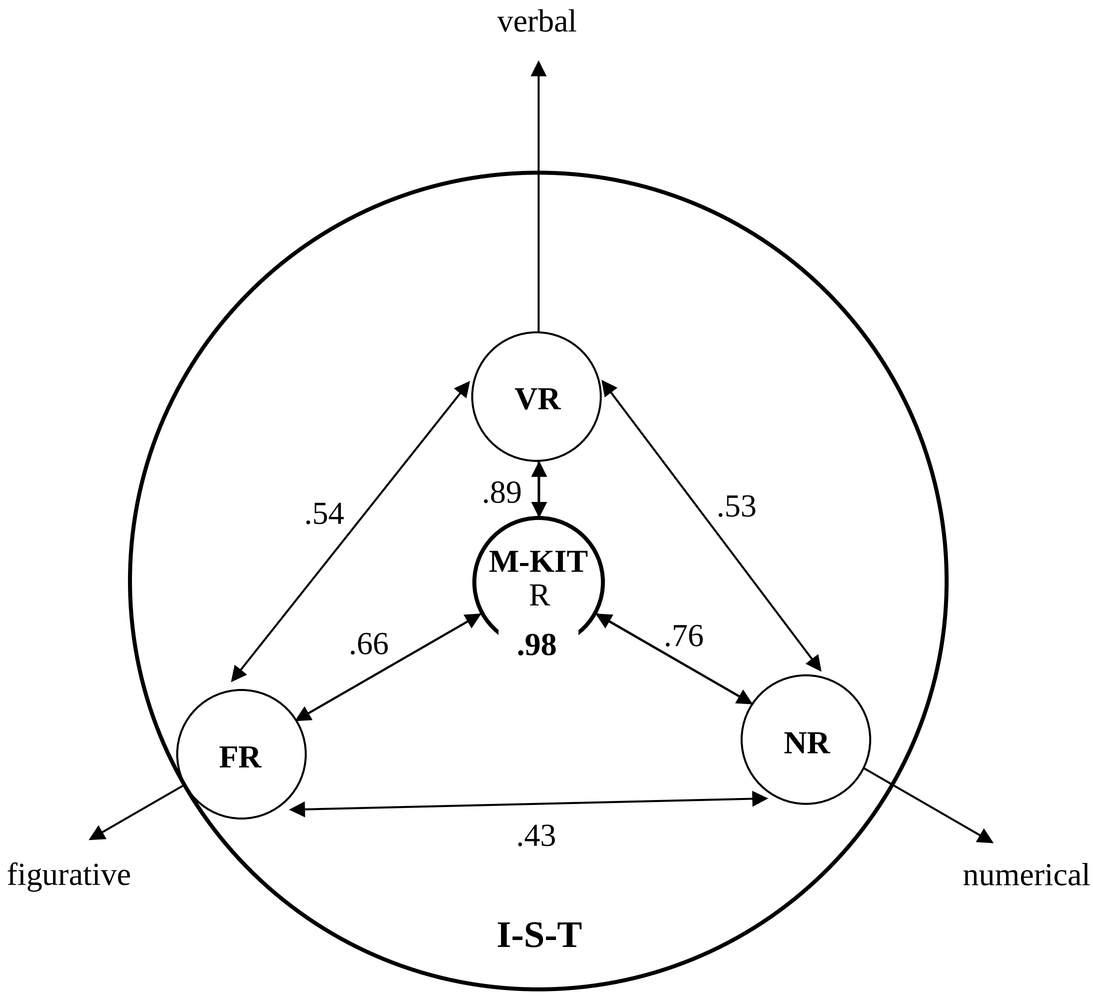

Figure 2. New visualization of factor structure. I-S-T = Intelligence Structure Test 2000 R; M-KIT = Modular Short Intelligence Kit; R = fluid reasoning; VR = verbal reasoning; FR = figurative reasoning; NR = numerical reasoning.

In the new visualization each circle represents a specific pool of task groups. The center of a circle represents its general factor, i.e. its main source of variance. Consequently, task group pools that share the same general factor are visualized with circles having the same center. Circles having the same center can possibly have different sizes to illustrate differences regarding lower order factors. This is also shown in Figure 2.

Both tests (task group pools) are visualized with circles having the same center. However, the circle representing the I-S-T 2000 R is larger than the circle representing the M-KIT, thereby taking into account its internal structure of three correlated factors (VR, FR, NR). The larger circle of the I-S-T 2000 R represents a comparatively greater weighting of verbal, figurative and numerical abilities and consequently a comparatively lower weighting of reasoning ability when using the overall test score. The new visualization emphasizes that the I-S-T 2000 R is a broader defined reasoning test than the M-KIT by using different sized circles.

Furthermore, the new visualization can be interpreted like a radar chart that orders task group pools along several dimensions. In Figure 2 the task group pool VR has a shorter distance to the center of the radar chart (representing R) than the task group pool FR. With regard to content that means that VR requires less specific verbal abilities than FR requires specific figurative abilities. VR is more similar to R than FR and consequently correlates somewhat higher with R than FR. In the following sections the possibilities of the new visualization are described in detail.

Naming the visualization

In the contex of psychological research using circles to visualize data is strongly assoziated with Venn diagrams. In Venn diagrams equality of manifest variables (R^2 = 1) is visualized with completely covered circles having the same size. In our visualization equality of latent factors (factor correlation = 1) is visualized with circles having the same center but possibly having different sizes depending on their additional sources of variance (their lower order factors). The idea of using a reference point when ordering task groups is based on Guttman’s (1954) Radex model. Guttman ordered task groups of intelligence test by using multidimensional scaling.

The link between multidimensional scaling and factor analysis was described by Marshalek, Lohman, and Snow (1983). They found that task groups located in the center of the radex estimated by multidimensional scaling tend to have higher loadings on the higher order factor (g). This suggests the possibility to order task groups along the results of factor analytic studies. The higher order factor (g) is equated with *fluid reasoning* (Gf; McGrew, 2005, 2009), referred to as R in this article. Because our visualization is inspired by Guttmans initial idea, it is called Factor Based Radex Visualization (FBRV).

Derivation of the factor structure

In this section we derive the factor structure of the M-KIT and the I-S-T 2000 R. In the following section we describe the creation of the FBRV presented in Figure 2.

The M-KIT was constructed to measure *fluid reasoning* (fluid intelligence [Gf]; McGrew, 2005, 2009) and consists of two verbal, two figurative, and two numerical task groups. One verbal, one figurative, and one numerical task group consist of (deductive) *sequential reasoning tasks* (Carroll, 1993). The other three task are called *non-sequential* tasks. This term includes *inductive* or *quantitative reasoning* tasks (Carroll, 1993) and tasks that also have substantial factor loadings but do not match these definitions (e.g., non-inductive figurative tasks that partly measure visual abilities like *perceptual speed* or *mental rotation*). When using SEM, the M-KIT shows a single factor structure (n = 1,054) with proper model fit (Comparative Fit Index [CFI = .99], Tucker-Lewis Index [TLI = .99], root mean square error of approximation [$RMSEA$ = .02], and standardized root mean square residual [$SRMR$ = .03]).

