

Perspectives on the Validity of the Thinking Styles Inventories

Florian Berding¹, Maike Masemann^{1,*}, Karin Rebmann¹ & Manuela Paechter²

¹Department of Business Administration, Economics, and Law, Carl von Ossietzky University, Oldenburg, Germany

²Department of Psychology, Educational Psychology Unit, Karl-Franzens University, Graz, Austria

*Correspondence: Department of Business Administration, Economics, and Law, Carl von Ossietzky University, Ammerlaender Heerstraße 114-118, 26129 Oldenburg, Germany. E-mail: maike.masemann@uni-oldenburg.de

Received: November 15, 2016

Accepted: November 28, 2016

Online Published: December 13, 2016

doi:10.5430/wje.v6n6p69

URL: <http://dx.doi.org/10.5430/wje.v6n6p69>

Abstract

The Thinking Styles Inventories (TSI) are questionnaires for assessing individual preferences in constructing knowledge. This paper identifies several problems concerning their validity, which range from an inadequate use of factor analysis, to missing information on the measurement model, to findings indicating a low discrimination between the thinking style scales. Against this background, two studies are conducted providing detailed insights into the measurement model of the TSI in German-speaking samples (Study I: 287 apprentices; Study II: 389 students). Although results indicate a high degree of reliability according to popular statistical rules, they confirm problems with the discriminant validity and criterion validity regarding achievement. The Thinking Styles Inventories should as a result be used with caution.

Keywords: *intellectual styles; thinking styles; validation; questionnaire; vocational education and training*

1. Introduction

In exploring individual differences in constructing knowledge and dealing with tasks, the concept of thinking styles by Sternberg (1997) has gained international attention because it reflects a concept for individual preferences in cognition, thinking, and learning (e.g., Zhang 2008; Zhang & Sternberg, 2005). For example, empirical studies report relationships between thinking styles and anxiety (e.g., Zhang, 2009), personality (e.g., Zhang, 2006), emotional intelligence (e.g., Murphy & Janeke, 2009), reflection (e.g., Chen, Kinshuk, Wei & Liu 2011), and academic achievement (e.g., Richmond & Conrad, 2012). In other words: thinking styles provide a concept for understanding learning differences.

Studies analyzing the relationships between thinking styles and learning phenomena that are based on Sternberg's (1997) theory generally use a version of the Thinking Styles Inventory (TSI; Sternberg & Wagner, 1992) for assessing the participants' thinking styles. Although the TSI has been revised twice (TSI-R and TSI-R2; Sternberg, Wagner & Zhang, 2007), and a large number of empirical work is built on this instrument, studies provide only limited information on the quality of the questionnaire, indicating validity problems as a result. This lack of information results from (1) an unusual application of factor analyses that ignores the relationships between a thinking style and its corresponding items (e.g., Dai & Feldhusen, 1999; Fjell & Walhovd, 2004; Cano-Garcia & Hughes, 2000; Sternberg, 1994; Zhang, 2005, 2008; Zhang & Higgins, 2008); and (2) these studies applying factor analysis in ways that produce conflicts between the assumptions of factor analysis (e.g., a higher-order variable causes the characteristics of a lower-order variable; Brown, 2015) and the assumptions of the Theory of Mental Self-Government. Findings also reveal (3) a low discrimination between the thinking styles' scales (e.g., Black & McCoach, 2008; Dai & Feldhusen, 1999; Zhang, 1999); and (4) an unstable number of factors as well as style compositions contradicting Sternberg's (1997) theory (e.g., Dai & Feldhusen, 1999; Fjell & Walhovd, 2004; Cano-Garcia & Hughes, 2000; Ngan Man Fon, 2013; Zhang, 2008; Zhang & Higgins, 2008). The question that results here is: Just how reliable and valid are in fact the Thinking Styles Inventories?

Because a measurement instrument is the foundation of all insights into empirical research, information on the reliability and validity of such an instrument is essential. To provide this information, this paper aims to give a detailed view into problems concerning the validity of the Thinking Styles Inventories, and concentrates on the

relationships between thinking styles and their corresponding items by applying the Thinking Styles Inventory in German-speaking samples. Before analyzing critical aspects of the validity, the following section gives a brief overview of Sternberg's (1997) Theory of Mental Self-Government and the Threefold Model by Zhang and Sternberg (2005). Both of these are used as arguments for the validity of the questionnaire (e.g., Zhang, 2005).

2. Sternberg's Theory of Mental Self-Government and the Threefold Model by Zhang and Sternberg

2.1 Theory of Mental Self-Government

The Theory of Mental Self-Government is based on the assumption that the different forms of governments found in the world are representations of different ways of thinking. Every individual has to organize its thinking, decide on priorities, and allocate its resources just like a government has to (Sternberg, 1997). In this tradition, an individual has to "govern" its thinking. An individual's preference to do this is denoted as a *thinking style*. These thinking styles are categorized into five dimensions ("functions," "forms," "levels," "scopes," and "leanings" of mental self-government).

The thinking styles within the "functions" dimension are the legislative, executive, and judicial styles. People favoring the *legislative* style tend to be more creative and prefer activities with high degrees of freedom in choosing their own strategies, whereas people with a tendency for the *executive* style show preferences for clear instructions and pre-structured tasks. The *judicial* style is characterized by a preference for evaluating situations, solutions, and performances.

The dimension "forms" includes the hierarchical, monarchic, oligarchic, and anarchic styles. The preferences described by these styles differ in terms of setting priorities and flexibility while working on tasks. The *hierarchical* style denotes preferences for prioritizing tasks, whereas the *anarchic* style is characterized by preferences for working on tasks at random. The *monarchic* style is indicated by preferences for focusing on one task at a time, and the *oligarchic* style includes preferences for dealing with several tasks without a clear hierarchy of goals or priorities.

The global and the local thinking styles fall into the "levels" dimension, which can be viewed as a dipole of preferences for details. The *global* style represents a preference for abstraction; the *local* style includes preferences for high levels of detail.

The dimension "scopes" represents a dipole in terms of sociological preferences. In this dimension of thinking styles, the *internal* style denotes preferences for working independently of others, while the *external* style indicates preferences for working in groups.

The "leanings" dimension delineates a contrast of styles in terms of preferences for tradition or innovation. The *liberal* style illustrates preferences for curiosity in order to find a proper solution, which stands in contrast to the *conservative* style that is marked by preferences for existing and established rules in performing tasks (e.g., Sternberg, 1997; Zhang & Sternberg, 2005, 2006, 2009).

2.2 The Threefold Model of Intellectual Styles

The Theory of Mental Self-Government is incorporated in the Threefold Model of Intellectual Styles by Zhang and Sternberg (2005). It provides a framework for systemizing the somewhat different, somewhat same constructs of styles within the style literature. To do this, it classifies the styles proposed by several researchers into three types of "intellectual styles." *Type I* styles carry preferences for tasks with higher degrees of complexity and less norm-favoring tendencies (e.g., Zhang & Sternberg, 2005). *Type II* styles are characterized by preferences for following existing rules and well-structured tasks. *Type III* intellectual styles show characteristics of both Type I and Type II styles; the specific requirements to solve a problem and to deal with a task, e.g. the social setting of a task, determine the "deployment" of these styles as well as the individual's interest in dealing with the specific situation (e.g., Zhang, 2013; Zhang & Sternberg, 2005).

