3. Aspects of the construct validity of the MLQ

In this chapter, some aspects of the construct validity of the MLQ will be examined. As such an examination can be divided into several facets, this chapter is divided into several parts as well. Cronbach and Meehl (1955), for example, propose five methods in order to examine the construct validity: the examination of group differences, an analysis of correlation matrices and factorial structures, studies of the internal structure, studies of change over occasions, and studies of process. Lienert and Raatz (1998) extend this proposal to seven aspects: the correlation with several external criteria, the correlation with tests that have a similar claim of validity (i.e. that measure the same construct), the correlation with tests that measure different constructs, a factorial analysis with external criteria as well as with tests that measure the same or a different construct, the analysis of interindividual differences, the analysis of interindividual changes in re-tests, and a content-logical analysis of the test elements.

Four aspects of the construct validity will be the center of this investigation. First of all, the factorial validity of the MLQ will be examined. Then a look at the relation to tests with a similar claim of validity, namely leadership behavior, will follow. And finally the relation to external criteria and individual differences will be discussed. The examined facets of the MLQ's construct validity will be presented in sub-chapters, every sub-chapter will have a short theoretical introduction that will be followed by the particular assumptions. The parts of this chapter examine the factorial validity of the MLQ, the convergent validity concerning alternative leadership constructs, and the convergent and discriminant validity dealing with external criteria. Although this theoretical introduction might be redundant when taking the introduction in chapter 2 into account, it was considered to be much more reader-friendly. After a description of the samples and the methods used, the results will be presented. In order to provide a proximal evaluation of the results, every part will end with a discussion. However, in order to appreciate all the results obtained, an integral discussion will succeed this chapter.

According to Lienert and Raatz (1998), the factorial structure of a test or questionnaire describes its factorial validity and is as such one part of the construct validity. As the factor structure of the MLQ has often been at the center of investigation, this facet of the MLQ's construct validity will be the first part of this investigation.

3.1 Part 1: Factorial Validity

Even though the MLQ has undergone a number of modifications and refinements since its first publication in 1985, the factor structure is still not satisfyingly confirmed. On the conceptual and the empirical side criticism arises because the allocation of MbEp as a more reactive behavior that does not imply any form of exchange to the transactional scales is not comprehensible. A consecutive empirical validation is still pending. The aforementioned results by Bass and Avolio (1995) show that from the one-factor to the nine-factor model the fit indices improved. According to the authors, only the nine-factor model met the criteria set in the literature ($\chi^2 = 2394$, $df = 558$, $p < .0001$, *RMSR* = .04, *AGFI* = .89, Bass & Avolio, 1995, p. 25). Avolio et al. (1999) examined different theoretical and empirically based factor models with the MLQ 5X and found a seven-factor model with a reduced number of transformational scales that includes three transformational scales, CR, MbEa, MbEp, and LF showed the best fit indices: $\chi^2 = 2497$, $df = 573$, $p < .0001$, *RMSR* = .04, *AGFI* = .90 (p. 451). Despite the good fit of those models that divide the transformational factor into several subscales, the general problem of the discriminant validity of the transformational scales and CR persists (the average correlation among the transformational scales in this sample was $r = .81$).

The high correlations among the transformational scales are, as seen above, often addressed in the literature – correlations between $r = .48$ and $r = .90$ are reported (Avolio et al., 1999; Den Hartog et al., 1997). The four I's of transformational leadership (Bass & Avolio, 1994) cannot be confirmed as four independent scales (Vandenberghe et al., 2002). CR shows correlations with the transformational scales as high as is shown by these amongst each other (Tejada et al., 2001; Vandenberghe et al., 2002). Even with regard to the correlation with several external criteria the transformational scales cannot be distinguished from CR (Lowe et al., 1996; Vandenberghe et al., 2001). Wofford et al.'s (1998) reanalysis of the affiliation of CR to either the transformational or the transactional construct might be one step in the right direction of grasping the problem of the position of the transactional scales within the frame of the full range of leadership. Due to the high correlations and the lack of discriminant validity of the transformational scales and CR, the theoretical underlying construct has been questioned and reexamined again and again as can be seen in chapter 2.3.2.

Some of the findings described above indicate that a reduced and modified set of factors might provide a better representation of the MLQ structure. For the German-speaking part especially the Geyer and Steyer (1998) results and the results of the PCA conducted by Felfe (2005) should be mentioned in support of this proposal. Furthermore, the results dealing with the correlation pattern of the MLQ point out that a reduced set of factors might provide a better representation of the underlying structure of the MLQ.

Thus, the following can be expected:

According to other studies (see above), I assume that there will be high positive correlations amongst the transformational scales and between the transformational scales and CR. On the other hand, MbEp and LF show a high positive correlation.

Therefore:

1. The fit of the confirmatory factor analysis for the structure of the MLQ as proposed by Bass and Avolio (1995) will not meet the criteria.

Based upon the correlation clusters mentioned above, I assume that a reduced set of factors will provide a better description of the leadership facets assessed with the MLQ.

2a. The transformational scales will associate with CR and break down into a reduced set of transformational factors. MbEp will be combined with LF and MbEa will build its own factor.

2b. CFA of a model with less correlated factors will meet the criteria for a good fit.

3.1.1. Method

Sample

This part of the investigation is based on three samples (examined over the period 2000 to 2004) from several public administration offices. All three studies were integrated in an overall organizational analysis, conducted in order to accompany a reform process in the organizations. Although the sample included several work areas (tax and registration offices), the organizational structure of the public administrations is comparable as there is a high bureaucratic status in all German public administrations. Altogether there were 2840 participants included in the analyses (Study 1: N_1 = 1311; Study 2: N_2 = 879; Study 3: N_3 = 650). The mean age of the entire composite sample is 37.8 years (*SD* =

9.9 years; for the age and gender distribution see Table 26). It is obvious that the three groups are equal with respect to age and gender. Furthermore, the occupational status of the three samples is also comparable. Only approximately 8% of the participants were working in their department for less than 5 years. In all three samples 7-8% of the participants had leadership functions. Rejection rates were 53.2% (Study 1), 63.5% (Study 2) and 55.3% (Study 3).

Table26. Age and gender distribution (percent values) for Study 1, Study 2, and Study 3

	Study 1 (N_1 = 1311)			**Study 2 (N_2 = 879)**			**Study 3 (N_3 = 650)**		
Age in years	**Total**	**Female**	**Male**	**Total**	**Female**	**Male**	**Total**	**Female**	**Male**
up to 25 years	2.3	1.9	3.2	1.9	1.7	2.3	1.7	1.2	2.7
26-35 years	25.2	26.0	23.1	24.5	25.3	23.2	22.7	26.0	17.0
36-45 years	38.9	39.9	37.2	39.9	39.6	40.3	35.5	37.3	32.8
46-55 years	24.4	24.4	24.7	21.9	22.3	20.8	26.6	24.0	30.6
56 years and older	8.8	7.6	11.7	11.8	11.0	13.4	13.4	11.5	17.0
Total		71.1	28.9		65.8	34.2		64.5	35.5

Measures

In the context of the first two studies the MLQ (Version 5X Short, Bass & Avolio, 1995a; translated and adapted by Felfe & Goihl, 2002b) was employed. According to Felfe and Goihl (2002b), the instrument comprises nine leadership scales and two second order factors. The nine scales and their allocation are described above. Each scale is measured with four items on a five-point Likert response scale ranging from *not at all* (1) to *frequently, if not always* (5). Although Felfe and Goihl (2002b) eliminate items 1, 5, and 23, these items are included in the present studies to assure a maximum degree of comparability with other studies. Moreover, the Laissez-Faire items are formulated according to the original American version (Bass & Avolio, 1995a).

All in all, the scales have a satisfactory internal consistency. Only MbEa does not meet the criterion of α = .70 ($\alpha_{Study\ 1}$ = .62 resp. $\alpha_{Study\ 2}$ = .65, see Table 27). In Study 3, a

reduced set of items, derived from the MLQ with respect to the results of Study 2, was used. The new questionnaire comprises 13 leadership items (for more details see below). The scales are represented by unweighted means of their items, as is recommended by Bass and Avolio (1997).

The scale intercorrelations for the transformational scales range from r = .65 to r = .82 in Study 1 and from r = .63 to r = .80 in Study 2. The correlations of CR with the transformational scales range from r = .69 to r = .76 in Study 1 and from r = .72 to r = .77 in Study 2 and are in parts higher than those of the transformational scales (see Table 27).

Table 27. Means, internal consistencies and scale intercorrelations for the MLQ scales

	Study 1 (N_1 = 1311)			Study 2 (N_2 = 879)											
Scales	***M***	***SD***	***α***	***M***	***SD***	***α***	**IIA**	**IIB**	**IM**	**IS**	**IC**	**CR**	**MbEa**	**MbEp**	**LF**
IIA	2.82	1.00	.84	2.79	1.02	.87	-	*.74*	*.66*	*.78*	*.79*	*.75*	*.18*	*-.61*	*-.64*
IIB	2.92	.82	.70	2.92	.90	.77	*.68*	-	*.80*	*.73*	*.70*	*.77*	*.24*	*-.51*	*-.58*
IM	2.89	.80	.75	2.79	.92	.81	*.66*	*.72*	-	*.66*	*.63*	*.74*	*.20*	*-.44*	*-.49*
IS	3.02	.88	.84	3.00	.93	.84	*.78*	*.69*	*.67*	-	*.78*	*.73*	*.25*	*-.61*	*-.64*
IC	2.96	.96	.84	2.83	.94	.76	*.82*	*.69*	*.65*	*.80*	-	*.72*	*.13*	*-.53*	*-.58*
CR	3.09	.87	.73	3.01	.84	.70	*.74*	*.75*	*.69*	*.76*	*.76*	-	*.27*	*-.48*	*-.58*
MbEa	2.89	.71	.62	3.00	.77	.65	*.09*	*.25*	*.15*	*.17*	*.06*	*.21*	-	*-.12*	*-.16*
MbEp	2.56	.89	.77	2.52	.94	.80	*-.56*	*-.39*	*-.42*	*-.57*	*-.54*	*-.48*	*.05*	-	*.65*
LF	2.71	.82	.79	2.57	.88	.71	*-.68*	*-.55*	*-.53*	*-.71*	*-.66*	*-.63*	*-.14*	*.57*	-