The I-S-T 2000 R is one of the most frequently used intelligence tests in German-speaking countries and also contains a reasoning subscale. It measures verbal, numerical, and figurative reasoning and consists of three task groups for each category. All these task groups are *non-sequential*. The reasoning module of the I-S-T 2000 R is indicated to be modelled by three correlated factors, one verbal, one numerical, and one figurative reasoning factor. Liepmann et al. (2007) confirmed this structure in their validation data (n = 2,208).

To assess the construct validity of the M-KIT, 145 participants of the whole validation sample completed the M-KIT and the reasoning module of the I-S-T 2000 R on the same day with a 45 minute break. The order of both tests was counterbalanced. Dantlgraber (2015) found a high correlation between both tests on the manifest level (r = .78) and a near perfect correlation on the latent level (r = .98), if one single reasoning factor (R) was extracted for each test (see Figure 3).

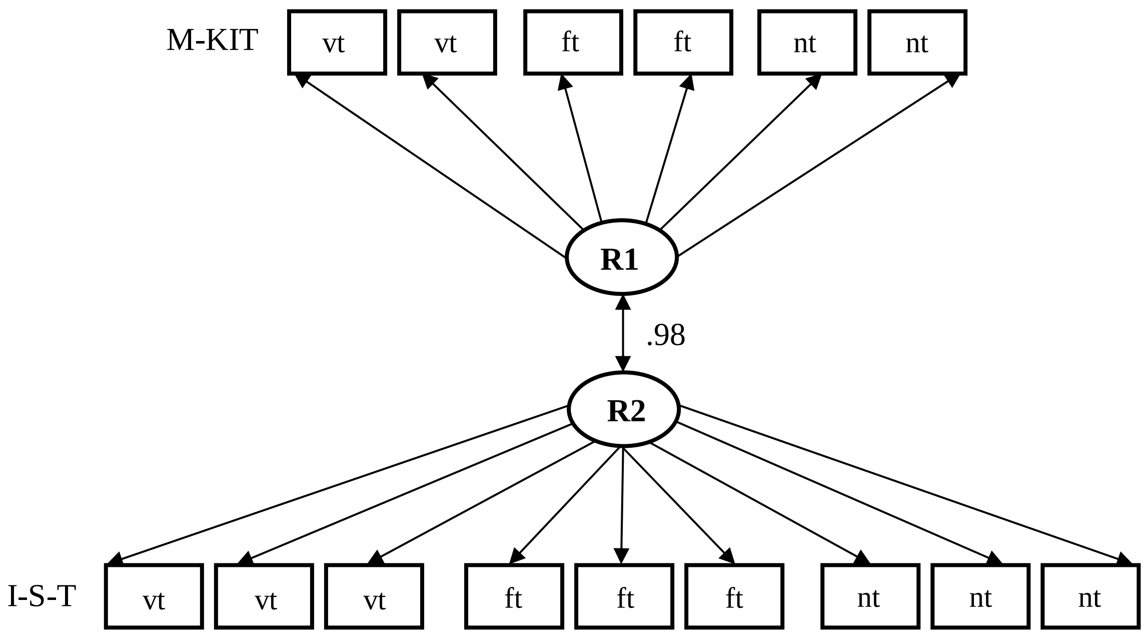

Figure 3. Latent correlation of the reasoning factors of both tests; I-S-T = Intelligence Structure Test 2000 R; M-KIT = Modular Short Intelligence Kit; R = fluid reasoning; vt = verbal task group, ft = figurative task group, nt = numerical task group.

In Figure 3 all task groups of both tests are integrated in one model. But this model extracts only one factor from the I-S-T 2000 R tasks and for this reason this model does not show a good fit (*CFI* = .84; *TLI* = .81; *RMSEA* = .10; *SRMR* = .07). In order to find a model with sufficiently high fit, the general factor of the I-S-T 2000 R needs to be separated into three correlated factors as indicated by the authors. Such a model improves the fit (*CFI* = .97; *TLI* = .96; *RMSEA* = .04; *SRMR* = .05) substantially, it is shown in Figure 4.

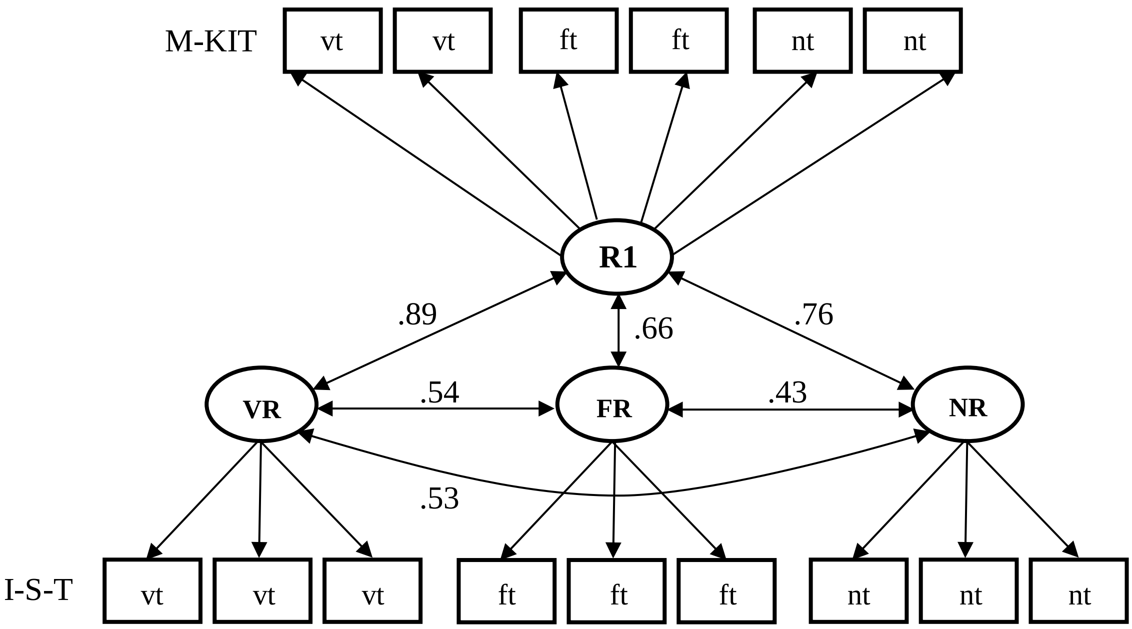

Figure 4. Visualization of the well-fitting four factor model. I-S-T = Intelligence Structure Test 2000 R; M-KIT = Modular Short Intelligence Kit; R = fluid reasoning; vt = verbal task group, ft = figurative task group, nt = numerical task group. VR = verbal reasoning; FR = figurative reasoning, NR = numerical reasoning.

Figure 4 is theory based and shows a good fit, therefore it could be recommend as a visualization in a published article. However, it does not show the valuable information of Figure 3, i.e. that both general factors R1 and R2 are almost identical. By taking this information into account and treating the general factors R1 and R2 as identical (*r* = .98) one can interpret the latent correlations of the factors VR, FR, NR with R1 (Figure 4) as latent correlations to their own higher order factor R2 (previously shown in Figure 1). In order to understand the relationship of both tests, Figure 1 is needed. It clarifies that both tests indeed require the same ability (*reasoning*), but differentiate regarding their additional required abilities.