Based on this model, the legislative, judicial, global, hierarchical, and liberal styles belong to Type I; the executive, local, monarchic, and conservative styles belong to Type II; and the oligarchic, anarchic, internal, and external styles are categorized as Type III (Zhang & Sternberg, 2005). The following section works out problems concerning the validity of the Thinking Styles Inventories based on these two theoretical approaches.

3. Critical Findings Regarding the Reliability and Validity of the Thinking Styles Inventories

3.1 Use of Factor Analysis for Validating the Thinking Styles Inventories

A central criterion for the quality of measurement instruments is its validity. It “is a unitary concept, which always refers to the degree to which empirical evidence and theoretical rationales support each other” (Lounsbury, Gibson, & Saudargas, 2006, p. 139). An important tool for investigating the validity of a measure is factor analysis (e.g., Brown, 2015). Studies examining the Thinking Styles Inventories use this tool as part of two different approaches (see Figure 1) (e.g., Fan, 2014; Fjell & Walhovd, 2004; Cano-Garcia & Hughes, 2000; Zhang & Higgins, 2008).

The first approach is based on *item-level* data and represents the “classical” application of factor analysis for validating questionnaires (e.g., Brown, 2015; Wang & Wang, 2012). This approach focuses on the relationships between an unobservable latent construct (e.g., the internal thinking style) and observable variables for assessing the latent construct (e.g., the item “I prefer situations where I can carry out my own ideas, without relying on others”). It assumes that the answers to the items are caused by the latent construct reflecting a person’s characteristic of that latent construct. Therefore, factor analysis on this level is used to validate the measurement model of a questionnaire (e.g., Brown, 2015).

In the case of exploratory factor analysis (EFA), this means showing that the number of assumed latent constructs is really represented in the data and that the items belong only to the latent constructs they were developed for. For example, if the questionnaire tries to measure the internal and external style, EFA should indicate two factors in the data: one for the internal and one for the external style. Furthermore, the items for the internal style should only belong to one factor and the items for the external style should only refer to the other factor (this is presented in Figure 1). If an item formulated for the external style belongs to the factor representing the internal style, this would indicate that the item is not able to assess the external style. Thus, factor analysis on an item-level provides a deep insight into the validity of an instrument. This is especially true for confirmatory factor analysis (CFA) which provides additional criteria for judging the reliability and validity of a questionnaire such as the item reliability, the average variance extracted, and the Fornell-Larcker criterion (e.g., Bagozzi & Baumgartner, 1994; Bagozzi & Yi, 1988; Fornell & Larcker, 1981; Wang & Wang, 2012).

Based on the definition of validity at the beginning of this section, factor analysis has to provide evidence for a correspondence between data and theory. On an item level, this means that the Theory of Mental Self-Government will show that 13 factors exist in the data, representing the 13 thinking styles of the theory, and that the items belong *only* to the thinking style they were developed for. Unfortunately, the item level of the Thinking Styles Inventories is rarely investigated, providing only limited information on the quality of the instrument (e.g., Black & McCoach, 2008; Fan, 2014).

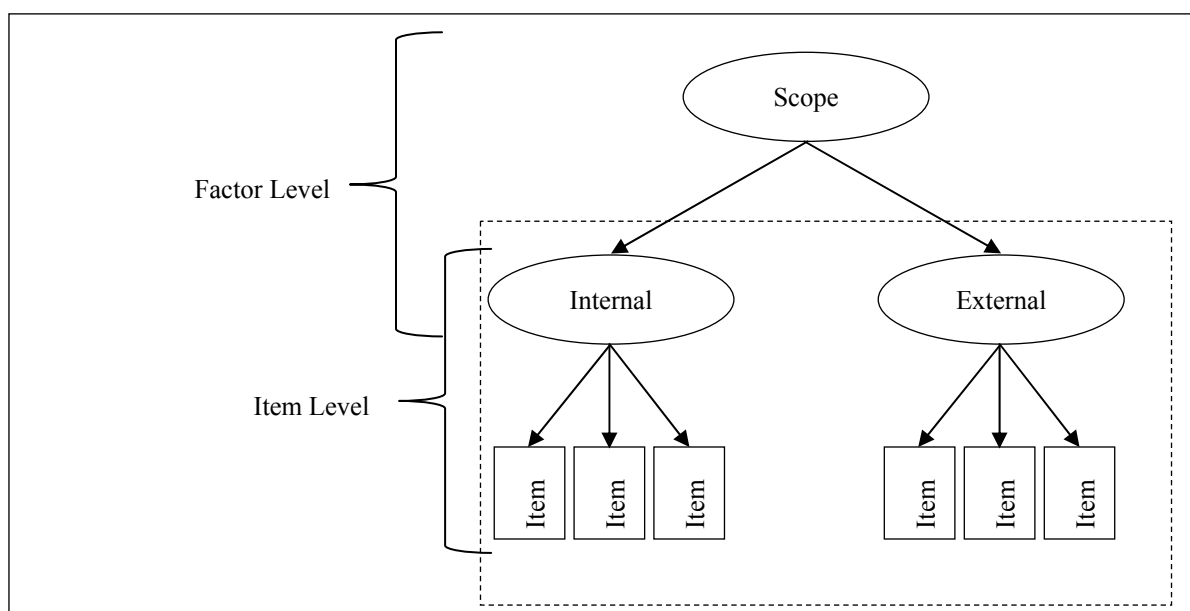


Figure 1. Levels of Factor Analysis in Validating the Thinking Styles Inventories

The second approach used in most studies validating Thinking Styles Inventories focuses on a *factor level* (e.g., Fjell & Walhovd, 2004; Cano-Garcia & Hughes, 2000; Zhang & Higgins, 2008). This approach investigates the relationships between latent constructs (e.g., the internal thinking style) and a higher-order latent construct (e.g., scopes) as illustrated in Figure 1. The idea behind this method of validating a questionnaire is to show that the measured latent constructs are caused by a higher-order latent construct as postulated by a theory (e.g., Black & McCoach, 2008). Scores for the thinking styles are computed by assuming that the items measure the thinking style they were developed for (item level).

Applied to the Theory of Mental Self-Government, five factors have to be expected (e.g., functions, forms, scopes) causing the corresponding thinking styles (e.g., legislative, executive, and judicial style for the dimension “functions”). For example, in Sternberg’s (1997) theory, the external and internal styles belong to the dimension “scopes.” So if the questionnaire is valid, the internal and external style should be caused by the “scopes” dimension. If the internal and external style cannot be reduced to this dimension, this would indicate invalidity.

Although this approach appears more than plausible for validating the Thinking Styles Inventories, it does not take into account the relationships between the items and their corresponding thinking styles. These relationships however are the basis *for all* analysis. Put another way, this approach assumes a valid measurement model without providing evidence for this assumption, leading to limited information about the quality of the questionnaire itself.