Note: IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire The correlations of Study 1 are displayed in the lower diagonal, the correlations of Study 2 in the upper diagonal; r > |.05|: p < .05, r> |.08|: p < .01

3.1.2. Results

Study 1

Study 1 (N_1 = 1311) was used in order to replicate the nine dimensions according to Bass and Avolio (1995). To examine the underlying structure a confirmatory factor analysis (LISREL 8.3; Jöreskog & Sörbom, 1999) was computed. This method is considered best for examining the factorial structure of an instrument, if an adequately large sample can be assured (Hoyle, 1995). Prior to this, as basis for the decision of the quality of the congruence of the theoretical model and the empirical data, the criteria for an acceptable fit were set: the comparative-fit index (*CFI*; Bentler, 1990), the goodness-of-fit index (*GFI*; Jöreskog & Sörbom, 1998) and the adjusted goodness-of-fit index (*AGFI*) should be > .90 Bentler, 1990), the root-mean-square error of approximation (*RMSEA,* Steiger, 1990) and the root mean square residual (*RMR*) < .05. Furthermore, a non-normed fit index (*NNFI*; Bollen, 1989; Byrne, 2001) > .90 was expected. These parameters meet the recommendations made in the literature (Kline, 1998; Medsker et al., 1994; Steiger, 1990). Starting point of the analyses was a covariance matrix, the factors were assumed to be correlated as the aforementioned American and German studies suggest high correlations between most of the scales. Furthermore, the error terms were assumed to be uncorrelated and a congeneric allocation of the items on their respective factor was assumed (Kline, 1998).

In addition to the nine-factor model, several other models that were already mentioned in the literature (see chapter 2.3.2) were tested in order to examine if one of those proposed models would show an acceptable fit for the data. Twelve alternative models that were plausible either theoretically or empirically were chosen (see Table 28).

Table 28. Tested MLQ models – scale allocation

Model	**Contents**
One factor	One general leadership factor
Two correlated factors	Active: Transformational Scales, CR and MbEa Passive: MbEp and LF
Two correlated factors	Active-constructive: Transformational Scales and CR Passive-corrective: MbE and LF

Table 28 continued

Model	Contents
Three correlated factors	Transformational Transactional Laissez-Faire
Three correlated factors	Transformational Transactional: CR and MbEa Passive-avoidant: MbEp and LF
Four correlated factors	Transformational CR MbEa Passive-avoidant: MbEp and LF
Five correlated factors	Transformational CR MbEa MbEp LF
Six correlated factors	Charisma/inspiration: IIA, IIB, IM IS IC CR MbEa Passive-avoidant: MbEp and LF
Six correlated factors	Charisma/inspiration: IIA, IIB, IM IS IC CR MbE: MbEa, MbEp LF
Seven correlated factors	Charisma/inspiration: IIA, IIB, IM IS IC CR MbEa MbEp LF
Eight correlated factors	Charisma (IIA+IIB) IM IS IC CR MbEa MbEp LF

Table 28 continued

Model	Contents
Eight correlated factors	IIA
	IIB
	IM
	IS
	IC
	CR
	MbEa
	Passive-avoidant

Note: IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire

Table 29 shows the pattern of factor loadings for the model with nine correlated factors. The fit indices are: $\chi^2 = 4457$, $df = 558$, $p < 0.01$, $N_1 = 1311$, $AGFI = 0.78$, $CFI = 0.85$, $RMSEA = 0.073$ (see also Table 30). Expectation 1 is retained. All in all, although the internal consistencies of the scales are acceptable, the fit of the nine-factor model is unsatisfactory.

Table 29. Factor loadings of the MLQ items on nine latent variables using CFA$^+$ (completely standardized solution, 36 items, $N_1 = 1311$)

Scales:	IIA	IIB	IM	IS	IC	CR	MbEa	MbEp	LF
Item*	*MLQ_10*	*MLQ_6*	*MLQ_9*	*MLQ_2*	*MLQ_15*	*MLQ_1*	*MLQ_4*	*MLQ_3*	*MLQ_5*
	0.82	0.45	0.61	0.70	0.76	0.66	0.29	0.75	0.67
Item	*MLQ_18*	*MLQ_14*	*MLQ_13*	*MLQ_8*	*MLQ_19*	*MLQ_11*	*MLQ_22*	*MLQ_12*	*MLQ_7*
	0.66	0.67	0.62	0.73	0.64	0.46	0.48	0.79	0.69
Item	*MLQ_21*	*MLQ_23*	*MLQ_26*	*MLQ_30*	*MLQ_29*	*MLQ_16*	*MLQ_24*	*MLQ_17*	*MLQ_28*
	0.81	0.61	0.75	0.79	0.79	0.74	0.74	0.47	0.69
Item	*MLQ_25*	*MLQ_34*	*MLQ_36*	*MLQ_32*	*MLQ_31*	*MLQ_35*	*MLQ_27*	*MLQ_20*	*MLQ_33*
	0.89	0.70	0.64	0.77	0.85	0.67	0.59	0.68	0.75

$^+$Method of estimation: maximum likelihood; phi-matrix symmetric and free (oblique solution)
**The order of the items of the American version of the MLQ 5X Short (Bass & Avolio, 1995) is maintained.*

Taking a look at the other tested models, one can see that the fit of the models increases with the number of factors in the model. However, none of the tested models reaches

the criteria of a good fit. Furthermore, the increase of fit from the four to the nine-factorial model is not very big, e.g. the *GFI* increases from .78 to .82. The χ^2-difference test, on the other hand, shows a significant improvement in fit for the nine-factor model in comparison to all other models.

Table 30. CFA of several factorial models of the MLQ (N_1 = 1311)

Model	χ^2	df	GFI	AGFI	CFI	NNFI	RMSEA	$\Delta\chi^2$
(1)	26749.40	630	.16	.11	.00	.00	.178	
(2)	6307.37	594	.73	.70	.78	.77	.086	1849.59 (p < .01)
(3)	5576.01	593	.75	.72	.80	.79	.082	1118.23 (p < .01)
(4)	5794.55	593	.75	.72	.80	.79	.082	1336.77 (p < .01)
(5)	5775.45	591	.75	.72	.80	.79	.082	1317.67 (p < .01)
(6)	6078.03	591	.74	.70	.79	.78	.084	1620.25 (p < .01)
(7)	5533.33	588	.78	.75	.81	.80	.080	1075.55 (p < .01)
(8)	5221.52	584	.79	.76	.82	.81	.078	763.74 (p < .01)
(9)	5408.54	579	.76	.72	.82	.80	.080	950.76 (p < .01)
(10)	5326.03	579	.78	.75	.82	.80	.079	868.25 (p < .01)
(11)	5011.05	573	.79	.76	.83	.81	.077	553.27 (p < .01)
(12)	4867.01	566	.80	.76	.84	.82	.076	409.23(p < .01)
(13)	4774.49	566	.81	.77	.84	.82	.075	316.71 (p < .01)
(14)	4457.78	558	.82	.78	.85	.83	.073	

Note: (1) Null-Model; (2) one general factor;(3) two correlated factors (active/passive); (4) two correlated factors (active constructive/passive corrective); (5) three correlated factors (transformational/transactional/passive-avoidant); (6)three correlated factors (transformational/transactional/laissez-faire);(7) four correlated factors; (8) five correlated factors; (9) six correlated factors (charisma, i.e. IIA + IIB + IM, IS_1, IC, CR_1, MbE, LF); (10) six correlated factors (charisma, i.e. IIA + IIB + IM, IS_1, IC, CR_1, MbEa, passive-avoidant); (11) seven correlated factors;(12) eight correlated factors [charisma (IIA+IIB)/IM/IS/IC/CR/MbEa/MbEp/LF];(13) eight correlated factors (IIA/IIB/IM/IS/IC/CR/MbEa/passive-avoidant); (14) nine correlated factors; for a detailed description of the models see Table 28; GFI = goodness of fit index; AGFI = adjusted GFI; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; $\Delta\chi^2$ was computed by the subtraction of the χ^2 value of the nine-factor model from the χ^2-value of the other tested models.

Furthermore, taking a closer look at the factor intercorrelations (the factor intercorrelations for the six- and the nine-factor model are displayed in Table 31) one can see that they are still very high. For example, the mean correlation of the transformational factors in the nine-factor model is r = .89, the mean correlation of the

transformational scales and CR is with r = .95 even higher than the average correlation of the transformational scales.

Table 31. Factor intercorrelations of the nine-factor and the six-factor model (N_1 = 1311)

	IIA	IIB	IM	IS	IC	CR	MbEa	MbEp	MbE	LF
CHA				.93	.95	.99			.72	-.85
IIA	-									
IIB	.85	-								
IM	.82	.98	-							
IS	.90	.89	.84	-	.97	.98			.76	-.90
IC	.94	.89	.83	.97	-	.98			.72	-.84
CR	.92	.98	.93	.97	.97	-			.73	-.86
MbEa	.29	.49	.35	.38	.25	.49	-			
MbEp	-.72	-.56	-.59	-.72	-.69	-.69	-.15	-		
MbE									-	-.82
LF	-.86	-.75	-.72	-.90	-.84	-.85	-.35	.80		-

Note: CHA = Charisma (IIA+IIB+IM); IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; all correlations are significant, p <.01.

Recent research by Antonakis et al. (2003) suggests that the fit of the CFA is better for homogenous samples (see also chapter 2.3.2). For example, in their study the fit indices improved when the authors tested the female and male rater samples separately, investigation of other homogenous contexts also showed an improvement in the fit indices. The here presented sample is already quite homogenous, as in Study 1 only one public administration branch was examined. However, more homogenous sub-samples were tested. The sample therefore was subdivided into male and female participants as the perception of leadership shows gender biases (Sczesny, 2003). Furthermore, the sample was divided into people who recently entered the administration office (they were working there for less than five years) and people who had been working there for more than 20 years. The time of exposure to certain leadership behavior can change the way it is experienced and hence observed (Felfe, 2005). And thirdly, the administration has different work areas. Depending on the field of administration work, different

leadership styles could be recommended. Seven work areas were identified and two of them were chosen in order to examine the nine-factor model of the MLQ.