The overall score of the M-KIT (Modular Short Intelligence Kit) mainly represents reasoning ability. For this reason results regarding external criteria or balanced sex differences remain similar when using various short forms. The overall score of the I-S-T 2000 R represents more verbal, figurative and numerical abilities than the overall score of the M-KIT. The I-S-T 2000 R defines the construct broader than the M-KIT and thus cannot be reduced to various short forms without departing from the main factor.

Note that the verified additional abilities of the I-S-T 2000 R and the M-KIT are not considered if researchers estimate latent parameters of one latent reasoning dimension. This is definitely possible, but using the broadly defined I-S-T 2000 R solely to estimate *reasoning ability* would be a very inefficient test use.

Creating a Factor Based Radex Visualization

The main idea of a FBRV is the comparison of a general factor model and a correlated factor model by using SEM, i. e. a comparison of an overall task group pool with its split into some smaller task group pools. FBRV illustrates in a radar chart how the main sources of variance vary between different task group pools. In this section we describe the creation of the FBRV presented in Figure 2.

First, in the general factor model a single factor is extracted of an overall task group pool including all fifteen task groups of both tests. In the correlated factor model three factors are extracted of three different task group pools including all verbal, numerical, and figurative task groups of both tests (see Table 1).

Table 1

Comparison of factor loadings of a general and a correlated factor model

Content	*Test*	*Task category*	*General factor model*	*Correlated factor model*	*Ratios of squared loadings*	*Center distance*
Verbal	I-S-T	nsq	.44	.50	1.27	0.27
		nsq	.66	.72	1.19	0.19
		nsq	.60	.65	1.17	0.17
	M-KIT	nsq	.57	.62	1.19	0.19
		sq	.74	.75	1.03	0.03
Numerical	I-S-T	nsq	.58	.72	1.56	0.56
		nsq	.56	.66	1.40	0.40
		nsq	.65	.76	1.39	0.39
	M-KIT	nsq	.58	.60	1.08	0.08
		sq	.72	.68	0.91	-0.09*
Figurative	I-S-T	nsq	.49	.55	1.29	0.29
		nsq	.46	.58	1.60	0.60
		nsq	.38	.48	1.58	0.58
	M-KIT	nsq	.59	.63	1.12	0.12
		sq	.64	.69	1.15	0.15

Note. I-S-T = Intelligence Structure Tests 2000 R; M-KIT = Modular Short Intelligence Kit; q = sequential reasoning task; nsq = non-sequential reasoning task. Ratios of squared loadings are based on four-digit factor loadings. *Center distances visualize the increase of factor loadings. Negative estimates are treated as 0 in a FBRV.

Table 1 shows that the factor loadings of the correlated factor model tend to be higher than the loadings of the general model. This is due to the correlated factor model taking the specific verbal, numerical, and figurative sources of variance into account. Table 1 also shows that the loadings of the I-S-T 2000 R task groups increase more than the loadings of the M-KIT task groups when switching from the general to the correlated factor model. This signifies that the task groups of the I-S-T 2000 R share more specific variance than the task groups of the M-KIT and explains why these two tests should be modeled differently.

To quantify the increase of the loadings, i.e. the specificity of the task groups, we use the ratios of the squared loadings for each task group. In order to calculate the ratios of squared loadings we divide the squared loadings of the correlated factor model by the squared loadings of the general factor model (see Table 1).

A ratio of squared loadings equal to one signifies that the estimated loadings of the respective task group are identical in the correlated and in the general factor model. Furthermore, it would mean that the respective task group does not have more variance in common with the smaller and more specific task group pool than with the overall task group pool. In other words, a ratio of squared loadings equal to one represents no difference. Therefore we take the ratios of the squared loadings for each task group and subtract 1 to represent a missing difference with 0. These values are called *center distances* (see Table 1).

In a FBRV the center of the radar chart represents the general factor of the overall task group pool. Center distances represent the comparatively association of each task group to its smaller and more specific task group pool. Center distances are used for locating the task groups in a radar chart (see Figure 5).

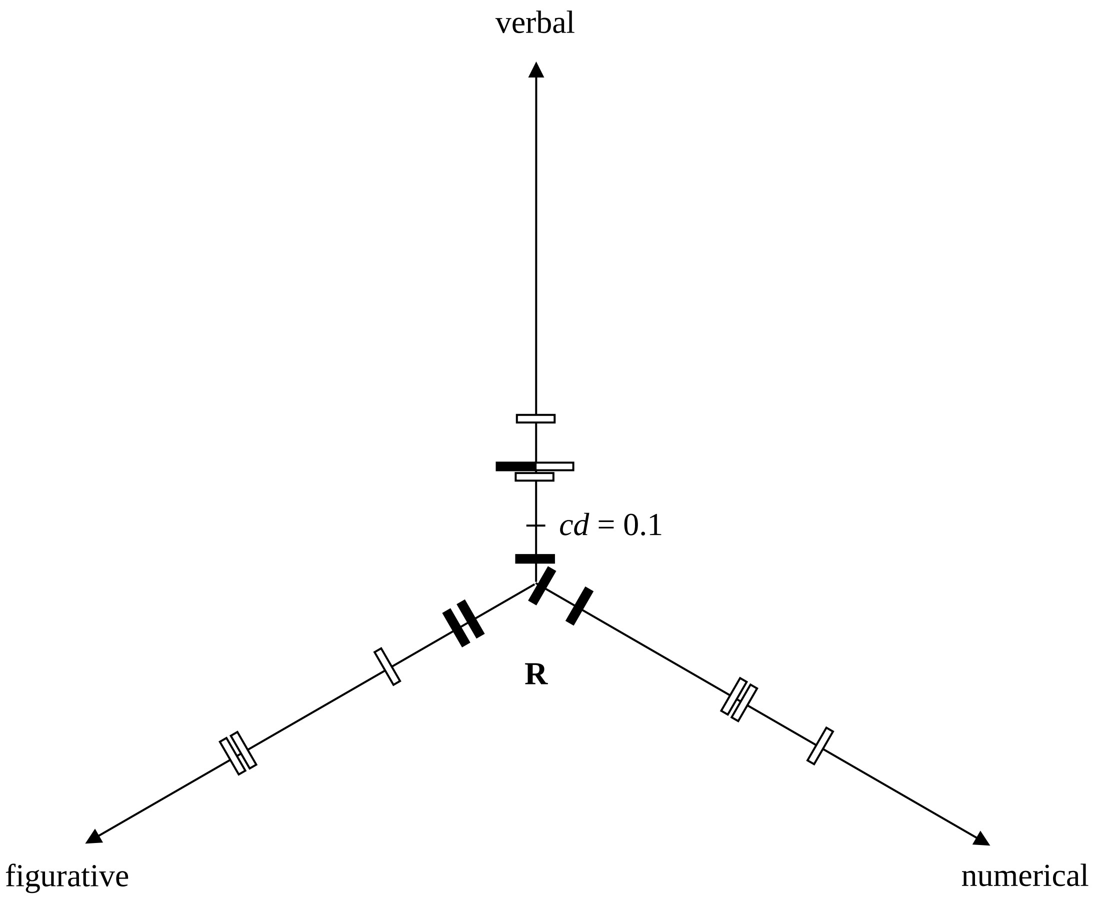

Figure 5. Task group location in a Factor Based Radex Visualization. M-KIT's task groups are drawn as black boxes and the I-S-T 2000 R's task groups as white boxes. R = fluid reasoning ability; cd = center distance.