To sum up, the common use of factor analysis for validating the Thinking Styles Inventories in studies conducted to date provides hardly any information about the quality of the underlying measurement model. Further studies should focus more strongly on the item level.

3.2 *Instable Factor Structures and their Interpretations*

Another problem concerning the validity of the Thinking Styles Inventories refers to the findings revealed by factor analysis. Sternberg’s (1997) theory assumes five dimensions of thinking styles. Thus, five factors have to be expected on a factor level. Studies using EFA however reveal inconsistent results regarding factor number and factor composition. For example, some studies have identified two (e.g., Zhang, 2008), three (e.g., Dai & Feldhusen, 1999; Ngan Man Fong, 2013), four (e.g., Cano-Garcia & Hughes, 2000; Zhang, 2005; Zhang & Higgins, 2008) or five factors (e.g., Fer, 2005; Fjell & Walhovd, 2004; Sternberg, 1994; Zhang, 1999; 2003).

Even in the case of five factors, their composition of thinking styles is not completely in line with the theory (e.g., Sternberg, 1994; Zhang, 1999). For example, Fjell and Walhovd (2004) report a factor which comprises the judicial, liberal, legislative, hierarchical, and monarchic styles. This factor contains styles belonging to the “functions,” “forms,” and “leanings” dimensions of the Theory of Mental Self-Government. This means styles of different dimensions are mixed, so the dimensions are not purely reflected by the data.

Although these studies provide no strong evidence for the validity of the questionnaire, these results are interpreted as indicators of a high instrument quality (e.g., Ngan Mang Fong, 2013; Zhang, 2008; Zhang & Higgins, 2008). For example, Zhang (2005) identified *four* factors, with the first factor containing Type I and Type III styles, the second Type II and Type III styles, the third contrasting the internal and the external, and the fourth comparing the global with the local style. Zhang (2005, p. 1919) draws the conclusion: “Taken together, these four factors lent strong support to the validity of the TSI-R for assessing the present research participants’ thinking styles.” Indeed, these factors provide *no* strong support since two models here are mixed into a new argument. The Type I, II, and III styles of the factors one and two reflect the Threefold Model of Intellectual Styles by Zhang and Sternberg (2005) in parts; while the global, local, internal, and external styles of the third and fourth factors reflect Sternberg’s (1997) Theory of Mental Self-Government in parts. The results neither reproduce the Threefold Model nor the Theory of Mental Self-Government completely. This can hardly be considered a strong argument for validity being the degree to which data and theory support each other (Lounsbury, Gibson, & Saudargas, 2006).

On an item level, the confirmatory factor analysis (CFA) conducted by Fan (2014) proves the measurement model of the TSI-R2 for all 13 thinking styles. In contrast, the study by Black and McCoach (2008) could not confirm the measurement model of the TSI. Their EFA revealed only nine factors, and several items showed relevant factor loadings to more than one factor. To sum up, it is not clear how valid the Thinking Styles Inventories are.

3.3 *Factor Analysis on a Factor Level as an Inadequate Analysis Tool*

Proving the validity of the Thinking Styles Inventories with factor analysis on a factor level is generally problematic for the Theory of Mental Self-Government. Factor analysis assumes a higher-order variable causing the characteristics of a lower-order variable (e.g., Brown, 2015). For the Thinking Styles Inventories, this means that the dimension “functions” causes the legislative, executive, and judicial styles; “forms” the monarchic, hierarchical,

oligarchic, and anarchic styles; “levels” the global and local; “scopes” the internal and external; and “leanings” the liberal and conservative styles.

This central assumption of factor analysis is meaningful for the “scopes,” “leanings,” and “levels” dimensions because their corresponding styles are “contrasted with each other” (Sternberg, 1997, p. 64). For example, a high value on the higher-order variable “scopes” could mean a low preference for the internal and a high preference for the external style, whereas a high value on the higher-order variable “levels” could indicate a low preference for the local and a strong preference for the global style. And in fact, this suggestion is partly reflected by exploratory factor analysis (e.g., Fjell & Walhovd, 2004; Sternberg, 1994; Zhang, 2003; Zhang & Higgins, 2008). Although Sternberg (1997) postulates that people do not have to be internally *or* externally exclusive in every situation, because the styles are conceptualized as at least partially socialized, flexible, and changeable in different situations, the respective descriptions of the styles and their corresponding items suggest the respective contrast.

For the other dimensions, this assumption is dubious. It is unclear how the dimensions “functions” and “forms” could cause the characteristics with the corresponding thinking styles, because every style is qualitatively different. The corresponding styles do not reflect dipoles and/or do not contrast each other. For example, the judicial style belonging to the dimension “functions” is characterized by a preference for working on tasks that allow for one’s evaluation, whereas the legislative style belonging to the same dimension is characterized by a preference for working on tasks that require creative strategies. These preferences do not exclude each other like the preferences of the internal and the external styles. And when it comes to the descriptions of the styles named, it seems plausible and explainable that an individual can show preferences for the judicial and legislative styles at the same time and within the same situation; only for the judicial style; only for the legislative style; or neither of them. Taking this example into account, there is no reason to conclude that a higher-order variable causes the characteristics of the “function” styles according to a stable pattern.

3.4 Low Discrimination between the Thinking Styles’ Scales

Another problem concerning the Thinking Styles Inventories is the discriminant validity of the scales. Discriminant validity can be described as “the degree to which a construct is discriminable (e.g., uncorrelated) from, and non-redundant with, other constructs” (Lounsbury, Gibson, & Saudargas, 2006, p. 139). High correlations between two thinking styles’ scales may indicate that both scales measure the same construct. According to Cohen (1988) an absolute of at least .10 indicates a small, .30 a medium, and .50 a strong relationship. Based on this classification, Dai and Feldhusen (1999) report at least moderate to strong correlations for 38%, Zhang (1999) for 44%, and Black and McCoach (2008) for 65% of the possible correlation pairs between the thinking styles. The average absolute correlations (via Fisher-Z transformation) in these studies are .295, .299, and .380, making it questionable whether the 13 thinking style scales are overlapping.

Taken together, the findings and arguments presented in this section challenge the validity of the Thinking Styles Inventories. Against this background, the current paper aims to provide an additional analysis for German-speaking samples on an item level to give a detailed insight into the quality of the underlying measurement model, and to provide recommendations for practical applications. Due to the fact that a German version is still missing, the next section outlines the development and validating process for the new questionnaire.

3.5 Translating and Testing a Thinking Style Inventory (TSI-GER) for German-Speaking Samples

As the detailed review of the existing literature showed, most studies focus on a factor level, providing only limited information on the compositions of items to factors (e.g., Fjell & Walhovd, 2004; Sternberg, 1994; Zhang, 2003; Zhang & Higgins, 2008). This is why the questionnaire analysis has to focus on an item level. This analysis should also anticipate that not all items are valid indicators. So in order to increase the chance of finding valid items, an item pool of 111 items is created consisting of all 65 items of the TSI-R2 (this is the latest revision of the Thinking Styles Inventories, and provides the most promising items for the German version of the questionnaire). 46 items from the TSI complete the item pool as “backup” items, i.e. they are used if items from the second revision of the TSI have to be deleted due to bad psychometric properties. All items were translated into German and, for control purposes, back-translated into English.