Table 32. CFA of the nine factor model in homogenous contexts (Study 1)

Sub-sample	N	χ^2	df	GFI	AGFI	CFI	NNFI	RMSEA
(1)	1311	4457.78	558	.82	.78	.85	.83	.073
(2)	927	3402.60	558	.81	.77	.85	.82	.074
(3)	376	1796.83	558	.77	.73	.83	.81	.077
(4)	120	1092.96	558	.68	.62	.77	.74	.090
(5)	325	1685.84	558	.76	.72	.83	.81	.079
(6)	470	2116.98	558	.78	.74	.83	.81	.077
(7)	244	1322.83	558	.75	.71	.86	.84	.075

Note: (1) complete sample; (2) female participants;(3) male participants; (4) working in the administration for less than 5 years; (5) working in the administration for more than 20 years; (6) work area 1;(7) work area 2; GFI = goodness of fit index; AGFI = adjusted GFI; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation.

The examination of more homogenous contexts, however, did not show an improvement in the fit indices for the nine-factor model in comparison to the complete sample (see Table 32). For all tested sub-samples, the fit indices remain below those of the nine-factor model and therefore also do not reach the criteria for a good fit.

An examination of the item intercorrelation shows that some items have higher loadings on other factors than their own, especially the transformational and CR items (see Appendix A).

In order to be responsive to these results, some of the proposed higher-order models (see chapter 2.3.2) were tested as well (see Table 33). Two models were chosen as basis for higher-order models, the six-factor and the nine-factor model. Based on the six-factor model, two models with two higher-order factors (one with uncorrelated and one with correlated higher-order factors) and one model with three uncorrelated higher-order factors were tested. Based on the nine-factor model, three models were also tested, two models with two higher-order factors, and one with three. In order to examine the explanation reached with the higher-order model for the covariation among the first-order factors, Marsh and Hocevar's (1985) target coefficient *T* was calculated.

Table 33. CFA for the MLQ data - higher order models (N_1 = 1311)

Model	χ^2	*df*	*GFI*	*CFI*	*NNFI*	*RMSEA*	*T*
Based on the six-factor model							
(1)	7054.33	594	.74	.75	.74	.091	.76
(2)	9332.96	594	.70	.67	.65	.106	.57
(3)	7049.69	593	.74	.75	.74	.091	.76
Based on the nine-factor model							
(4)	6123.86	585	.78	.79	.77	.085	.73
(5)	6133.65	584	.78	.79	.77	.085	.73
(6)	5276.44	584	.79	.82	.81	.078	.84

Note: (1) 2 uncorrelated higher-order factors (active constructive/passive corrective); (2) 3 uncorrelated higher order factors (transformational, developmental/transactional, passive-avoidant); (3) 2 correlated higher order factors (Charisma/Inspiration, developmental/transactional) and one uncorrelated higher-order factor (MbE + LF); (4) 2 uncorrelated higher-order factors (active constructive/passive corrective); (5) 2 correlated higher order factors (IIA + IIB + IM + IS_1, IC + CR_1) and one uncorrelated higher-order factor (MbE + LF); (6) 2 correlated higher order factors (transformational and transactional) and LF(correlated); GFI = goodness of fit index; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation, T = target coefficient.

As can be seen, none of the higher-order models reaches the criteria for a good fit. Furthermore, the target coefficient stays beneath the required cut-off of .90 (Marsh & Hocevar, 1985). The model with the relatively best fit, though, describes the theoretical foundation of the underlying theory. This model combines the transformational scales and the transactional scales into higher-order factors and has one first-order factor of LF.

Considering these results, the pattern of factor correlations and the misfit of the proposed theoretical models suggest another structure with a new allocation of items to factors and, altogether, less factors. Hence, a possible accurate estimation of the number of factors that can be extracted should be undertaken. Therefore, a parallel analysis, as one of the most promising methods to determine the number of factors (cf. Beauducel, 2001; Velicer, Eaton, & Fava, 2000), was conducted. For this parallel analysis, the eigenvalues of 500 solutions with 36 random Gaussian variables were averaged. Three factors clearly exceed the random solution eigenvalue, the fourth factor only moderately exceeds the random solution eigenvalue (empirical solution eigenvalue: 1.2447, *SD* = .0148, random solution eigenvalue: 1.2434; see Figure 04). If one takes the 95th percentile instead of the mean eigenvalue of the random solution (O'Connor, 2000),

then only three factors are suggested (for the third factor: empirical solution eigenvalue: 1.597, random solution eigenvalue: 1.227). Therefore, based on the results of these comparisons, three factors can be extracted from this item pool.

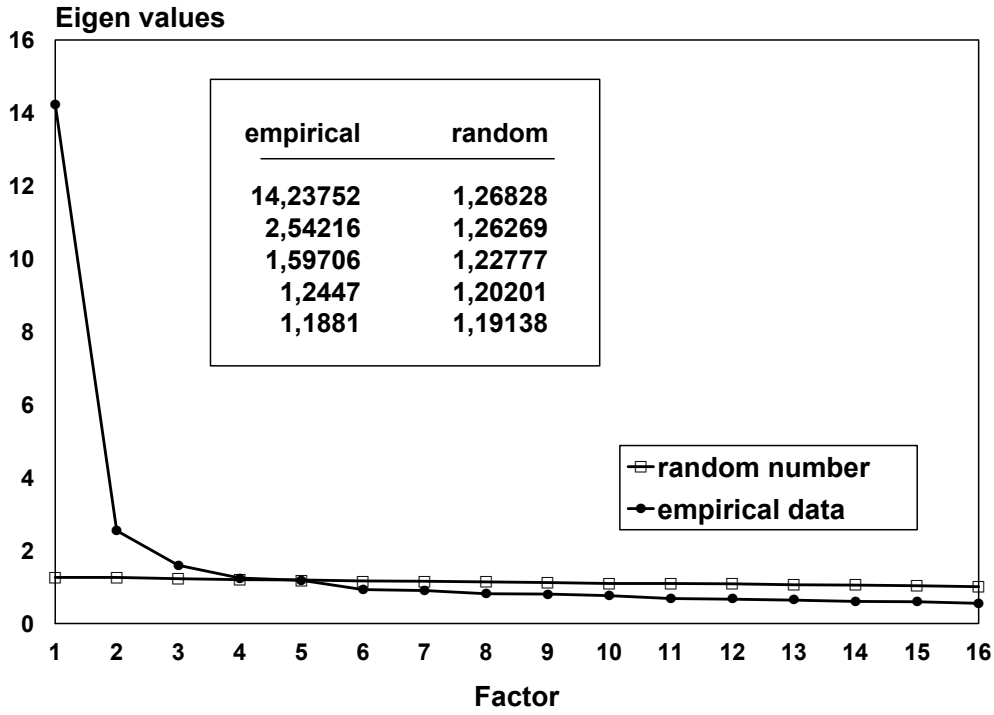

Figure 04. Parallel analysis with 36 items – screeplot for empirical (MLQ data, N_1 = 1311) and random data

Study 2

Taking this result, an empirical three-factor solution is the basis for the next step. The exploratory factor analysis [principle axis analysis - as this model is a common factor model with reflective indicators (Podsakoff, MacKenzie, Podsakoff, & Lee, 2003) - extraction of three factors, oblique solution] within a second, independent study (N_2 = 879) shows the structure presented in Table 34. Only loadings >.60 are taken into account. In order to get an unambiguous solution, items with high cross-loadings (> .20) are excluded.

The first factor contains six items of the former scales IIB, IM, and CR (two items each). The transformational scales IIA, IIB and IM are sometimes combined in one factor called Charisma/Inspiration (cf. Avolio et al., 1999) according to the first

operationalization of the transformational leadership theory (Bass, 1985). As IIA is not represented in this scale, only items that concern the leader behavior remain. The two CR items deal with future goals (clarifying rewards for goals that are met, and being satisfied if expectations are fulfilled). All in all, this factor combines future and goal oriented behavior, vision, and team orientation. Briefly summarized this would be charismatic, goal-oriented behavior; therefore this factor is called 'charismatic goal orientation'. The second factor, on the other hand, is based on four items and includes the former factors MbEp (three items) and LF (one item). It can be named 'passive-avoidant leadership' (cf. Antonakis et al., 2003; Avolio et al., 1999). MbEa is confirmed as a sovereign factor. Hence, due to the insufficient factor loading, item 22 is excluded.

Table 34. Pattern Matrix for the principle axis analysis with 36 items, factor loadings of the selected items (13) and factor loadings of the three latent variables on CFA (completely standardized solution)

		Principle axis analysis N_2 = 879			**Confirmatory factor analysis** N_3 = 650		
	Original MLQ Scale	**F I***	**F II***	**F III***	**F I***	**F II***	**F III***
MLQ_14	**IIB**	**.79**	.12	.06	0.60		
MLQ_26	**IM**	**.72**	.12	-.03	0.78		
MLQ_16	**CR**	**.69**	.10	-.05	0.69		
MLQ_9	**IM**	**.67**	-.01	.06	0.78		
MLQ_35	**CR**	**.63**	-.05	-.07	0.72		
MLQ_34	**IIB**	**.60**	.02	-.19	0.67		
MLQ_12	**MbEp**	.10	**.87**	.03		0.79	
MLQ_3	**MbEp**	.01	**.74**	.03		0.78	
MLQ_20	**MbEp**	.01	**.69**	-.08		0.73	
MLQ_5	**LF**	-.09	**.69**	.04		0.69	
MLQ_27	**MbEa**	.00	-.16	**.76**			0.64
MLQ_4	**MbEa**	-.12	.18	**.63**			0.66
MLQ_22	**MbEa**	.10	-.19	**.61**			0.63

** F I: charismatic goal orientation, F II: passive-avoidant leadership, F III: management by exception; IIB = idealized influence behaviour, IM = inspirational motivation, CR = contingent reward, MbEp = management by exception passive, LF = laissez-faire, MbEa = management by exception active.*

As this factor structure does not differentiate between the two facets of management by exception behavior, namely active and passive, and as the passive facet is completely enclosed in the 'passive-avoidant' factor, this factor will simply be named 'management by exception'.