In accordance with the traditional visualization manifest variables of the task groups are illustrated as boxes. M-KIT's task groups are drawn as black boxes and the I-S-T 2000 R's task groups as white boxes. Figure 5 shows that M-KIT's task groups are more centered than I-S-T 2000 R's task group. It thereby visualizes why both tests should be modeled differently although both tests consist of verbal, figurative and numerical task groups.

Note that the sequential (deductive) numerical task group of the M-KIT is slightly better represented by the general reasoning factor than by the numerical reasoning factor (see Table 1). Center distances only visualize the increase of factor loadings. Negative estimates are not illustrated in a FBRV and treated as 0. Thefore the numerical task group is centered in Figure 5. FBRV offers two possibilities to combine task groups to task groups pools: shifted circles or centered circles. Both options are described in the following.

Shifted Circles

Task groups that are located on the same dimension can be combined to a task group pool and be visualized by a shifted circle. The circle representing the task group pool is located by using the *mean center distance* of the included task groups. The mean center distance is illustrated as the distance from the circle edge to the center of the radar chart. In Figure 6 the verbal, figurative and numerical task groups of the I-S-T 2000 R are combined to three task group pools (VR, FR, NR), visualized by shifted circles. The latent correlations of these pools are also included.

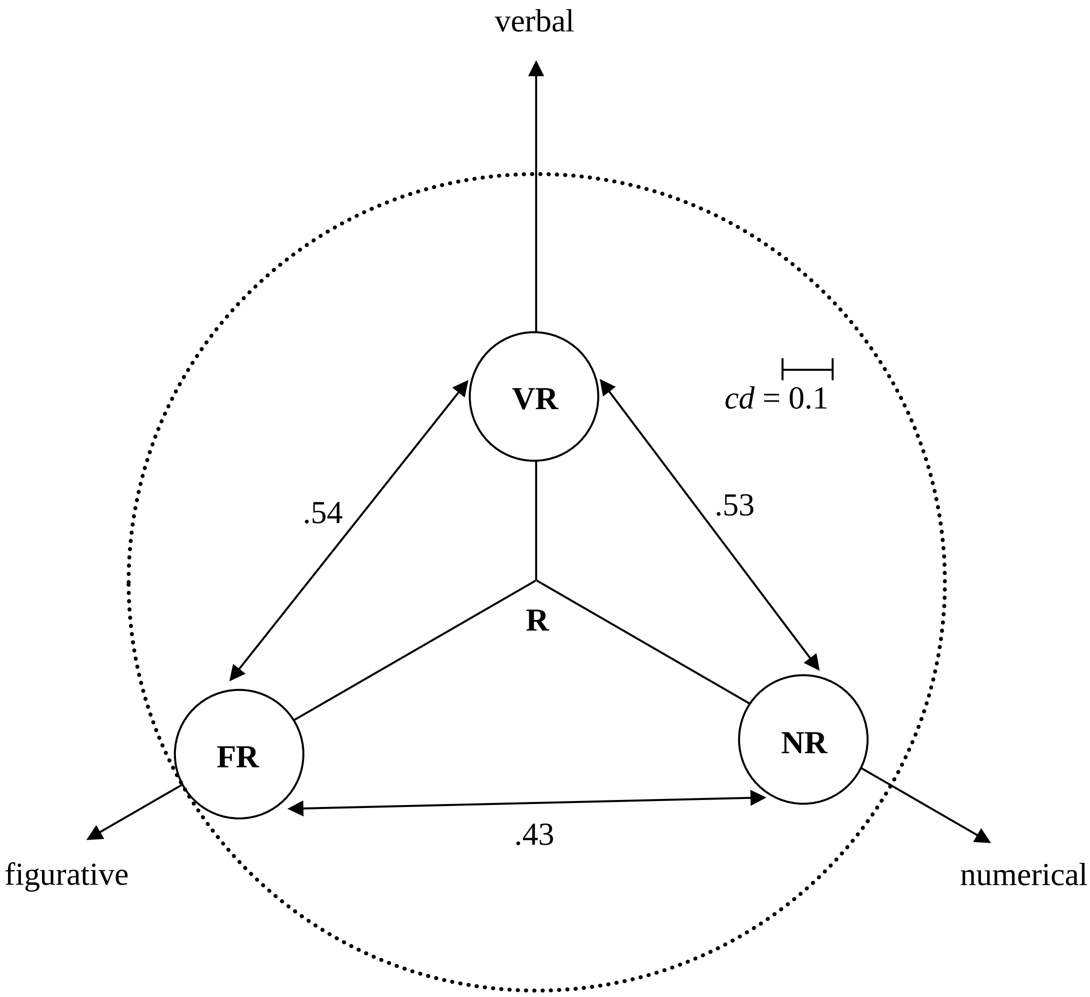

Figure 6. Location of shifted circles in Factor Based Radex Visualization. R = fluid reasoning; VR = verbal reasoning; FR = figurative reasoning; NR = numerical reasoning; cd = center distance.

Centered Circles

The center of the radar chart represents the general factor of all included task groups. All included task groups can be combined to an overall task group pool and visualized by one centered circle. If the overall task group pool can be split into two or more smaller task group pools representing equal factors (equal main sources of variance), the smaller pools can also be visualized by centered circles. As a cutoff criterium to determine factor equality we use a latent correlation of $r > .95$ between the smaller pools.

If the overall task group pool is split into the two original test pools (M-KIT / I-S-T 2000 R), the original test pools show a correlation of $r = .98$ and therefore can be visualized with centered circles. Figure 2 results from adding centered circles and additional latent correlations to Figure 6.

Requirements for FBRV

FBRV has two requirements. (1) The data should allow SEM. (2) All estimated factor loadings should be positive to ensure the interpretation of the squared loadings. However, if all items of the same pool are negative, it is possible to recode them and rename the pool (see Procedure section). Center distances can be negative but are treated as 0 (see Table 1 and Figure 5).

Comparing scales with FBRV

In the sections above we showed how to visualize the structure of two psychometric tests having similar theories. Both intelligence tests shared an identical general factor (*reasoning ability*) and all task groups of both tests could be arranged along the same three dimensions (verbal, figurative, numerical).

When comparing similar scales they often do not have comparable theories. They often do not share an identical general factor and vary with regard to the number and the kind of their facets. For this reason we show how to visualize the structure of scales having correlated but not identical general factors and differing in their lower order factors. We use a large dataset containing three different scales that measure positive-self concepts, namely the Domain Specific Self-esteem Inventory (DSSEI: Hoyle, 1991), the Rosenberg Self-esteem

Scale (RSES: Rosenberg, 1965), and the Sports Mental Toughness Questionnaire (SMTQ: Sheard, Golby, & van Wersch, 2009).