A CFA is conducted to close the item level information gap. CFA offers the following additional criteria to judge the reliability and validity of an instrument, and addresses most of the issues outlined at the beginning of this paper.

- The item reliability (IR) is the amount of item variance that is explained by the corresponding latent variable (Wang & Wang, 2012). It should exceed 40% (Bagozzi & Baumgartner, 1994).
- The composite reliability (CR) uses all items belonging to a latent variable and describes the reliability of the

scale as a whole (Wang & Wang, 2012). Values above .600 are recommended here (Bagozzi & Yi, 1988).

- The average variance extracted (AVE) should be at least .500 (Bagozzi & Yi, 1988).
- The strict Fornell-Larcker criterion can be used to judge the discriminant validity of a construct. It requires that the squared correlation between two constructs be lower than the average variance extracted from them (Fornell & Larcker, 1981).

Additionally, a questionnaire has to show criterion validity. It describes “how well a test corresponds with a particular criterion. Such evidence is provided by high correlations between a test and a well-defined criterion measure” (Kaplan & Saccuzzo, 2013, p. 138). The concept of achievement is used for evaluating the criterion validity of the German version of the TSI, as several studies report relationships between thinking styles and achievement (e.g., Cano-Garcia & Hughes, 2000; Richmond & Conrad, 2012; Zhang, 2001, 2002, 2004a; Zhang & Sternberg, 2000). For example, Zhang (2004a) examined that the hierarchical style significantly contributed to higher achievement scores in the social sciences and humanities, whereas the judicial style contributed to higher natural science scores of secondary school students in Hong Kong.

Two studies are carried out in the process of the instrument’s development. Study I (exploration) uses a CFA approach to develop a composition of reliable items that represents the 13 thinking styles in its factor structure. Furthermore, the discriminant validity and criterion validity regarding achievement is tested. Study II (confirmation) investigates whether the factorial structure and psychometric properties identified in Study I can be replicated in another sample. In addition, Study II reports the retest reliability. The method, samples, and results are described in the following section. To avoid redundancies, the results of both studies are discussed at the end of the paper.

4. Study I: Exploring and Optimizing the Factor Structure and Psychometric Properties of the Preliminary TSI-GER

4.1 Participants and Context

204 males and 83 females completing their vocational education within the dual education system participated in the study. The apprentices were on average 20.57 years old ($SD = 3.11$) (one missing case), and were training in either technical ($n = 212$) or economic vocational education ($n = 73$; two missing cases).

4.2 Measures

Thinking Styles. The preliminary TSI-GER includes 111 items. Individuals assess how well an item describes their behavior on a seven-point scale ranging from 1 (*not at all well*) to 7 (*extremely well*). High values indicate a high preference for a certain thinking style.

Achievement. The apprentices’ grades in their last German, mathematics, and economics tests were collected. Grades could vary between 1 (*very good*) and 6 (*insufficient*).

4.3 Analysis

A CFA with 13 latent variables representing the 13 thinking styles is performed to test the quality of the instrument. Items are assigned to the factors as the *theory* expects them to be. Criterion validity with reference to achievement is investigated by means of multiple linear regression analysis.

4.4 Results

The distribution of the data is first analyzed. The univariate skews of the items do not exceed the absolute value of 0.744, and the absolute values of the univariate kurtosis’ are lower than 0.891. They are within the range for a moderate violation of normal distribution ($|\text{skew}| < 2$, $|\text{kurtosis}| < 7$) and allow the application of the maximum likelihood algorithm (West, Finch, & Curran, 1995).

A first computation of a CFA model with 13 latent variables and 111 items lead to estimation problems because the latent variable covariance matrix is not positively definite, indicating negative variances, linear dependencies, or correlations of at least one between latent variables. Indeed, some correlations are exceptionally high (e.g., internal and legislative $r = .914$). Similar estimation problems are reported by Black and McCoach (2008). Against this background, the complete model was divided into three sub-models to provide information on at least parts of the questionnaire representing the dimensions of the Theory of Mental Self-Government. Here, Model 1 comprises the legislative, executive, and judicial styles (“functions”); Model 2 the monarchic, hierarchical, oligarchic, and anarchic styles (“forms”); and Model 3 the global, local, internal, external, liberal, and conservative styles (“levels,” “scopes,” and “leanings”).

CFA reveals the following results for the global model fit after successively removing items with low item reliability: Model 1 “functions” ($\chi^2(51) = 96.318, p < .01$; RMSEA .056; SRMR .060; CFI .931), Model 2 “forms” ($\chi^2(164) = 323.798, p < .01$; RMSEA .058; SRMR .058; CFI .867), and Model 3 “levels,” “scopes,” and “leanings” ($\chi^2(335) = 549.906, p < .01$; RMSEA .047; SRMR .064; CFI .912). According to Hu and Bentler’s (1999) combinational rule, RMSEA below .06 and SRMR below .09 indicate global model fit for all three parts of the questionnaire. CFI values below .950 point to a global model misfit in all calculated models (e.g., Hu & Bentler, 1999). However, the model evaluation is based only on the combinational rule because CFI compares the hypothetical model with a baseline model. The baseline model assumes all observed variables to be uncorrelated, which is usually inappropriate for many scientific applications (Kline, 2005). As Heene, Hilbert, Draxler, Ziegler, and Bühner (2011) argue, decreased factor loadings generally imply reduced correlations between observed variables and lead to a lower CFI. Psychological tests however only seldom achieve high factor loadings. Thus, a research model that assumes low factor loadings, i.e. a model similar to the baseline model, achieves only a low CFI even though it is not necessarily mis-specified or meaningless. In the current study, low to moderate factor loadings do in fact appear (see Table 2). Thus, RMSEA and SRMR are used for model evaluation without regard to CFI. The final solution consists of 60 items (TSI-R2: 41; TSI: 19). The number of items per scale ranges between three to eight items, and at least three items per factor are recommended for standard CFA (Kline, 2005). Table 1 reports Cronbach’s α , the composite reliability, and the average variance extracted for Study I.

Table 1. Cronbach’s α , Composite Reliability, Average Variance Extracted, and Retest Reliability for the Scales from Study I and II

Style	Study I			Study II			
	α	CR	AVE	α	CR	AVE	RT
Legislative	.704	.710	.384	.740	.740	.417	.786
Executive	.699	.714	.386	.786	.791	.495	.639
Judicial	.706	.704	.374	.700	.702	.376	.640
Monarchic	.699	.710	.455	.668	.717	.480	.777
Hierarchical	.695	.709	.294	.716	.715	.311	.745
Oligarchic	.777	.776	.410	.803	.810	.469	.642
Anarchic	.647	.648	.239	.592	.608	.208	.648
Global	.573	.585	.324	.636	.635	.371	.726
Local	.634	.635	.369	.804	.808	.514	.859
Internal	.667	.669	.291	.725	.725	.347	.733
External	.706	.736	.415	.854	.856	.599	.713
Liberal	.884	.887	.497	.900	.900	.529	.844
Conservative	.805	.803	.449	.821	.817	.476	.642

Note: CR = Composite Reliability; AVE = Average Variance Extracted; RT = Retest Reliability.