The examination of the full model with nine correlated factors in this sample again showed unsatisfactory fit-indices ($\chi^2 = 3569$, $df = 558$, $p < 0.01$, $N_2 = 879$, $AGFI = 0.78$, $CFI = 0.84$, $RMR = 0.078$).

Study 3

The third step of this investigation (Study 3, $N_3 = 650$) serves the goal of affirming the obtained results by confirmatory factor analysis. As can be seen in Table 34, the structure found in Study 2 is clearly replicated using an independent sample. The model meets the criteria for a good fit set in the literature ($\chi^2 = 217$, $df = 62$, $p < 0.01$, $N_3 = 650$, $AGFI = 0.93$, $CFI = 0.95$, $RMR = 0.042$). Hence, expectations 2a and b can be maintained.

The internal consistency of the three factors is in parts satisfactory. Cronbach's alpha amounts to $\alpha = .86$ for 'charismatic goal orientation' and $\alpha = .85$ for 'passive-avoidant leadership'. Only 'management by exception' ($\alpha = .60$) does not meet the Nunnally (1978) criterion. MbEa did not show a satisfactory internal consistency in Study 1. As the scale did not change considerably but even lost one item, this result is not surprising.

The pattern of correlations is now more clear-cut. The 'transformational' facet shows a relatively high negative correlation ($r = -.53$, $p < .01$) with 'passive-avoidant leadership' and a smaller positive correlation with 'management by exception' ($r = .20$, $p < .01$). The two former transactional and non-leadership facets correlate moderately negative with each other ($r = -.12$, $p < .01$).

A reanalysis of the three-factor model with the first sample yielded good fit indices as well ($\chi^2 = 263$, $df = 62$, $p < 0.01$, $N_1 = 1311$, $AGFI = 0.91$, $CFI = 0.93$, $RMR = 0.061$); the reliability coefficients were $\alpha = .84$ for 'charismatic goal orientation', $\alpha = .79$ for 'passive-avoidant leadership', and $\alpha = .60$ for 'management by exception'.

Short digression No. 3: Using positively formulated Laissez-Faire items

Unlike in the original version, Felfe and Goihl (2002b) re-formulated the LF items in a positive way so that 'the person I am rating is absent when needed' turns into 'the person I am rating is always there when needed'. This was done as a self-rating version was developed at the same time, and the authors were of the opinion that the passive formulation of the items would not be rated in an adequate way by the leaders. However, the following question arises: Do the more positively formulated items tap the same area of behavior as the originally formulated items or do the items now measure something more active or even transformational? A further question then would be: Is LF only the passive formulation of transformational leadership and therefore not a unique behavioral facet? Taking earlier results on item wording effects into account, it is questionable if the regularly formatted and the polar opposite formatted items load on a common factor. For example, Pilotte and Gable (1990) and Miller and Cleary (1993) analyzed the effects of item wording and "found that positively- and negatively-worded items formed separate and distinct factors" (Schriesheim & Eisenbach, 1995, p. 1179). Schriesheim and Eisenbach (1995) state that their "EFA results show that the polar opposite … item formats are clearly capable of obtaining problematic loadings in an exploratory factor analysis" (p. 1188). In their study, the differently worded LBDQ items all loaded on separate method factors in a CFA. Although the correlational pattern of the recoded LF scale resembles that of the originally formulated items, the mean LF value seems to be higher for the recoded version than for the original items (Felfe & Goihl, 2002b).

However, one can assume that:

3. The positively formulated and the polar opposite LF items form unique factors in an exploratory factor analysis.

Method

Sample

This part of the study is based on another sample from a public administration office. This study, too, was integrated into an overall organizational analysis, conducted in order to accompany a reform process in the organization. The sample consists of N_4 =

404 participants. The mean age is 38.8 years (*SD* = 9.7 years), 68.2% of the sample are female. 15% of the participants had leadership functions. The rejection rate was 34%.

Measures

The MLQ (Version 5X Short, Bass & Avolio, 1995a; translated and adapted by Felfe & Goihl, 2002b) was used in this study as well. Moreover, the LF items were also presented in the original format. As the MLQ comprises, besides the leadership items, nine additional items in order to measure success criteria and Felfe and Goihl (2002b) also added some charisma items to the leadership items, there were a lot of possibilities to mix in the originally formulated LF items. None of the respondents complained about a repetition of items. Furthermore, the number of missings of these eight items was comparable (ranging from 10 to 16; only item 5, describing a person who avoids getting involved when important issues arise, had 25 missings). In the following section, the originally formulated items are labeled with their respective item number, the recoded (polar opposite formatted) items are labeled with the item number of the originally formatted correspondent item and the annex 'rec'.

Results

The originally formulated and recoded LF scales show significantly different means (see Table 35; $p < .01$), whereby the originally formulated scale has a lower mean than the more positively formulated and recoded items. Furthermore, the originally formulated scale has a higher reliability, although both scales show a satisfactory α. The lower α value of the recoded scales stems from the weak correlations of item MLQ 5 rec with the other items of the scale (the correlations range from $r = .22$ to $r = .23$). The two scales correlate with each other with $r = .72$.

Table 35. Means and standard deviations of the MLQ laissez-faire scales ($N_4 = 404$)

Scale	**M**	**SD**	**α**
laissez-faire (originally formulated)	2.29	.90	.83
laissez-faire (recoded)	2.66	.88	.76

An exploratory factor analysis (EFA) showed only one factor for the originally formulated items and the recoded items of the Felfe and Goihl (2002b) version (see

Table 36). 51.3% of the variance was explained by this factor. All items, except the recoded item 5 show high loadings on the factor.

Table 36. Factor loadings of the MLQ laissez-faire items (N_4 = 404)

	Factor I
MLQ_5	.80
MLQ_7	.71
MLQ_28	.79
MLQ_33	.78
MLQ_5 rec	.33
MLQ_7 rec	.82
MLQ_28 rec	.79
MLQ_33 rec	.87

In order to take a look at the connection of the LF items with the transformational items, an EFA with the LF and the transformational items was performed. The result of the PAA with oblique rotation (eigenvalue > 1) shows that most of the LF (originally formulated and recoded) items load on one factor. However, four transformational items, two IIA and two IS items also load on this factor. It should be noted, though, that these items also have high loadings on Factor 3 and that their loadings on Factor 1 are, except for item 25, noticeably lower than those of the LF items (see Table 37).

On the other hand, the last recoded LF item, MLQ 5 rec (that also shows the weakest factor loading on the LF factor in the preceding analysis) loads on a factor with three IC items, one IS item, and two IIA and IIB items, but has a comparable loading on the first factor as well.

Table 37. Factor loadings of the MLQ transformational and laissez-faire items (N_4 = 404)

	Factor I	**Factor II**	**Factor III**
MLQ_28	***-.88***		.19
MLQ_33	***-.76***		
MLQ_33 rec	***-.76***		
MLQ_28 rec	***-.71***		
MLQ_5	***-.70***		-.11

Table 37 continued

	Factor I	Factor II	Factor III
MLQ_7 rec	***-.64***		-.21
MLQ_7	***-.61***	.12	-.13
MLQ_25	***.60***	.14	.21
MLQ_21	***.41***	.11	.40
MLQ_8	***.37***	.19	.23
MLQ_2	***.32***	.31	.26
MLQ_13		***.95***	-.10
MLQ_14		***.73***	.11
MLQ_26	.18	***.66***	
MLQ_9		***.65***	
MLQ_15	.21	***.39***	.29
MLQ_34		***.37***	.27
MLQ_32	.34	***.34***	.23
MLQ_29			***.88***
MLQ_19			***.72***
MLQ_31		.21	***.66***
MLQ_30		.21	***.62***
MLQ_18	.29		***.45***
MLQ_23	.21	.17	***.38***
MLQ_10	.27	.25	***.33***
MLQ_6		.23	***.28***
MLQ_5 rec	-.14		***-.17***

Taking all the leadership items of the MLQ into the analysis, most of the LF items form one factor together with the (also negative) MbEp items. However, as in the analyses above, item MLQ 5 rec does not load on this factor, but on a factor that is mainly built up by transformational items.

A finer analysis of the item allocation can be provided by CFA (Schriesheim & Eisenbach, 1995). Three models were tested according to the procedure suggested by Schriesheim and Eisenbach (1995), one model with two method factors and one trait factor, one model with two method factors (the respective formulation of the items) and one model with only one trait factor. Only the method factors were allowed to be correlated with each other. The results show that, although all models yield good fit indices, the model with the two separate method factors and the trait factor has the best

fit, and shows significant improvement in the χ^2 values over the other two tested models (see Table 38).

Table 38. CFA for the laissez-faire items (N_4 = 404)

Model	χ^2	df	GFI	AGFI	CFI	NNFI	RMSEA	$\Delta\chi^2$
(1)	50.04	12	.97	.92	.98	.95	.09	
(2)	91.50	19	.95	.90	.96	.93	.10	41.46 ($p < .01$)
(3)	129.50	20	.93	.87	.93	.91	.12	79.46 ($p < .01$)

Note: Model 1: two correlated method factors and one trait factor, model 2: two correlated method factors, model 3: one trait factor; GFI = goodness of fit index; AGFI = adjusted GFI; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation.

Taking all the results into account, the originally and recoded item formulations do not seem to tap different leadership behaviors. The EFAs show that both LF versions load on one factor, with the exception of one item. This item, however, seems to be particularly difficult, as it also has the highest missing rate. The CFA support the assumption of a common trait factor as well. Therefore, assumption 3 can be maintained.