Method

Sample

The sample was drawn in German speaking countries (mainly Austria and Germany) and consists of 2,272 participants who filled out the DSSEI, the RSES, and the SMTQ. The sample constitutes a community-based sample and is age-stratified. 1,265 (56%) were women, 980 (43%) were men, 27 (1%) did not state their sex. Participants' mean reported age was 39.8 years (*SD* = 17.7). Regarding highest educational level, 547 (24%) had a university degree, 783 (34%) had matura (high school graduation), 512 (23%) had an apprenticeship certificate, 333 (15%) had a mandatory degree, and 86 (4%) had no degree.

Procedure

First, the items of the scales were trated as the task groups before. Second, we separately estimate three FBRVs for the DSSEI, the SMTQ, and the RSES. The FBRV of each test was calculated as described above. First, a general factor model and a correlated factor model were estimated for each test by using SEM. In the next step ratios of squared loadings and center distances of the items were calculated and then a mean center distance for each facet.

Within the DSSEI and the SMTQ, items are grouped to facets as indicated by the original authors of the scales. Within the RSES, items are also grouped to two subscales as suggested in the past by some researchers (positive vs. negative self-esteem; for a recent discussion, see Supple, Su, Plunkett, Peterson, & Bush, 2013). In order to meet the requirements of a FBRV, all items of the facet *negative self-esteem* were recoded and the facet was renamed to *lack of negative self-esteem*.

In a FBRV all item pools that are not indicated to be further divided are drawn by small circles having the same size. In this example, this regards the facets of the three tests. The circles are arranged along the dimensions of a radar chart using mean center distances. Mean center distances are illustrated as distances from the cicle edges to the center of the radar chart. The absolut circle size of the smallest circles and the scaling are defined by authors. All facets (item pools) of a single test are represented by a larger circle. The size of the larger circle is adjusted to the facet that is furthest from the center. In the end, the circles are named and arrows representing latent correlations are added. Figure 7 shows a FBRV for each test.

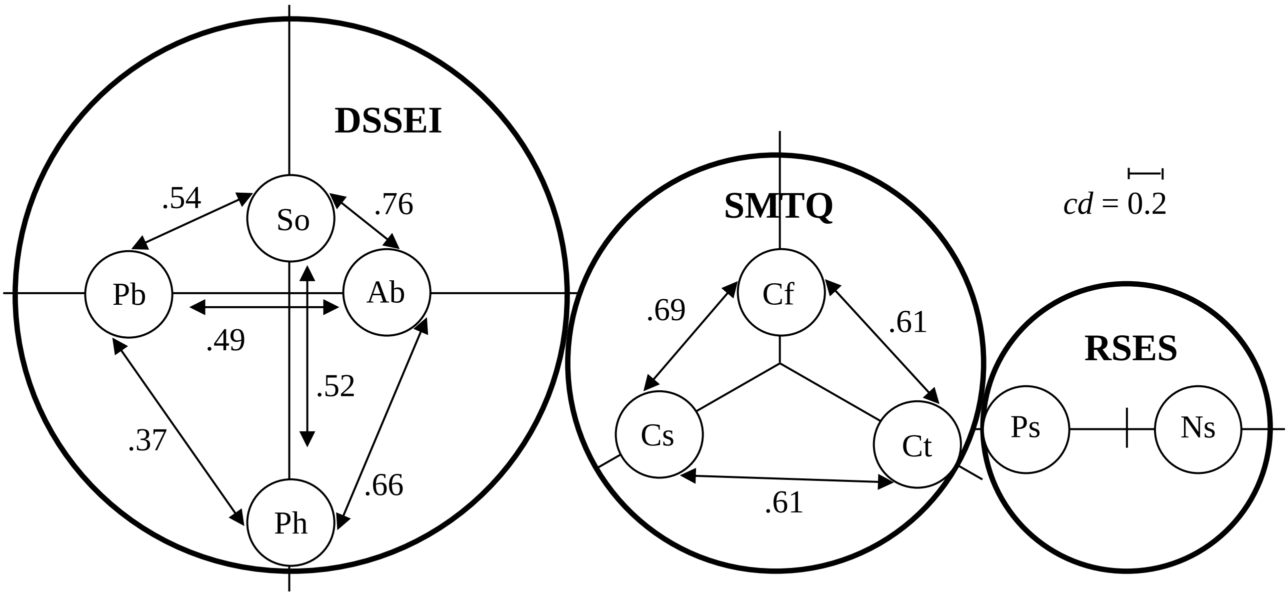

Figure 7. Factor Based Radex Visualizations of the Domain Specific Self-Esteem Inventory (DSSEI), the Sports Mental Toughness Questionnaire (SMTQ), and the Rosenberg Self-Esteem Scale (RSES). So = Social Competence; Pb = Public Presentation; Ab = Task-Related

Abilities; Ph = Physical Appeal; Cf = Confidence; Cs = Constancy; Ct = Control; Ps = Positive self-esteem; Ns = Lack of Negative self-esteem; cd = center distance.

In the section above a single FBRV was estimated for each test (see Figure 7). Now we interpret each test as a facet of a superordinate FBRV. The calculation steps are the same as previously described. The loadings of a superordinate factor model including all items of all three tests are compared to the loadings of three correlated factors representing the general factors of the DSSEI, the SMTQ, and the RSES. Next, ratios of squared loadings and center distances were calculated for each item (see Table 1). Then a mean center distance was calculated for each test and used for locating the tests in a superordinate radar chart.

Results

Figure 8 shows the result of the comparison of the DSSEI, the SMTQ, and the RSES using FBRV. Note that the center distances within the three tests are drawn on the same scale, but the center distances of the superordinate model are twice that in order to avoid overlapping circles.

The item "*In general, I feel confident about my abilities*" of the DSSEI facet *Ability* and the item "*I feel confident about my social behavior*" of the DSSEI facet *Social* have the highest loadings on the superordinate general factor including all items of the three tests (r = .66, r = .62). For this reason we named the superordinate factor *self-confidence*.