Cronbach’s α ranges between .573 for the global style and .884 for the liberal style (e.g., Table 1). For an internally consistent scale, Adams and Lawrence (2015) demand the Cronbach’s α value to be at least .700. The legislative, executive, judicial, hierarchical, monarchic, oligarchic, external, liberal, and conservative styles accomplish this rule. Except for the global style, the composite reliability exceeds the critical value of .600 with a minimum of .635 (local style) and a maximum of .887 (liberal style). However, the item reliabilities are low and vary within a scale ($Min = .127, Max = .628, Mdn = .381$). Only the liberal style consists completely of items whose variance is at least explained by 40% of the latent variable. Similar to the item reliabilities, the average variance extracted is low and ranges between .239 for the anarchic style and .497 for the liberal style.

To investigate the discriminant validity of the questionnaire, the strict Fornell-Larcker criterion is applied. Table 2 shows the inter-scale correlations below the diagonal, the squared inter-scale correlations above the diagonal, and the average variance extracted on the diagonal.

Table 2. Inter-Scale Correlations, Squared Inter-Scale Correlations, and Average Variance Extracted for Study I

Style	Legislative		Executive		Judicial	
Legislative	.384		.012		.283	
Executive	.108		.386		.036	
Judicial	.532**		.191*		.374	
Style	Monarchic		Hierarchical		Oligarchic	Anarchic
Monarchic	.455		.371		.082	.042
Hierarchical	.609**		.294		.250	.123
Oligarchic	.286**		.500**		.410	.378
Anarchic	.204*		.351**		.615**	.239
Style	Global	Local	Internal	External	Liberal	Conservative
Global	.324	.006	.371	.084	.031	.195
Local	-.075	.369	.216	.058	.336	.029
Internal	.609**	.465**	.291	.001	.285	.058
External	.290**	.241**	.038	.415	.128	.043
Liberal	.177*	.580**	.534**	.358**	.497	.092
Conservative	.442**	.171*	.240**	.208**	-.304**	.449

Note. Values below the diagonal: inter-scale correlations; values above the diagonal: squared inter-scale correlations; values on the diagonal: average variance extracted.

* $p < .05$, ** $p < .01$

21 out of 24 thinking styles' pairs comply with the strict Fornell-Larcker criterion. Although the average variance extracted from the monarchic style is higher than the squared correlation between the monarchic and the hierarchical styles, the average variance extracted from the hierarchical style is lower than the squared correlation between them, indicating a violation of the Fornell-Larcker criterion. The same problem occurs for the oligarchic and anarchic styles, as well as for the global and internal styles.

To further test the discriminant validity of these three pairs of thinking styles, a Wald-chi-squared test is performed, testing the hypothesis that the correlation between one of the thinking styles' pairs equals one. Results for the monarchic and hierarchical ($\chi^2(1) = 38.817, p < .01$), oligarchic and anarchic ($\chi^2(1) = 36.604, p < .01$), and global and internal styles ($\chi^2(1) = 21.986, p < .01$) reject the hypothesis, indicating no perfect relationship between the thinking styles' pairs. Different but perhaps overlapping constructs can therefore be assumed.

A multiple linear regression is conducted to investigate the criterion validity with reference to achievement. As can be seen in Table 4, only the local style contributes significantly to achievement in mathematics. The more apprentices prefer to concentrate on details, the better their grades in this subject are. Students favoring the executive style and refusing the oligarchic style are more successful in German. Furthermore, the external style is negatively correlated and the anarchic style is positively correlated with achievement in economics.

Taken together, the first study shows advantages and disadvantages of the German version. The advantages are formed by partly high values for Cronbach's α , and the composite reliability as well as the confirmation of the factor structure for *parts* of the questionnaire. Disadvantages arise by failing to prove the factor structure for the *complete* questionnaire, and low values for the average variance extracted.

One reason for failing to prove the complete factor structure may be a large number of invalid items within the 111-item pool. Thus, the optimized and reduced item solution of Study I has the potential to provide a valid measurement model for the complete questionnaire. This hypothesis will have to be validated with another sample. Furthermore, new items for the global, local, monarchic, and anarchic styles are developed in order to improve the psychometric properties of these scales. The method and results of Study II are outlined in the following section.

Table 3. Summary of Multiple Linear Regression Analysis for Achievement Regressed on Thinking Styles in Studies I and II

Thinking Style	Mathematics (n = 114)/(n = 222)			German (n = 272)/(n = 223)			Economics (n = 194)/(n = 210)		
	b	SE b	β	b	SE b	β	b	SE b	β
Legislative	-.175	.144	-.146	-.029	.084	-.028	-.028	.104	-.026
	.252*	.110	.200*	.051	.085	.054	-.243*	.099	-.218*
Executive	-.148	.149	-.151	-.174*	.086	-.183*	.036	.112	.034
	-.073	.119	-.069	.032	.091	.040	-.151	.104	-.163
Judicial	.065	.130	.066	.057	.086	.055	-.012	.111	-.011
	-.035	.097	-.031	.000	.074	.000	.148 ⁺	.087	.149 ⁺
Monarchic	.097	.105	.105	-.042	.060	-.050	-.042	.077	-.047
	.157*	.079	.169*	-.049	.062	-.069	.017	.071	.020
Hierarchical	.106	.131	.096	-.086	.083	-.082	-.149	.108	-.134
	-.147	.100	-.121	-.137 ⁺	.078	-.149 ⁺	-.223*	.090	-.202*
Oligarchic	-.085	.107	-.087	.205**	.073	.209**	.057	.097	.052
	.036	.077	.035	-.004	.059	-.006	-.059	.070	-.062
Anarchic	.054	.134	.047	.081	.087	.071	.307**	.113	.251**
	.145	.104	.114	.097	.078	.102	.128	.093	.111
Global	-.042	.136	-.037	-.095	.080	-.091	-.089	.105	-.075
	-.002	.096	-.002	-.033	.077	-.036	.090	.086	.084
Local	-.306**	.110	-.334**	-.129 ⁺	.071	-.140 ⁺	-.087	.089	-.086
	-.044	.087	-.043	.010	.069	.013	-.107	.078	-.113
Internal	-.120	.122	-.121	.053	.088	.052	-.180	.119	-.159
	-.261*	.100	-.231*	-.159*	.079	-.184*	.187 ⁺	.095	.180 ⁺
External	.044	.099	.046	-.076	.067	-.081	-.172*	.083	-.167*
	-.037	.077	-.040	.012	.059	.017	-.123 ⁺	.068	-.149 ⁺
Liberal	.123	.138	.139	-.063	.094	-.067	-.130	.120	-.130
	-.254*	.122	-.237*	.039	.094	.047	-.121	.112	-.123
Conservative	.170	.139	.182	.094	.088	.095	.022	.111	.020
	-.198	.130	-.182	.173 ⁺	.099	.209 ⁺	.005	.121	.005
R ²	.150/.103*			.112**/.080			.187**/.172**		
adjusted R ²	.040/.047			.067/.022			.129/.117		

Note: The first value belongs to Study I, and the second to Study II.