3.1.3. Discussion

Lately, more emphasis has been put on a more differentiated view of leadership. With the research into transformational leadership, the spectrum of leadership behaviors has become much wider. In addition to initiating structure and consideration - the standard 'range' of leadership behaviors for decades - charisma, motivation, and intellectual stimulation have become more focused in the last decade. However, the finer differentiation and wider range of leadership behavior facets in the MLQ has not found broad empirical support in the literature. In this study, once again, the nine-factorial model of the MLQ could not be supported. Although this model shows superior fit-indices in comparison to others (empirically and theoretically supported factor solutions with fewer factors) it does not reach the criteria for a good fit. Furthermore, high correlations are obtained for the transformational scales and their correlation with CR. A recently suggested procedure of examining the factorial models in homogenous contexts also failed to show any improvement in fit for the nine-factor model. The

model fitted worst for the group of people who worked in the administration less than five years. This is not surprising as these people might not have gathered enough leadership experience and therefore have not yet developed cognitive schemes that simplify judging their leader's behavior. Furthermore, models with second-order factors, as proposed by Bass (1985) and anchored in the theory of transformational leadership, also did not show satisfactory fit-values.

Hence, these results were the starting point for a more thorough examination of the underlying factor structure. It seemed that the item used to scale allocation had to be changed, as some items also showed higher loadings on and correlations with other scales than their own. The therefore conducted parallel analysis points out that a three-factor solution is sufficient to describe the leadership behavior facets that are assessed using the MLQ 5X. The obtained reduced set of factors, consequently, contains less items and does not comprise all the facets of transformational leadership that are proposed by Bass and Avolio (1995a). IS, IC, and IIA are not contained in the 13 item version. Within the new structure, the expected pattern emerges: the transformational facets associate with CR and are broken down into one factor, i.e. IM and IIB, the behavioral facets of the former charisma scale (Avolio et al., 1999) go together with the goal oriented items of CR, a transactional scale. The other transactional scales and LF are still part of the reduced version: MbEa stays the same, and MbEp and LF go together. LF, however, is only represented by one item. As this scale is always highly correlated with MbEp and as MbEp and LF show comparable negative correlations with positive external criteria, it is doubtful if additional information is gained when assessing the full LF scale. The proposed names of the three factors are: *charismatic goal orientation, management by exception,* and *passive-avoidant leadership.* Confirmatory factor analysis shows a better fit for this solution than for the original nine-factor structure.

Whatever model is chosen for further analysis, the three-factor model for empirical reasons or the nine-factor model, as often argued, for training purposes, the formulation of the LF items (originally or polar opposite formatted) does not seem to matter when taking the results of the EFA. However, taking a look at the means of the scales, there seems to be a significant difference between them. This is further supported by the results of the CFA that suggest that there is a common trait factor but also two method factors. The model with only the trait factor, which would indicate that all the items can

be subsumed in one trait, does not meet the fit criteria. Furthermore, one recoded item seems to be difficult to answer, and shows differences to the other scales in more than one way (correlation pattern or EFA results, when the transformational scales are also taken into account). I therefore would suggest that the originally formulated items should be taken in order to stay as close to the American version as possible, as long as this questionnaire is only transferred and not completely adapted to the German culture.

3.2 Part 2: Relation to questionnaires that measure the same construct

Two leadership instruments will be essential to this part of the investigation of the construct validity of the MLQ. First, a questionnaire that is supposed to also assess transformational leadership, the Transformational Leadership Behavior Inventory by Podsakoff et al. (Podsakoff et al., 1990). And second, a questionnaire that was mostly used to assess leadership behavior before the MLQ was developed, the Leadership Behavior Description Questionnaire (Stogdill, 1963).

3.2.1. The Transformational Leadership Behavior Inventory (TLI)

The TLI is, in comparison to the MLQ, a quite unknown questionnaire. Therefore, the structure of the TLI as well as the results obtained so far will be briefly discussed in order to provide a better basis for the understanding of the results in this study and the subsequent discussion.

As already mentioned in chapter 2.2.3, Podsakoff et al. (1990) agreed on six key behaviors of transformational leadership after a thorough review of the literature:

- *Identifying and Articulating a Vision (IAV)* – Behavior on the part of the leader aimed at identifying new opportunities for his or her unit/division/company, and developing, articulating, and inspiring others with his or her vision of the future.

- *Providing an Appropriate Model (PAM)* – Behavior on the part of the leader that sets an example for employees to follow that is consistent with the values the leader espouses.
- *Fostering the Acceptance of Group Goals (FAGG)* – Behavior on the part of the leader aimed at promoting cooperation among employees and getting them to work together toward a common goal.
- *High Performance Expectations (HPE)* – Behavior that demonstrates the leader's expectations for excellence, quality, and/or high performance on the part of followers.
- *Providing Individualized Support (PIS)* – Behavior on the part of the leader that indicates that he/she respects followers and is concerned about their personal feelings.
- *Intellectual Stimulation (IS_2[4])* – Behavior on the part of the leader that challenges followers to re-examine some of their assumptions about their work and rethink how it can be performed. (Podsakoff et al., 1990, p. 112)

Not all of the above mentioned scales are categorized as being core transformational leadership behaviors by all researchers. However, each of the above mentioned behavioral facets was identified as an important element of the transformational process. For some of the facets, though, the agreement is higher than for others. For example FAGG and PAM were identified as important transformational facets by four of the authors cited by Podsakoff et al. (1990, see page 114), IS is only mentioned by Bass (1985). As the authors wanted to guarantee that the domain of transformational leadership behaviors is adequately tapped, scales such as IS were also included in the questionnaire.

[4] The supplement "_2" is added in order to avoid confusion with the abbreviations of the MLQ scales.

Of the initial item pool of 100 items, generated by searching the literature for previous operationalizations and comparing them to the defined construct definitions, only 23 items remain after the Q-sort (criterion was an 80% agreement on the item's coding). Furthermore, transactional leadership is assessed with the TLI using one five-item scale [Contingent Reward (CR_2)].

Validation

The TLI is not as widely-used as the MLQ in assessing transformational leadership. Therefore, up to now, only three studies have examined the validity of the TLI.

Structure

Podsakoff et al. (1990) found acceptable fit indices for their model (transformational and transactional scales included) in the initial study. All items had significant and substantial loadings on their hypothesized factors. However, the correlations amongst three of the transformational scales were very high (approaching or exceeding $r = .90$). Therefore, the authors examined a second-order latent construct with three first-order factors as indicators (IAV, PAM, FAGG). All items had significant loadings on their first-order factors; IAV, PAM, and FAGG had significant loadings on the second-order factor. The second-order construct was named *core* transformational leader behavior and includes leadership activities identified as being *core* transformational by other researchers as well (e.g. Kouzes & Posner, 1987), i.e. "developing a vision and getting employees to accept it, being a good example of the values and behaviors that are essential to fulfilling the vision, and getting employees to put the interests of the group or organization above their self-interest" (Podsakoff et al., 1990, p. 134). PIS and IS_2 are consistent with the MLQ scales IC and IS, HPE is comparable to House's (1977) notion that having high performance expectations is a characteristic of a transformational leader. All in all, the model had a good fit [χ^2 (337) = 877.07, $p < .01$; *NNFI*: .97]. The coefficient alpha internal consistency reliability for the five scales (four transformational scales and CR_2) ranged from $\alpha = .78$ to $\alpha = .92$. The correlations between the transformational scales ranged from $r = .23$ to $r = .84$, the correlations between the transformational scales and CR_2 from $r = .37$ (HPE) to $r = .73$ (*core* transformational leader behaviors).

In the second study, Podsakoff et al. (1996) aimed at confirming the scale's psychometric properties. A six-factor model yielded good fit indices [χ^2= 965.51, *df* = 194, $p < .01$; *GFI* = .91, *NNFI* = .93; the CR_2 scale was not included in this study]. All items had statistically significant loadings (all .60 or above) on the hypothesized factors. A comparison with alternative models, where the hypothesized factors were combined (two to six at a time), suggested that the hypothesized six-factor model had the best fit.

Table 39. Means, *SD,* α, and correlations for the TLI scales (Podsakoff et al. 1996, p. 270)

	M [+]	***SD***	**IAV**	**PAM**	**FAGG**	**HPE**	**PIS**	**IS_2**
IAV	5.08	1.23	.87					
PAM	5.01	1.48	.68	.84				
FAGG	5.25	1.37	.63	.63	.89			
HPE	5.39	1.22	.49	.40	.39	.80		
PIS	4.88	1.55	.40	.54	.49	.12	.90	
IS_2	4.96	1.24	.61	.52	.46	.40	.31	.82

Note: IAV: identifying and articulating a vision, PAM: providing an appropriate model, FAGG: fostering the acceptance of group goals, HPE: high performance expectations, PIS: providing individual support, IS_2: intellectual stimulation, CR_2: contingent reward; r > .11: p < .01, coefficient α *is printed in diagonal*

[+] *scales range from (1) "strongly disagree" to (7) "strongly agree"*

The coefficient alpha internal consistency reliability for the six transformational scales ranged from α = .82 to α = .90, the intercorrelations ranged from *r* = .12 to *r* = .68 (see Table 39). The comparison between these two studies, however, is limited when dealing with the IS_2 scale as two different sets of items were used to assess this leadership facet. Unfortunately, Podsakoff et al. (1996) do not explain why this was done.

The third investigation, conducted by Schriesheim, Castro, Williams, Medsker, & DeChurch (1997, cited in Castro, 1998), was divided into three studies. In all three studies, the six-factor structure was generally supported and most items loaded strongly on their hypothesized factors. It should be noted, though, that on the IAV, HPE and PIS scales some factor loadings were < .70 in at least two of the three samples. Therefore, Castro (1998) suggests that those items should be reviewed or even replaced. The coefficient alpha internal consistency reliability for the six transformational scales in the three studies ranged from α = .74 to α = .95.

External Criteria

In the two aforementioned studies by Podsakoff and his colleagues (Podsakoff et al., 1996; Podsakoff et al., 1990), the correlations with several external criteria were examined as well.