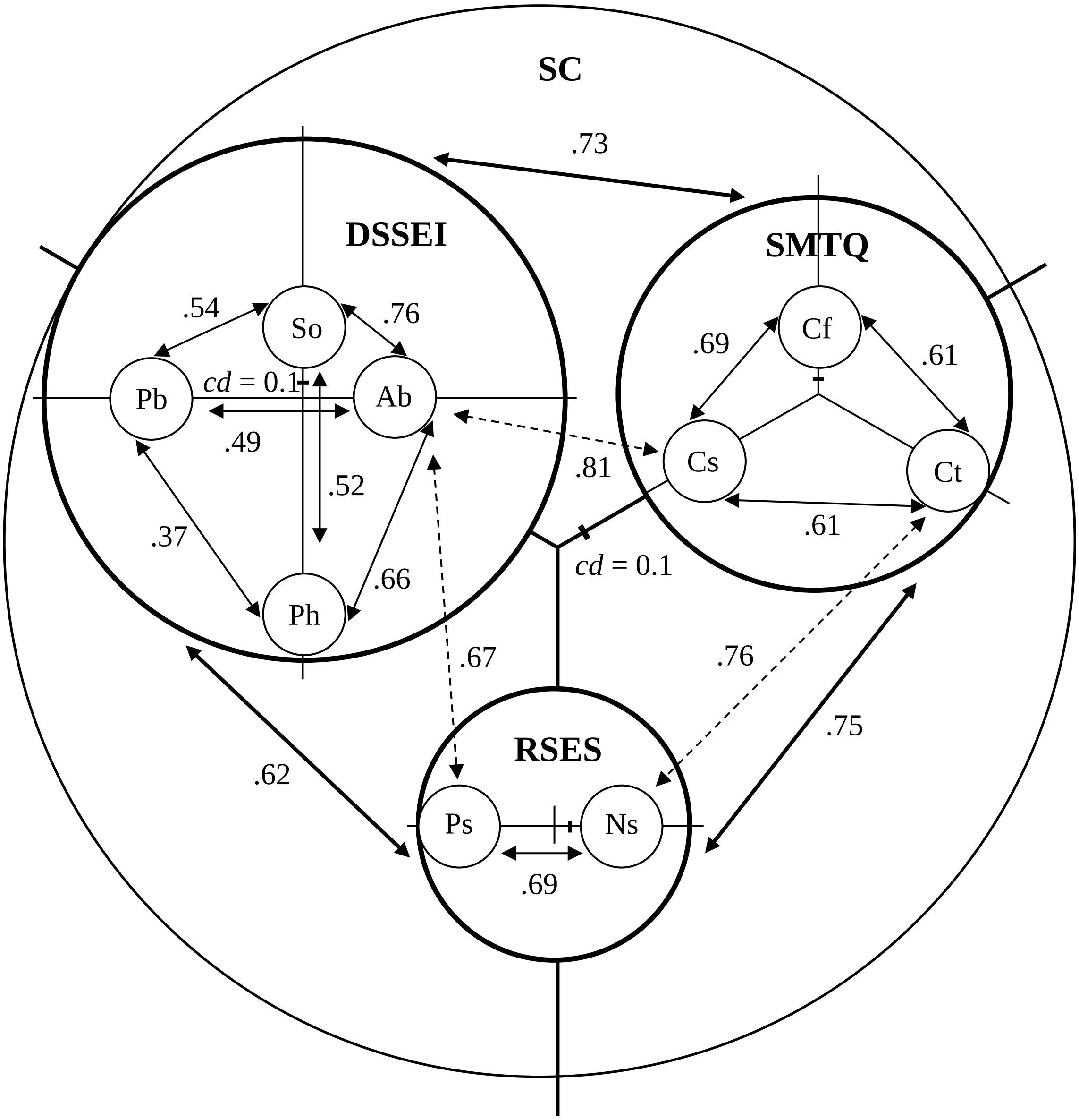

Figure 8. Connected Factor Based Radex Visualization of the Domain Specific Self-Esteem Inventory (DSSEI), the Sports Mental Toughness Questionnaire (SMTQ), and the Rosenberg Self-Esteem Scale (RSES). So = Social Competence; Pb = Public Presentation; Ab = Task-Related Abilities; Ph = Physical Appeal; Cf = Confidence; Cs = Constancy; Ct = Control; Ps = Positive self-esteem; Ns = Lack of Negative self-esteem; cd = center distance.

Figure 8 integrates three radar charts in one superordinate radar chart. As every radar chart FBRV is interpreted by analyzing distances along the dimensions. We recommend to begin the interpretation of an FBRV from the center of the superordinate radar chart (*self-confidence*, Figure 8). The center of the superordinate radar chart has the shortest distance to the edge of the overall circle of the DSSEI meaning the superordinate factor *(self-confidence)* is best represented by the general factor of the DSSEI.

Next, the facet level is interpreted. The center of the DSSEI has the shortest distance to its facets *Social Competence* and *Task-Related Abilities* meaning these facets represent the general factor of the DSSEI best. Furthermore, as described above, the general factor of the DSSEI represents the superordinate factor (*self-confidence*) best. This implies that the facets *Social Competence* and *Task-Related Abilities* are strongly associated with self-confidence.

In some cases facets between different tests are correlated stronger than their general factors. This can be viewed as an indicator that the additional facet dimensions of the different tests have something in common and are stretching in a similar direction. Facet correlations between different tests that are higher than their general factor correlations are drawn with dotted arrows in a FBRV. Figure 8 shows that the additional dimension *Task-Related Abilities* of the DSSEI has something in common with the additional dimensions *constancy* and *positive self-esteem* of the SMTQ and the RSES.

Interpretation of general factors

Using FBRV facilitates the interpretation of general factors. In Figure 8 the general factor of the RSES (representing *self-esteem*) shows a lower correlation to the general factor of the DSSEI (also representing *self-esteem*) than to the general factor of the SMTQ (representing *mental toughness*). This is surely not an expectable result, because both, the RSES and the DSSEI are indicated to measure self-esteem.

However, Figure 8 shows that the general factor of the RSES is better represented by its facet *lack of negative self-esteem* than by the facet *positive self-esteem* (*lack of negative self-esteem* is more centered). Furthermore, the facet *lack of negative self-esteem* shows a high correlation to the facet *control* of the SMTQ. This result is not as surprising and can explain the high correlation of their general factors. Figure 8 shows that using FBRV facilitates the understanding of tests and their operational definition, i.e. FBRV helps to understand what tests really measure when they are used. Consequently FBRV helps to interpret correlations between tests and to external criteria.

Discussion

Using FBRM in test selection

FBRV can be used to illustrate the structure of a single test or to compare different tests. In the sections above we showed that FBRV facilitates interpretations of overall test scores of multifaceted constructs. Now we want to elaborate this aspect and explain why this is strongly associated with a appropriate test selection in psychological research and practice.

Overall test scores of multifaceted constructs could be calculated on the basis of different facet scores (i. e. the scores are not sufficient). The individual facet scores are not reconstructible when analyzing overall scores. For this reason it is important to know how the different facets are represented in the overall score. The correlated parts of all facets explains variance in each item and thus explains the most variance in the overall score. In a FBRV the correlated parts of the facets (the main source of variance) is represented by the center of the radar chart.

As described in the section *Interpretation of general factors*, the main source of variance of the RSES is similar to its facet *lack of negative self-esteem* and the main source of variance of the DSSEI is similar to its facets *Social Competence* and *Task-Related Abilities*

(Figure 8). Researchers who want to use a self-esteem scale can evaluate which aspect of self-esteem is the most important for them, i.e. which test is the most valid one regarding their hypotheses and research goals. In those circumstances usually the more reliable test is used. But the more reliable test is not necessarily the more valid when focusing on less relevant facets. The same applies to unidimensional scales (Rasch, 1960, 1961). A scale can be unidimensional without being particularly valid when focusing on a less relevant aspect of the investigated construct.