** $p < .01$; * $p < .05$; ⁺ $p < .10$

5. Study II: Confirming and Optimizing the Factor Structure and Psychometric Properties of the TSI-GER

5.1 Participants and Context

389 students from two universities in Germany and one university in Austria participated in the study (207 bachelors and 180 masters students, 2 missing cases). They were between 18 and 46 years old ($Mdn = 24$; $M = 24.92$; $SD = 4.17$) and mostly female (71.2%; 3 missing cases).

5.2 Measures

Thinking Styles. The questionnaire comprises the 60 items from Study I and 24 items that were additionally formulated.

Achievement. Grades in mathematics, German, and economics were collected to test the criterion validity with reference to achievement.

5.3 Analysis

A CFA is applied to test the factor structure from Study I and to investigate the reliability/validity of the questionnaire. All 13 thinking styles' scales are simultaneously analyzed. Criterion validity to achievement is investigated by means of multiple linear regression analysis.

5.4 Results

An analysis of the data shows that the absolute univariate skew does not exceed .841, and the absolute univariate kurtosis is not more than 1.218. They are within the range for a moderate violation of the normal distribution ($|\text{skew}| < 2$, $|\text{kurtosis}| < 7$) and allow the application of the maximum likelihood algorithm for CFA (West, Finch, & Curran, 1995). CFA results ($\chi^2(1632) = 3612.508$, $p < .01$; RMSEA .056; SRMR .082; CFI .797) indicate a global model fit for the item solution from Study I (Hu & Bentler, 1999). To further improve the questionnaire, 24 new items are added and gradually removed if they do not contribute to model fit or the reliability of the scales. At the end of this process, only one added item for the local style remained. CFA proves a global model fit for the optimized 61-item solution ($\chi^2(1691) = 3704.520$, $p < .01$; RMSEA .055; SRMR .081; CFI .799).

Table 1 reports Cronbach's α , the composite reliability, the average variance extracted, and the retest reliability for Study II.

Cronbach's α ranges between .592 for the anarchic style to .900 for the liberal style. Composite reliability exceeds the critical value of .600 recommended by Bagozzi and Yi (1988) for all scales ($Min = .608$, $Max = .900$). However, only the local, external, and liberal style show an average variance extracted above the cutoff value of .500 (Bagozzi & Yi, 1988). Item reliabilities vary between .078 and .737 ($Mdn = .439$), and 35 items show values above .400. But only the external style scale consists completely of items whose variance is at least explained by 40% of the underlying factor. Test-retest reliability is calculated for a sub-sample of 61 students and ranges from .639 (executive style) to .859 (local style).

Table 4 reports the inter-scale correlations, squared inter-scale correlations, average variance extracted, and the average absolute correlations for the 13 thinking styles' scales (via Fisher-Z transformation). 71 out of 78 thinking style pairs are in line with the strict Fornell-Larcker criterion. Problems with discriminant validity involve the legislative and liberal, legislative and internal, executive and conservative, judicial and liberal, anarchic and local, anarchic and liberal, and the monarchic and hierarchical styles. To test the discriminant validity with a less strict criterion, Wald-chi-squared tests are performed, testing the hypothesis that the correlation between two thinking styles' pairs equals one. Results for the legislative and liberal ($\chi^2(1) = 53.981$, $p < .01$), legislative and internal ($\chi^2(1) = 9.187$, $p < .01$), executive and conservative ($\chi^2(1) = 6.543$, $p < .05$), judicial and liberal ($\chi^2(1) = 63.681$, $p < .01$), anarchic and local ($\chi^2(1) = 62.212$, $p < .01$), anarchic and liberal ($\chi^2(1) = 14.353$, $p < .01$), and monarchic and hierarchical styles ($\chi^2(1) = 42.039$, $p < .01$) reject the hypothesis, indicating no perfect relationship between the thinking styles' pairs. Different but perhaps overlapping constructs can therefore be assumed.

Finally, the criterion validity when it comes to achievement is investigated by means of multiple linear regression analysis. Table 4 presents the results. As can be seen, students favoring the liberal and internal styles, and rejecting the legislative and monarchic styles are more successful in mathematics, whereas learners preferring the legislative and hierarchical styles show higher achievement scores in economics. For German, only the regression coefficient concerning the internal style becomes significant. The thinking styles explain only 4.7% of the variances in mathematics and only 2.2% of the differences in German. For grades in economics, they have more explanatory power as the adjusted R^2 of 11.7% indicates. The following section discusses the results of both studies while factoring in the issues outlined at the beginning of the paper.

Table 4. Inter-Scale Correlations, Squared Inter-Scale Correlations, and Average Variance Extracted for the 13 Thinking Styles' Scales from Study II

	Legislative	Executive	Judicial	Global	Local	Liberal	Conservative	Hierarchical	Monarchic	Oligarchic	Anarchic	Internal	External
Legislative	.417	.039	.353	.103	.089	.461	.066	.050	.033	.018	.120	.699	.009
Executive	-.198**	.495	.015	.033	.000	.276	.903	.306	.256	.001	.001	.003	.007
Judicial	.594**	-.122	.376	.108	.132	.433	.093	.064	.025	.000	.150	.071	.181
Global	.321**	.181**	.329**	.371	.327	.017	.017	.073	.052	.011	.002	.037	.019
Local	.298**	-.020	.364**	-.572**	.514	.119	.004	.004	.004	.000	.297	.138	.000
Liberal	.679**	-.525**	.658**	.132	.345**	.529	.436	.000	.003	.007	.246	.127	.073
Conservative	-.257**	.950**	-.305**	.131	-.062	-.660**	.476	.193	.181	.009	.005	.002	.000
Hierarchical	.224**	.553**	.253**	.270**	.063	-.009	.439**	.311	.552	.020	.015	.068	.043
Monarchic	.181**	.506**	.159*	.229**	.061	-.050	.425**	.743**	.480	.030	.016	.061	.033
Oligarchic	-.133*	.036	-.011	.103	.014	.082	.096	-.142*	-.174**	.469	.134	.017	.027
Anarchic	.346**	-.030	.387**	-.043	.545**	.496**	-.069	-.121	-.128	.366**	.208	.130	.030
Internal	.836**	.059	.267**	.193**	.372**	.356**	.045	.260**	.246**	-.132*	.360**	.347	.112
External	.097	.083	.425**	.138*	.018	.271**	.022	.208**	.181**	.163**	.173*	-.334**	.599

Note: Values below the diagonal: inter-scale correlations; Values above the diagonal: squared inter-scale correlations; Values on the diagonal: average variance extracted.