In the first study, the correlations of the transformational leadership scales with trust and satisfaction ranged from $r = .41$ to $r = .87$, the correlations with CR_2 were $r = .72$ resp. $r = .64$ (Podsakoff et al., 1990, see Table 40). A structural model revealed a significant impact of the *core* transformational behaviors, PIS, and IS_2 on trust and satisfaction. However, IS_2 had a negative impact on both criteria. Podsakoff et al. (1990) noted that one possible explanation of these effects might be due to the fact that

> although intellectual stimulation may produce desirable effects in the long run, it may be that in the short run, leaders who continually urge or exhort followers to search for new and better methods of doing things create ambiguity, conflict, or other forms of stress in the minds of those followers. (p. 135)

Unfortunately, no information about team duration or task duration was given in this context. CR_2 showed no significant effect on trust or satisfaction.

Furthermore, the impact of transformational leadership on citizenship behavior, being the extra-role performance this leadership style is said to evoke (usually denoted as performance beyond expectations), was analyzed. The results indicated that transformational leadership behaviors do affect organizational citizenship behaviors. However, those effects were indirect, mediated by the followers' trust in the leader.

The second study examined a variety of external criteria (e.g. satisfaction, trust, loyalty, commitment, performance and organizational citizenship behavior). Furthermore, leadership substitutes were analyzed. The correlations of the transformational leadership scales and the external criteria were mostly positive (except the correlations with role conflict) and ranged from $r = .06$ to $r = .63$ (Podsakoff et al., 1996, for examples see Table 40). Regression analyses with the transformational leadership behaviors and substitutes for leadership showed 71% explained variance for satisfaction and 48% for organizational commitment. However, only 7% resp. 1% explained variance is uniquely attributable to the transformational leadership behaviors.

The regression coefficients show that five leader behaviors had significant main effects on the followers' general satisfaction. Four facets had positive effects (PIS: β = .12, PAM: β = .10, IAV: β = .09, FAGG: β = .05), HPE (β = -.05) had a small but negative effect on satisfaction. Only one of the leader behaviors had a significant impact on organizational commitment (IAV: β = .10). For satisfaction as well as for commitment six substitutes for leadership had significant effects. Referring to the other criterion variables (trust in leader, role clarity, role conflict, in-role performance, altruism, conscientiousness, sportsmanship, courtesy, and civic virtue) of the leadership behaviors, only IAV had significant effects on all of these variables. Therefore, IAV "appears to be a particularly important determinant of employee attitudes, role perceptions and behaviors" (Podsakoff et al., 1996, p. 290).

Table 40. Correlations between TLI scales and external criteria (Podsakoff et al., 1990, p. 126; Podsakoff et al. 1996, p. 270)

	Podsakoff et al., 1990	*Posdakoff et al., 1996*	
	Satisfaction	**Satisfaction**	**Commitment**
IAV		.49	.34
PAM		.48	.28
FAGG		.46	.27
CORE	.77		
HPE	.48	.27	.20
PIS	.65	.45	.25
IS_2	.61	.36	.26
CR_2	.64		

Note: IAV: identifying and articulating a vision, PAM: providing an appropriate model, FAGG: fostering the acceptance of group goals, CORE: core transformational leadership, HPE: high performance expectations, PIS: providing individual support, IS_2: intellectual stimulation, CR_2: contingent reward; $r > .19$; $p < .01$

Although the TLI is not based on such a comparably broad data basis as the MLQ is, it seems justified to use it as a measure for approaching the convergent validity of the MLQ. First of all, the inclusion of formerly identified important facets of transformational leadership seems to provide a broad coverage of transformational behavior. Secondly, the factorial structure of the instrument appears to be stable. And thirdly, the MLQ does not only assess transformational leadership, but also facets of

transactional leadership. With its own CR_2 scale, the TLI therefore provides the possibility to examine this facet as well. Most of the other instruments that measure transformational leadership as, for example, the Transformational Leadership Questionnaire (Alban-Metcalfe & Alimo-Metcalfe, 2000; Alimo-Metcalfe & Alban-Metcalfe, 2001) do not measure any other leadership facets, for example transactional leadership. Furthermore, taking the correlations with the external criteria, the TLI shows the expected correlations, for example high correlations with satisfaction, trust and commitment.

However, choosing the TLI also holds some problems. For example, the criticism expressed because of the high correlations of the MLQ transformational scales and with CR are not solved within the TLI. Here also high correlations among these scales appear.

As the MLQ and the TLI are both declared to measure the same leadership construct, the following assumptions can be made:

5. The transformational scales show high correlations with each other.

6. The transformational scales of the MLQ show higher correlations with the transformational scales of the TLI than the transactional scales of the MLQ and LF.

On the basis of the scale contents, it can be seen that some of the MLQ and TLI scales tap the same facet of transformational behavior. For example, Intellectual Stimulation is part of both questionnaires. Furthermore, IC and PIS both measure leader behavior that indicates the respect of leader for his followers and his concern about their personal feelings.

Therefore:

7. IS and IS_2 as well as IC and PIS show higher correlations with each other than with the rest of the transformational scales of the respective other instrument.

8. CR and CR_2 also shows a higher positive correlation than CR has with any other scale.

3.2.2. Method

Sample

Basis of this part of the investigation was the sample Study 4 that was already described in chapter 3.1.2. All in all, 404 respondents were taken into account.

Measures

As in the studies before, the MLQ translation of Felfe and Goihl (2002b) was used. The TLI comprises, as the two reported studies show, different scales for the two different versions. For the translation into German (see Heinitz & Rowold, in print), the more recent version was chosen (Podsakoff et al., 1996). As transactional leadership is supposed to be the basis of transformational leadership and, in order to cover a broad range of leadership behaviors, the CR_2 scale from the older version (Podsakoff et al., 1990) was also included in the translation. The questionnaire was translated by a bilingual person, compared with the original and inconsistencies were discussed. The version of the TLI used here therefore comprises 26 items and measures the scales *Identifying and Articulating a Vision (IAV* – 5 items*)*, *Providing an Appropriate Model (PAM* – 3 items*)*, *Fostering the Acceptance of Group Goals (FAGG* – 4 items*), High Performance Expectations (HPE* – 3 items*)*, *Providing Individualized Support (PIS* – 4 items*)*, *Intellectual Stimulation (IS_2* – 3 items*)* and *Contingent Reward (CR_2* – 4 items). In order to provide a comparable response format with the other presented scales, the response format was changed from a 7-point Likert-scale into a 5-point Likert-scale ranging from *never* (1) to *always* (5).

An examination of the factorial structure of the TLI shows the factor loadings presented in Table 41. The specifications of the CFA are the same as the ones reported for the factorial validation of the MLQ.

Table 41. Factor loadings and factor intercorrelations for the TLI scales, completely standardized solution (N_4 = 404)

Scales	IAV	PAM	FAGG	HPE	PIS	IS_2	CR_2
Item 2	.73						
Item 12	.79						
Item 4	.78						
Item 18	.88						
Item 21	.76						

Table 41 continued

Scales	IAV	PAM	FAGG	HPE	PIS	IS_2	CR_2
Item 22		.82					
Item 10		.73					
Item 25		.82					
Item 7			.80				
Item 9			.75				
Item 11			.84				
Item 17			.85				
Item 5				.75			
Item 20				.62			
Item 24				.40			
Item 8 rec					.65		
Item 13					.83		
Item 16					.89		
Item 23 rec					.70		
Item 1						.75	
Item 3						.82	
Item 14						.80	
Item 6							.90
Item 15 rec							.53
Item 19							.89
Item 26							.86
IAV							
PAM	.91						
FAGG	.93	.91					
HPE	.60	.42	.54				
PIS	.61	.69	.71	.14			
IS_2	.83	.79	.73	.48	.55		
CR_2	.67	.65	.71	.31	.64	.60	

Note: IAV = identifying and articulating a vision; PAM = providing an appropriate model; FAGG = fostering the acceptance of group goals; HPE = high performance expectations; PIS = providing individualized support; IS_2 = intellectual stimulation; CR_2 = contingent reward; $r > .13$, $p < .05$; $r > .30$, $p < .01$.

The factors show high correlations, only PIS and HPE correlate moderately with $r = .14$. The confirmatory factor analyses also show that the proposed seven-factor model has a satisfactory fit. None of the other tested models, a model with one general factor and a

model with one higher-order factor according to the core-transformational factor of Posdakoff et al. (1990), shows a better fit than the seven-factor model (see Table 42).

Table 42. CFA of the TLI (N_4 = 404)

Model	χ^2	*df*	*GFI*	*AGFI*	*NNFI*	*CFI*	*RMSEA*	$\Delta\chi^2$	Δdf
(1)	1946.96	299	.662	.603	.739	.760	.117	1281.79 (p < .01)	11
(2)	665.17	278	.880	.849	.934	.944	.059		
(3)	718.94	286	.868	.838	.937	.928	.061	53.63 (p < .01)	8

Note: GFI = goodness of fit index; AGFI = adjusted GFI; NNFI = non-normed fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation. (1) one general factor; (2)seven-factor model; (3) model with one higher-order factor (TLI-CORE), that comprises the scales IAV, PAM and FAGG. $\Delta\chi^2$ was computed by subtracting the χ^2 value of the seven-factor model of the χ^2-value of the respective model.

All in all, the scales have a satisfactory internal consistency. Only HPE does not meet the criterion of α = .70 (α = .61, see Table 43). As with the MLQ, the scales of the TLI are represented by unweighted means of their items.

The scale intercorrelations for the transformational scales range from r = -.02. to r = .83. The correlations of CR with the transformational scales range from r = .15 to r = .64 (see Table 43).