The described association between reliability and validity is the main reason why FBRV illustrates additional sources of variance. Each kind of systematic variation in the data can increase reliability and can be illustrated in a FBRV. Whether the additional sources of variance are viewed as valid depends on the research purpose and the underlying theory. Usually the facets of a published test are part of the theory suggested by the test authors. For this reason a test comparison is often equivalent with a theory comparison. By facilitating test comparisons and the differentiation between reliability and validity FBRV supports formation and examination of theorys.

Comparison to bifactor models

In the section above we described, why FBRV illustrates additional sources of variance. In our calculation we use ratios of squared loadings instead of bifactor modelling (Reise, Moore, & Haviland, 2010). Bifactor models also differentiate between the main source of variance and additional shared sources of variance and therefore are also easier to interpret than a single correlated factor model (Dantlgraber, Wetzel, Schützenberger, Stieger, & Reips, 2016). However, modelling a general factor and some additional uncorrelated factors bifactor models assume that the additional sources of variance are uncorrelated. This is problematic because many published tests are based on correlated factor models that do not have this assumption. In order to illustrate published tests without additional assumptions ratios of squared loadings are used in a FBRV. Furthermore, the fact that two separate models are estimated when using ratios of squared loadings offers the possibility of estimation in smaller samples and enables the connection of different samples.

Creating large networks

FBRV offers the opportunity to zoom in and out of constructs, visualizing psychological questions on different hierarchical levels. For example, the item "*In general, I feel confident about my abilities"* has the highest loading on the facet *Task-Related Abilities* of the DSSEI (Hoyle, 1991). Further studies could investigate which kind of abilities are highly correlated with self-esteem by analyzing an expanded item pool. The further separation of this facet could be illustrated in an enlarged FBRV (see Figure 9).

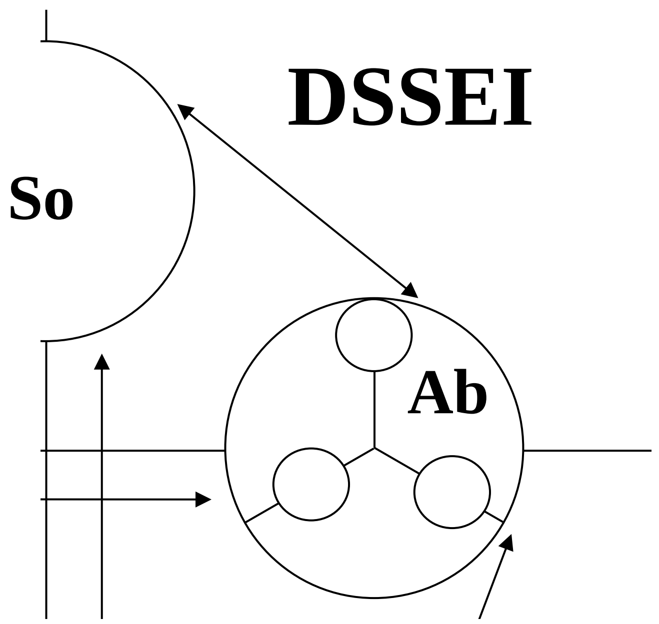

Figure 9. Enlarged Factor Based Radex Visualization. DSSEI = Domain Specific Self-Esteem Inventory; So = Social Competence; Ab = Task-Related Abilities.

Apart from that, *self-confidence* (the superordinate factor of Figure 8) could be compared with another correlated factor representing another correlated construct (see Figure 10). That means, the hierarchical structure of FBRV offers the opportunity to extend the illustration with further correlating tests building a large network of constructs.

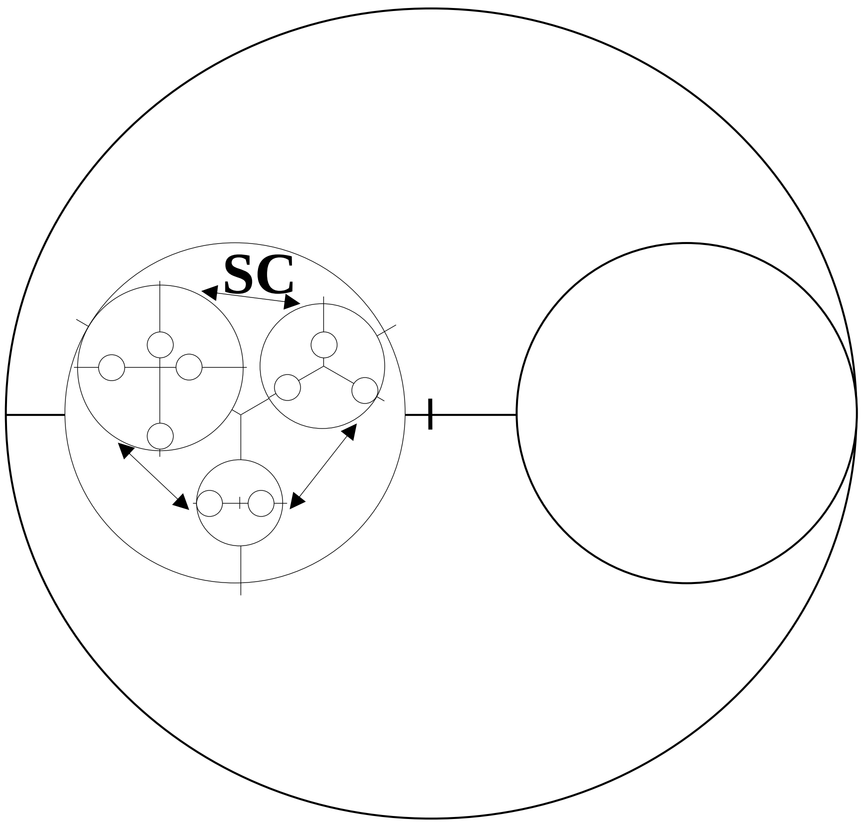

Figure 10. Minimized Factor Based Radex Visualization. SC = Self Confidence

The potential to create large network seems to be limited by the reasonable length of a study. Naturally, participants cannot fill out endless questionnaires. However, the hierarchical structure of FBRV can also be used to connect comparable but different samples. For example, the participants of our sample filled out all items of the DSSEI, the SMTQ, and the RSES representing all facets. We located the three tests and its facets in a superordinate radar chart (see Figure 8). Its center represents the superordinate factor (*self-confidence*). This sample could be used to create a short form representing the same superordinate factor by splitting the whole item pool into a smaller and a larger item pool that are perfectly correlated (latent correlation >.95). In a further study the smaller item pool could be used as a short form to locate the network of the three tests (see Figure 8) in relation to another construct (see Figure 10).

Comparison to network analysis

The emphasis on overall commonalities and the possibility to create large networks naturally reminds of network analysis (see Cramer et al., 2012). Nevertheless their basic concepts are different. Network analysis estimates a network of all included items on the basis of partial correlations. Therefore network analysis is similar to exploratory factor analysis and can be used to identify new relationships or to prove assumed relationships without restrictions. However, because of its exploratory approach, network analysis can deliver confusing and inconclusive results. This regards in particular comparisons of two or more multifaceted tests.