* $p < .05$, ** $p < .01$

6. General Discussion and Conclusion

6.1 Reliability of the TSI

In both studies, most of the thinking styles' scales reveal high and stable values for Cronbach's α of about at least .700, following the recommendations by Adams and Lawrence (2015). Problematic are the anarchic, global, local, and internal thinking styles. Their reliability is low in at least one study. Problems with low Cronbach's α values for the monarchic, anarchic, and local styles were the reasons for the two revisions of the original Thinking Styles Inventory (Zhang, 2004b; Zhang & He, 2011). Thus, the German version of the TSI shows already-known strengths and weaknesses.

Except for the global style in Study I, all scales are in line with the minimum requirements of .600 for the composite reliability (Bagozzi & Yi, 1988), indicating that the item compositions reliably assess their corresponding thinking style. Focusing on single items however reveals less-than-optimal results. Based on Bagozzi and Baumgartner (1994), item reliabilities should exceed 40%. In Study I these values range between .127 and .628 with a central tendency below .400 ($Mdn = .381$). In Study II, better values could be achieved, ranging from .078 to .737 with a central tendency above .400 ($Mdn = .439$). In Study I only the liberal scale and in Study II only the external scale consist completely of items with sufficient item reliability. As a consequence, the TSI-GER still contains items contributing only slightly to the assessment of specific thinking styles. Taking into account that most studies focus on a factor level that fails to provide information on the item reliabilities (e.g., Fjell & Walhovd, 2004; Cano-Garcia & Hughes, 2000; Zhang & Higgins, 2008), this finding leads to the conclusion that the Thinking Styles Inventories may consist of less desirable items. Further studies should therefore concentrate on the item reliabilities of the TSI, TSI-R, and TSI-R2 in order to separate preferable items from less preferable items in an overall attempt to improve these instruments.

Problematic is also the average variance extracted because only the local, external, and liberal scales in Study II show values above the cutoff value of .500 (Bagozzi & Yi, 1988). Due to the fact that most of the studies do not report the average variance extracted, it remains unclear whether this is a problem only for the German version of the TSI, or a more overall weakness of the Thinking Styles Inventories.

Furthermore, some of the retest reliabilities are quite low ($r < .700$) in terms of the executive, judicial, oligarchic, anarchic, and conservative styles, although this finding is in line with the theoretical foundation of the thinking styles formulated within the Theory of Mental Self-Government. Sternberg (1997) in fact postulates a general malleability of the thinking styles (see also Zhang, 2013).

6.2 Validity of the TSI

Whereas the reliability of the TSI-GER seems to be improvable and therefore less challenging, the validity of the instrument is more problematic. Understanding validity as the degree of correspondence between data and theoretical considerations (Lounsbury, Gibson, & Saudargas, 2006), the global and local model fit within CFA indicate a valid measurement model. A more detailed analysis of the data, however, questions this judgment.

In both studies, at least three pairs of thinking style scales are not in line with the strict Fornell-Larcker criterion, complying only with the less strict Wald-chi-squared tests. Furthermore, according to Cohen (1988), some of the correlations between the thinking styles' scales are extremely high ($r \geq .50$). In both studies, strong relationships between the legislative and judicial ($r = .532$, $r = .594$), and monarchic and hierarchic styles ($r = .609$, $r = .743$) could be identified. An inspection of the corresponding *item compositions* for the legislative and judicial styles shows similarities since the items of both styles address the degrees of freedom in decision making and working on tasks. In addition, both the monarchic and hierarchical styles are characterized by a strong preference for pursuing a specific goal. Dai and Feldhusen (1999) also report medium to strong correlations for both pairs, whereas Black and McCoach (2008) report only one medium to strong relationship for the monarchic and hierarchical styles. Correlations reported by Zhang (1999) are generally lower for these two pairs of thinking styles. In Study II, the correlation between the executive and conservative styles is .950, between the legislative and internal styles .836, and between the legislative and liberal styles .679. Correlations of at least .50 are also reported by Black and McCoach (2008), Dai and Feldhusen (1999), and Zhang (1999). An inspection of the item compositions for these thinking styles' pairs shows high similarities. The legislative and liberal styles share high degrees of freedom in working on tasks. The executive and the conservative styles have in common the reliance on well-established rules and methods. Finally, the preference for relying on one's own points of view is a similarity between the legislative and internal styles. This is why these results generate doubt about whether the scales actually assess different constructs. The high correlations may indicate the existence of superordinate factors combining these different thinking styles into one single construct. These findings also raise questions about the number of constructs assessed by the TSI and postulated in the Theory of Mental Self-Government. Either the operationalization of the TSI needs further revisions in order to differentiate between 13 thinking styles, or the theory has to be altered to fit its respective data requirements. With this in mind, it's plausible enough to expect less than 13 thinking styles.

On the one hand, some of the thinking styles predict achievement in mathematics, German, and economics. On the other hand, no stable pattern of predictors could be identified across the two studies. This basically means that the criterion validity of the TSI-GER is problematic. One reason for this could be the partly low retest reliabilities leading to varying assessments of the thinking styles. As mentioned in the reliability section of this paper, this is in line with the theoretical postulation concerning the malleability of thinking styles. However, the lack of stable patterns questions the usefulness of a questionnaire as an instrument of practical findings. After all, teaching practice cannot be derived from significant results this way. If thinking styles are in fact as flexible and malleable as the Theory of Mental Self-Government postulates at various points (e.g., Sternberg, 1997; Zhang, 2013), the evaluation and measurement of thinking styles would require other instruments and research designs. Instead of focusing on different domains of knowledge as investigated in several studies (e.g., Cano-Garcia & Hughes, 2000; Zhang, 2001; 2004a), research should focus on the advantageousness and stability of thinking styles in specific (classes of) activities and tasks.

6.3 Conclusions

Summing up, most studies focus on a factor level of the Thinking Styles Inventories, ignoring the relationships between the thinking styles and the corresponding items. The current paper closes this gap by providing a detailed insight into the psychometric properties of specific items and the measurement model of a German version of the Thinking Styles Inventories. The results indicate some problems with the structural, discriminant, and criterion validity of the questionnaire. In particular, high correlations between some thinking styles could be replicated and are in line with findings from other investigations (e.g., Black & McCoach, 2008; Dai & Feldhusen, 1999; Zhang, 1999). Due to the lack of studies focusing on the quality of the measurement model, contradictory assumptions of Sternberg's (1997) theory and factor analyses on a factor level, and less-than-optimal findings in the current study, researchers should use the Thinking Styles Inventories with caution. As the measurement instrument forms the basis

for all empirical research, knowledge generated to date using the Thinking Styles Inventories should be closely examined for its trustworthiness. This is one of the key reasons why the current paper emphasizes the need to develop other measures that assess thinking styles more validly and/or even modify the Theory of Mental Self-Government.