Table 43. Means, internal consistencies and scale intercorrelations for the TLI scales (N_4 = 404)

Scale	No. of items	*M*	*SD*	α	IAV	PAM	FAG	HPE	PIS	IS_2
AV	5	2.66	.90	.89						
PAM	3	2.94	.96	.83	.78					
FAGG	4	3.00	.95	.88	.83	.78				
HPE	3	3.18	.79	.61	.42	.25	.37			
PIS	4	3.48	.92	.86	.50	.57	.58	-.02		
IS_2	3	2.48	.86	.83	.72	.66	.63	.31	.43	
CR_2	4	2.92	1.00	.87	.59	.56	.64	.15	.58	.51

Note: IAV = identifying and articulating a vision; PAM = providing an appropriate model; FAGG = fostering the acceptance of group goals; HPE = high performance expectations; PIS = providing individualized support; IS_2 = intellectual stimulation; CR_2 = contingent reward.
r > .15: p < .05, r> .28: p < .01

3.2.3. Results

The transformational scales of the MLQ and the TLI show high correlations with each other. The highest correlations appear for the core-transformational facets IAV, PAM, and FAGG. HPE shows lower correlations with the transformational scales of the MLQ, PIS and IS_2 also show high but lower correlations with the MLQ scales than the core-transformational scales (see Table 44). All correlations are significant, assumption 5 is therefore maintained.

Table 44. Scale intercorrelations of the MLQ and TLI scales (N_4 = 404)

TLI	IAV	PAM	FAGG	HPE	PIS	IS_2	CR_2
MLQ 5x							
IIa	.74	.78	.71	.21	.61	.63	.56
IIb	.69	.61	.69	.36	.44	.55	.49
IM	.72	.57	.62	.38	.34	.56	.45
IS	.75	.70	.69	.28	.51	.65	.55
IC	.69	.65	.68	.20	.63	.61	.65
CR	.71	.63	.69	.33	.53	.59	.63
MbEa	.08	.05	.08	.30	-.12	.14	-.05
MbEp	-.61	-.56	-.58	-.16	-.52	-.46	-.50
LF	-.62	-.65	-.61	-.16	-.55	-.51	-.48

Note: IAV = identifying and articulating a vision; PAM = providing an appropriate model; FAGG = fostering the acceptance of group goals; HPE = high performance expectations; PIS = providing individualized support; IS_2 = intellectual stimulation; CR_2 = contingent reward; IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; r > |.08|: p < .05, r> |.12|: p < .01.

Comparable to the correlational pattern within the MLQ, CR shows high correlations with the core transformational scales of the TLI, MbEa shows the lowest correlations and MbEp shows negative correlations with all the scales. For the core transformational scales as well as PIS and IS_2, the correlations of CR show no significant differences to those of the transformational scales of the MLQ. However, CR and MbEa correlate higher with the transformational scale HPE than the transformational scales IIA and IC. The difference between these correlations is significant for CR ($p < .01$), but not for the

MbEa scale. CR also has higher correlations with the PIS and the IS_2 scales than some of the transformational scales of the MLQ (e.g. than IIB, IM and IS in case of PIS). Significant differences in the correlations with PIS result for IM and IIB, for IS_2 the differences are not significant. MbEa correlates significantly lower with these two scales than any of the transformational scales of the MLQ. Furthermore, MbEp and LF show negative correlations with all TLI scales. Assumption 6 is not supported.

IS and IS_2 correlate with $r = .65$, IC and PIS with $r = .63$. Although these correlations are already high, both scales show higher correlations with other transformational scales. Both facets of the MLQ transformational leadership correlate higher with the core-transformational scales of the TLI. The differences of the correlations are significant for IS/IS_2 with IAV ($p < .01$) and PAM ($p < .05$) and for IC/PIS with IAV ($p < .05$). CR correlates with CR_2 with $r = .63$, higher values are obtained for the correlations of CR with IAV and PAM, both differences are significant (for IAV: $p < .01$, for PAM: $p < .05$). Assumptions 7 and 8 are therefore not supported.

In order to further examine the convergence of the MLQ and TLI transformational and transactional concepts, a principal axis analysis was computed. The eigenvalue > 1 criterion leads to 10 factors, approximately 60% of the variance was explained (see Table 45).

Table 45. Factor loadings of the MLQ and TLI items ($N_4 = 404$), PAA, varimax-rotation

Item	Factor Scale	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**
TLI_10	PAM	**.64**	.21		.16	.13	.16	.11			
TLI_25	PAM	**.64**	.27	.12	.22	.20	.14	.22			
TLI_22	PAM	**.62**	.39	.25	.19	.12	.26	.16			
TLI_12	IAV	**.60**	.26	.19	.21	.12	.24	.11	.26		
TLI_18	IAV	**.59**	.29	.44	.13	.24	.13	.18	.15		
TLI_11	FAGG	**.58**	.29	.21	.14	.20	.22	.13	.26		.19
TLI_17	FAGG	**.57**	.31	.20	.20	.19	.31		.17		.22
TLI_7	FAGG	**.55**	.24	.22	.14	.34	.20		.12		.19
TLI_21	IAV	**.51**	.11	.41	.15	.25		.25	.16		
TLI_9	FAGG	**.50**	.12	.24	.13	.23	.19	.11	.28		.33
MLQ_10	IIA	**.48**	.31	.33	.31	.15	.18	.17			
TLI_4	IAV	**.46**	.27	.41	.16	.21		.22	.22		
MLQ_21	IIA	**.45**	.42	.24	.42	.14	.21	.14			
MLQ_18	IIA	**.42**	.27	.13	.39	.15	.23				

Table 45 continued

Item	Factor Scale	I	II	III	IV	V	VI	VII	VIII	IX	X
MLQ_12	MbEp	-.27	**-.68**	-.16	-.26	-.16	-.13				-.15
MLQ_20	MbEp		**-.66**	-.19	-.11	-.19	-.14	-.10	-.17		
MLQ_28	LF	-.28	**-.66**	-.18	-.11		-.20			-.16	
MLQ_5	LF	-.34	**-.63**	-.12	-.21	-.12	-.11				-.14
MLQ_3	MbEp	-.32	**-.62**	-.13	-.19	-.16	-.16	-.10			-.16
MLQ_17	MbEp		**-.60**	-.20		-.11	-.11		-.13	.14	
MLQ_33	LF	-.26	**-.59**	-.20	-.16		-.21	-.15			
MLQ_7	LF	-.22	**-.52**		-.23		-.13	-.15			-.13
MLQ_25	IIA	.48	**.51**	.28	.35		.20		-.12	.11	
MLQ_32	IS	.29	**.42**	.35	.26	.17		.34		.13	.25
TLI_2	IAV	.38	**.40**	.35	.10	.12		.25	.26		
MLQ_8	IS	.33	**.38**	.23	.36		.15				.11
MLQ_13	IM	.20	.17	**.74**	.14	.11			.17	.12	.14
MLQ_26	IM	.35	.32	**.65**	.21	.11		.16			
MLQ_14	IIB	.25	.24	**.60**	.31	.12	.12	.12	.19	.13	.15
MLQ_9	IM	.16	.20	**.55**	.20			.10	.11		
MLQ_16	CR	.37	.35	**.41**	.37	.22		.16	.12	.10	
MLQ_36	IM	.16	.30	**.32**	.30	.22	.16			.19	.31
MLQ_29	IC	.16	.22	.14	**.61**	.27	.30	.20		-.10	.16
MLQ_30	IS	.26	.28	.24	**.53**	.15		.40			.20
MLQ_19	IC	.19	.20	.13	**.51**	.24	.36				.14
MLQ_23	IIB	.35	.25	.18	**.50**		.17			.14	
MLQ_31	IC	.27	.30	.30	**.46**	.33	.21	.35			.17
MLQ_11	CR	.12	.18	.21	**.43**	.18	.15	.11	.30	.18	
MLQ_15	IC	.39	.32	.34	**.42**	.16		.11		.17	
MLQ_2	IS	.31	.41	.30	**.41**		.17	.16	.13	.14	
MLQ_1	CR	.33	.36	.20	**.40**	.23	.33	.11			
MLQ_35	CR		.25	.24	**.39**	.36	.15			.22	.36
MLQ_6	IIB	.12	.11	.18	**.35**	.12			.19	.10	
TLI_19	CR_2	.29	.14	.12	.20	**.76**	.19	.16			
TLI_6	CR_2	.25	.21	.11	.18	**.76**	.22				
TLI_26	CR_2	.27	.16	.18	.21	**.70**	.17	.15			.11
TLI_15rec	CR_2		.34		.11	**.39**	.30		-.18		
TLI_16	PIS	.35	.17		.30	.21	**.68**		.11		
TLI_13	PIS	.28	.18		.27	.23	**.66**	.10			
TLI_23rec	PIS	.15	.31		.11	.17	**.65**		-.24	-.11	
TLI_8rec	PIS	.19	.29			.13	**.56**		-.15		

Table 45 continued

Item	Factor Scale	I	II	III	IV	V	VI	VII	VIII	IX	X
TLI_1	IS_2	.32	.26	.24	.16	.15	.12	**.53**	.12		
TLI_14	IS_2	.39	.15	.24	.14	.20	.15	**.51**	.14		
TLI_3	IS_2	.42	.14	.26	.15	.17	.10	**.51**			
TLI_24	HPE						-.11		**.50**		
TLI_20	HPE	.12		.24					**.43**	.31	
TLI_5	HPE	.27	.13	.25		.14		.12	**.39**	.24	.13
MLQ_24	MbEa	.11	.24	.12	.16				.16	**.54**	
MLQ_22	MbEa		-.14							**.50**	
MLQ_4	MbEa	-.18	-.38		-.14	-.14	-.19			**.46**	
MLQ_27	MbEa		.18		.13			.30		**.46**	.19
MLQ_34	IIB	.33	.20	.31	.24				.13		**.56**

Note: the items carry the instrument, as well as the number of their appearance in the instrument; IAV = identifying and articulating a vision; PAM = providing an appropriate model; FAGG = fostering the acceptance of group goals; HPE = high performance expectations; PIS = providing individualized support; IS_2 = intellectual stimulation; CR_2 = contingent reward; IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; only loadings > .09 are displayed.

The TLI scales IAV, PAM and FAGG build the first factor together with three of the MLQ IIA items. The other TLI scales build up their own factors, whereas only the MLQ MbEa items load on an individual factor. The rest of the MLQ items mixes up and breaks down into three bigger factors. CR as a transactional facet is combined with the transformational items of the MLQ, the CR_2 items, however, build their own factor. Apart from IIA and one IAV item, the transformational facets of the TLI seem to tap different contents than the transformational facets of the MLQ.