FBRV is similar to confirmatory factor analysis and locates theory based item pools, i. e. the effect of items when they are combined to scales or tests as indicated. We think the latter is more relevant for practical purposes because items are rarely solely used because of reliability alone.

Compatibility with traditional visualization

To conclude we want to briefly mention that FBRV is compatible with the traditional visualization of factor structure. Like the traditional visualization, FBRV uses cicles to represent latent factors. For this reason it is possible to integrate a FBRV in a traditional visualization. Figure 11 shows an integration of the superordinate factor (*self-confidence*) into a possible mediation analysis.

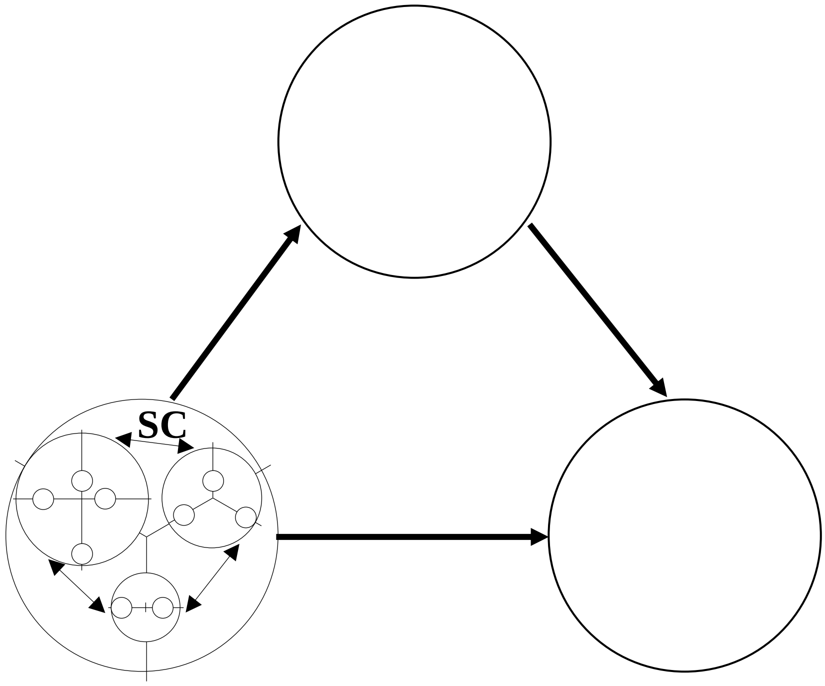

Figure 11. Integration of a Factor Based Radex Visualization into a traditional one

Moreover, FBRV illustrates a reference system of various factors. Factor loadings of the items vary depending on the reference factor. For this reason we recommend the use of tables showing various loadings of the same item regarding various reference factors (see Table 1). Nevertheless, it is possible to integrate factor loadings in a FBRV. Figure 5 shows that FBRV can be used to locate items. Such a figure could also be complemented with factor loadings and information about the reference factor.

Limitations

In genral FBRV has some limitations but first we want to discuss a specific limitation in this article. The sample used to demonstrate task group location (see Figure 5) is relatively small (n = 145). Therefore there is some uncertainty about the exact location of the task groups. However, there is no doubt that M-KIT's task groups are more centered than I-S-T 2000 R's task groups. As mentioned before the M-KIT showed a single factor strucutur in its validation data (n = 1,054) and the I-S-T 2000 R showed a structure of three correlated factors in its validiation data (n = 2,208). These structures were replicated with the small sample (n = 145) and are illustrated with more and less centered task groups in Figure 5.

Regarding limitations in general, FBRV are illustrations of item groups and subgroups and the relations among themselves. There is no estimating process involved that helps decide if the implemented groups are reasonable. Obviously, analyzing FBRV offers some hints whether regrouping could be appropriate. Nevertheless, FBRV does not replace traditional factor analysis. Exploratory or confirmatory factor analysis is needed to investigate possible facets of a construct or other kinds of systematic of variation in the data.

FBRV illustrates commonalities of items by using circles. Some researcher may miss the possibility to arrange complex structures of positive and negative correlated constructs and facets that cannot be easily recoded (see *Requirements of a FBRV* and *Procedure*). Addressing this issue, we want to point out the compatibility of FBRV with the traditional visualization (see Figure 11).

The amount of information contained in Figure 8 might also confuse some readers. In our opinion, FBRV would show its whole potential in a computer based version in which users can easily zoom in and zoom out and switch on and off information any time. In addition, it should be kept in mind that the traditional visualization is arranged more complicated than FBRV (see Figure 1 and Figure 2). When using the traditional visualization comparisons of different tests are often reduced to comparisons of general factors. As described in the introduction section that ignores the specific operational definition of each test and leads to misdirected test use.

References

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511571312

Cramer, A. O. J., Van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., …, Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality, 26*, 414–431.

Dantlgraber, M. (2015). *Modularer Kurzintelligenztest (M-KIT) [Modular Short Intelligence Kit]*. Bern: Hogrefe.

Dantlgraber, M., Wetzel, E., Schützenberger, P., Stieger, S., Reips, U.-D. (2016). Simple construct evaluation with latent class analysis: An investigation of Facebook addiction and the development of a short form of the Facebook Addiction Test (F-AT). *Behavior Research Methods, 48*, 869–879. doi:10.3758/s13428-016-0716-2

Guttman, L. A. (1954). A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258-348). Illinois: Glencoe.

Hoyle, R. H. (1991). Evaluating measurement models in clinical research: Covariance structure analysis of latent variable models of self-conception. *Journal of Consulting and Clinical Psychology, 59*, 67-76.

Kaplan, D. (2001). *Structural equation modeling: Foundations and extensions*. Sage, Thousand Oaks, CA, USA. doi:10.1016/B0-08-043076-7/00776-2

Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test-2000-R (I-S-T 2000 R). Manual (2. erweiterte und überarbeitete Aufl.). [Intelligence Structure Test 2000 R. Manual (2nd. ed)]*. Göttingen: Hogrefe.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7,* 107-127. doi:10.1016/0160-2896(83)90023-5

McGrew, K. S. (2005). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136-181). New York: Guilford.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37,* 1-10.

Rasch, G. (1960). *Studies in mathematical psychology: I. probabilistic models for some intelligence and attainment tests*. Oxford, UK: Nielsen and Lydiche.

Rasch, G. (1961). *On general laws and the meaning of measurement in psychology*. Berkeley: University of California Press.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92,* 544-559. doi:10.1080/00223891.2010.496477

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Sheard, M., Golby, J., & van Wersch, A. (2009). Progress toward construct validation of the Sports Mental Toughness Questionnaire (SMTQ). *European Journal of Psychological Assessment, 25*, 186-193. doi:10.1027/1015-5759.25.3.186

Supple, A. J., Su, J., Plunkett, S. W., Peterson, G. W., & Bush, K. R. (2013). Factor structure of the Rosenberg Self-Esteem Scale. *Journal of Cross-Cultural Psychology, 44*, 748-764. doi:10.1177/0022022112468942

Wright, S. (1921). Correlation and causation. *Journal of agricultural research, 20*, 557–585.