References

- Adams, K. A., & Lawrence, E. K. (2015). *Research methods, statistics, and applications*. Thousand Oaks, CA: Sage.
- Bagozzi, R. P., & Baumgartner, H. (1994). The evaluation of structural equation models and hypothesis testing. In R. P. Bagozzi (Ed.), *Principles of marketing research* (pp. 386–422). Cambridge, MA: Blackwell Business.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94. <https://doi.org/10.1007/BF02723327>
- Black, A. C., & McCoach, D. B. (2008). Validity study of the Thinking Styles Inventory. *Journal for the Education of the Gifted*, 32(2), 180–210.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.
- Cano-Garcia, F., & Hughes, E. H. (2000). Learning and thinking styles: An analysis of their interrelationship and influence on academic performance. *Educational Psychology*, 20(4), 413–430. <https://doi.org/10.1080/713663755>
- Chen, N.-S., Kinshuk; Wei, C.-W., & Liu, C.-C. (2011). Effects of matching teaching strategy to thinking style on learner's quality of reflection in an online learning environment. *Computers & Education*, 56(1), 53–64. <https://doi.org/10.1016/j.compedu.2010.08.021>
- Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dai, D. Y., & Feldhusen, J. F. (1999). A validation study of the Thinking Styles Inventory: Implications for gifted education. *Roper Review*, 21(4), 302–307. <https://doi.org/10.1080/02783199909553981>
- Fan, J. (2014). *Thinking styles' socialization and their roles in student development*. Retrieved May 19, 2015, from the University of Hong Kong Web site: <http://hdl.handle.net/10722/196455>
- Fer, S. (2005). Validity and reliability of the Thinking Styles Inventory. *Educational Sciences: Theory & Practice*, 5(1), 55–68.
- Fjell, A. M., & Walhovd, K. B. (2004). Thinking styles in relation to personality traits: An investigation of the Thinking Styles Inventory and NEO-PI-R. *Scandinavian Journal of Psychology*, 45(4), 293–300. <https://doi.org/10.1111/j.1467-9450.2004.00408.x>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobserved variables and measurement errors. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, B. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological methods*, 16(3), 319–336. <https://doi.org/10.1037/a0024917>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kaplan, R. M., & Saccuzzo, D. P. (2013). *Psychological testing: Principles, applications, and issues* (8th ed.). Belmont, CA: Wadsworth.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.
- Lounsbury, J. W., Gibson, L. W., & Saudargas, R. A. (2006). Scale Development. In F. T. L. Leong, & J. T. Austin (Eds.), *The Psychology Research Handbook* (pp. 125–146). Thousand Oaks, CA: Sage.
- Murphy, A., & Janeke, H. C. (2009). The relationship between thinking styles and emotional intelligence: An exploratory study. *South African Journal of Psychology*, 39(3), 357–375. <https://doi.org/10.1177/008124630903900310>
- Ngan Man-Fong, O. (2013). Do local and international school students in Hong Kong have different thinking styles?

- Retrieved May 19, 2015, from the University of Hong Kong Web site: <http://hdl.handle.net/10722/198882>
- Richmond, A. S., & Conrad, L. (2012). Do thinking styles predict academic performance of online learning? *International Journal of Technology in Teaching and Learning*, 8(2), 108–117.
- Sternberg, R. J. (1994). Thinking styles: Theory and assessment at the interface between intelligence and personality. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 169–187). Cambridge, UK: Cambridge University Press.
- Sternberg, R. J. (1997). *Thinking styles*. New York, NY: Cambridge University Press.
- Sternberg, R. J., & Wagner, R. K. (1992). *Thinking Styles Inventory*. Unpublished test, Yale University.
- Sternberg, R. J., Wagner, R. K., & Zhang, L. F. (2007). *Thinking Styles Inventory – Revised II*. Unpublished test, Tufts University.
- Wang, J., & Wang, X. (2012). *Structural equation modeling*. Chichester, UK: Wiley.
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables. In R. H. Hoyle (Ed.), *Structural equation modeling* (pp. 56–75). London: Sage.
- Zhang, L. F. & Sternberg, R. J. (2006). *The nature of intellectual styles*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Zhang, L. F. (1999). Further cross-cultural validation of the theory of mental self-government. *The Journal of Psychology*, 133(2), 165–181. <https://doi.org/10.1080/00223989909599731>
- Zhang, L. F. (2001). Do thinking styles contribute to academic achievement beyond self-rated abilities? *The Journal of Psychology*, 135(6), 621–637. <https://doi.org/10.1080/00223980109603724>
- Zhang, L. F. (2002). Thinking styles: Their relationship with modes of thinking and academic performance. *Educational Psychology*, 22(3), 331–348. <https://doi.org/10.1080/01443410220138557>
- Zhang, L. F. (2003). Contributions of thinking styles to critical thinking dispositions. *The Journal of Psychology*, 137(6), 517–544. <https://doi.org/10.1080/00223980309600633>
- Zhang, L. F. (2004a). Revisiting the predictive power of thinking styles for academic performance. *The Journal of Psychology*, 138(4), 351–370. <https://doi.org/10.3200/JRLP.138.4.351-370>
- Zhang, L. F. (2004b). Thinking styles: University students' preferred teaching styles and their conceptions of effective teachers. *The Journal of Psychology*, 138(3), 233–252. <https://doi.org/10.3200/JRLP.138.3.233-252>
- Zhang, L. F. (2005). Validating the theory of mental self-government in a non-academic setting. *Personality and Individual Differences*, 38(8), 1915–1925. <https://doi.org/10.1016/j.paid.2004.11.009>
- Zhang, L. F. (2006). Thinking styles and the big five personality traits revisited. *Personality and Individual Differences*, 40(6), 1177–1187. <https://doi.org/10.1016/j.paid.2005.10.011>
- Zhang, L. F. (2008). Thinking styles and emotions. *The Journal of Psychology*, 142(5), 497–515. <https://doi.org/10.3200/JRLP.142.5.497-516>
- Zhang, L. F. (2009). Anxiety and thinking styles. *Personality and Individual Differences*, 47(4), 347–351. <https://doi.org/10.1016/j.paid.2009.04.001>
- Zhang, L. F. (2013). *The malleability of intellectual styles*. New York, NY: Cambridge University Press.
- Zhang, L. F., & He, Y. F. (2011). Thinking styles and the Eriksonian Stages. *Journal of Adult Development*, 18(1), 8–17. <https://doi.org/10.1007/s10804-010-9101-z>
- Zhang, L. F., & Higgins, P. (2008). The predictive power of socialization variables for thinking styles among adults in the workplace. *Learning and Individual Differences*, 18(1), 11–18. <https://doi.org/10.1016/j.lindif.2007.03.002>
- Zhang, L. F., & Sternberg, R. J. (2000). Are learning approaches and thinking styles related? A study in two Chinese populations. *The Journal of Psychology*, 134(5), 469–489. <https://doi.org/10.1080/00223980009598230>
- Zhang, L. F., & Sternberg, R. J. (2005). A threefold model of intellectual styles. *Educational Psychology Review*, 17(1), 1–53. <https://doi.org/10.1007/s10648-005-1635-4>
- Zhang, L. F., & Sternberg, R. J. (2009). Revisiting the value issue in intellectual styles. In L. F. Zhang, & R. J. Sternberg (Eds.), *Perspectives on the nature of intellectual styles* (pp. 63–85). New York, NY: Springer Publishing Company.