A further examination of the overlapping of the transformational concepts of the MLQ and the TLI by means of CFA showed an unsatisfactory fit for a model with two correlated higher-order factors of transformational and transactional leadership where the respective MLQ and TLI scales were combined (χ^2 = 4287.77, *df* = 1579, *GFI* = .70, *AGFI* = .67, *CFI* = .83, *NNFI* = .82, *RMSEA* = .065). This was mainly due to the high overlapping of the MLQ transformational scales. A model where all the transformational items loaded on one factor and all the transactional items on another

correlated factor, however, showed an even worse fit to the data (χ^2 = 6044.03, *df* = 1595, *GFI* = .56, *AGFI* = .53, *CFI* = .73, *NNFI* = .72, *RMSEA* = .083).

Short digression No. 4: Taking the three-factorial model as a basis

As in the aforementioned study, the MLQ nine-factorial model showed an unsatisfactory fit for these data as well (χ^2 = 1723.14, *df* = 558, *GFI* = .80, *AGFI* = .76, *CFI* = .87, *NNFI* = .85, *RMSEA* = .072). The model proposed in chapter 3.1.2 with three correlated factors and fewer items again shows satisfactory fit indices (χ^2 = 196.20, *df* = 62, *GFI* = .93, *AGFI* = .90, *CFI* = .93, *NNFI* = .91, *RMSEA* = .073). Therefore the correlations of this model with the TLI will be examined here as well. The three-factorial model combines the transformational items in one factor (charismatic goal orientation), and has one pure transactional factor (management by exception). MbEa, which is the only content of management by exception, shows only weak correlations with the MLQ scale CR, another transactional facet or the transformational scales of the MLQ. Therefore, the following assumptions can be made:

9. Charismatic goal orientation shows high correlations with the transformational scales of the TLI and CR_2.

10. Management by exception shows weak to zero correlations, passive-avoidant leadership will show negative correlations with the TLI scales.

Results

The correlations show the expected pattern for Charismatic goal orientation (see Table 46). All correlations are significant. As seen for the nine-factor model of the MLQ, the core-transformational scales of the TLI show the highest correlations with Charismatic goal orientation. HPE and PIS correlate significantly lower with this scale, CR_2, however, shows high correlations with this transformational facet. Management by exception also shows significant correlations with the TLI scales.

Table 46. Scale intercorrelations of the three-factor model of the MLQ and the TLI scales (N_4 = 404)

TLI	IAV	PAM	FAGG	HPE	PIS	IS_2	CR_2
MLQ							
Charismatic goal orientation	.78	.63	.72	.39	.43	.61	.54
Management by exception	.25	.23	.26	.32	.07	.30	.11
Passive-avoidant leadership	-.62	-.58	-.59	-.16	-.52	-.48	-.51

Note: IAV = identifying and articulating a vision; PAM = providing an appropriate model; FAGG = fostering the acceptance of group goals; HPE = high performance expectations; PIS = providing individualized support; IS_2 = intellectual stimulation; CR_2 = contingent reward; r > |.08|: p < .05, r> |.12|: p < .01.

Only the correlation with PIS does not reach the level of significance. As expected, the passive-avoidant leadership behaviors show negative correlations with the active transformational leadership behaviors. Only HPE shows a significantly lower negative correlation with passive-avoidant leadership than the other TLI scales. Hence, assumption 9 is maintained, assumption 10 is, due to the significant correlation of MbE with the TLI scales, retained.

3.2.4. MLQ and LBDQ

A second very popular approach in leadership research goes back to the OHIO studies (see chapter 2.2.4) and differentiates between Consideration (C) and Initiating Structure (InS). Although the C/InS concept and the transformational leadership concept can probably be considered the two most popular leadership theories, only two studies up to date examined differences and parallels between these leadership styles. One reason for this may be that, when dealing with C and InS, “in most comprehensive reviews of the leadership literature, one is left with the impression that these constructs are archaic. The Ohio State studies, contemporary thinking argues, are to be remembered only to better understand where we are today” (Judge et al., 2004, p. 37).

Concerning their similarities, Bass (1999b) states that the “transactional/ transformational paradigm is independent conceptually from the concepts of directive

vs. participative leadership, leader-member exchange (LMX), and the factor of consideration as measured by the Leader Behavior Description Questionnaire, although empirical correlations with them may be found to some extent" (p. 13). However, one strong similarity is that Consideration and Individualized Consideration share nearly the same name. According to Bass (1999b) "LBDQ consideration focuses on friendliness, approachability, and participative decision making; Individualized consideration deals with concern for each follower as an individual and with the follower's development." (p. 14). Seltzer and Bass (1990) discuss this in more detail:

> Individualized consideration was unlike the consideration construct in that it focused on the individual development of followers. Despite the use of the same term, "consideration" in both the prior conceptualization and in Bass' framework, it was not expected that the two scales would necessarily be correlated. Consideration is measured by a wide array of relations-oriented behaviors of the leader, which arguably provides in exchange, acceptance of the leader and satisfaction with him or her. Individualized consideration measures behavior that is transforming through its attention to the individual members and their development. We can note that the intercorrelation of the consideration and individualized considerations scales in this study was .60 ($p < .01$). A principal component factor analysis of the items from the two scales was performed with varimax rotation and a two-factor solution. All 10 individualized consideration items loaded uniquely on the first factor. Nine of the consideration items loaded uniquely on the second factor. The remaining consideration item loaded on both factors. We conclude that empirically as well as conceptually consideration and individualized consideration are distinct. (p. 695)

Furthermore, Bass (1995) states that "LBDQ initiation and consideration conceptually may substitute for transactional, but not for transformational leadership, for much additional variance in effectiveness is accounted for by transformational leadership when multiple regression analyses are applied" (p. 473). However, high correlations of

transformational leadership and C have been found. Keller (1992) reports correlations that vary between $r = .51$ and $r = .76$ for the scales charismatic leadership, intellectual stimulation and consideration. In the study of Seltzer and Bass (1990) the correlations between charisma, IS and C were $r = .69$ and $r = .47$ resp., IC and C correlated with $r = .60$.

No study up to date considered the link between the transformational/transactional paradigm and InS in detail. This might be due to the fact that the structuring of tasks and the specification of methods to get jobs done is contrary to the transformational understanding of leadership, where old ways are questioned and new ideas of how to get things done are encouraged. However, structuring tasks is considered to be one important facet of leadership. For example, Yukl (2002) defines fourteen managerial practices that also contain the categories *clarifying roles and objectives* as well as *planning and organizing*. Although one could argue that management and leadership are different, the category *inspiring and motivating*, hence transformational leadership behavior, is also part of the taxonomy. Keller (1992) stated that "Initiating structure fits Bass' (1985) notion of transactional leadership because this leader behavior focuses on the clarification and assignment of role requirements" (p. 491). Yet, CR focuses more on the clarification of rewards if certain standards are met, MbEa and MbEp monitor, but don't clarify structures, standards or ways of problem solving. However, as CR and InS both consider standards, they should correlate positively with each other. Unfortunately, none of the reported two studies reports correlations of transactional leadership and InS. The correlations of InS with the transformational scales are lower than the correlations reported with C. In the Keller (1992) study the correlations vary between $r = .12$ and $r = .21$, in the study of Seltzer and Bass (1990), InS correlated with the transformational facets from $r = .39$ to $r = .48$.

In the discussion of the superiority of the transformational concept over the 'old-fashioned' ideas of leadership, the question arose if transformational and transactional leaders showed any difference in InS or C behaviors. Bass (1995) stated that "the widely used LBDQ measures of initiation and consideration did not differentiate between transformational and transactional leaders" (p. 469). Hence it seems that these behaviors measure something different to transformational and transactional leadership. Transformational leadership as well as the C/InS concept both have a predictive value for several success criteria and are correlated, but do not differentiate between each

other. However, one has to note that Bass' (1995) results are only based on the answers of 102 MBA students. Therefore, further examination of this question is pending.

Based on the aforementioned results, the following assumptions are made:

11. The transformational scales show high correlations with C and lower correlations with InS, MbEp and LF show negative correlations with the LBDQ scales.

12. CR and MbEa will show higher correlations with InS than with C.

13. Factorial analyses identify the two factors of C and IC as individual factors.

As already mentioned, Bass (1995) states that the C/InS approach does not differentiate between transformational and transactional leaders. However, transformational and transactional leadership are, in the conception of Bass (1985), not two ends of the same continuum. Hence, transformational leaders can also show a high frequency of transactional behaviors. What is, in my opinion, implied in the statement, though, is the differentiation between high transformational and low transformational (hence transactional) leaders. It can therefore be assumed that:

14. High and low transformational leaders will show no differences in the frequency of C and InS behaviors.

3.2.5. Method

Sample

Basis of this part of the investigation was the sample of Study 1 that has already been described in chapter 3.1.1. All in all, 1311 respondents were taken into account.

Measures

As already reported, the MLQ translation of Felfe and Goihl (2002b) was used. On the side of the LBDQ, several versions are reported in the literature. The LBDQ version used in this study was developed by Halpin (see Bryant, 2002) and used in his doctoral dissertation (for the German translation see Appendix B). All in all 40 items are given, 15 C items, 15 InS items, and 10 fill in-items. For the translation into German only the C and InS items were chosen due to economic reasons as the whole questionnaire for the organizational analysis was already quite long. The response format for this scale was a 5-point Likert-scale ranging from *never* (1) to *always* (5).

An examination of the factorial structure of the LBDQ items per CFA yielded poor fit indices (see Table 47 – again the specifications of the CFA were the same as the ones reported for the factorial validation of the MLQ). The reliability of the two full scales were α = .92 for C and α = .85 for InS. In order to examine the convergent validity of the MLQ, only the pure contents of the InS and C scales were needed. Therefore, the data set was randomly split into two halves. The modification indices provided by LISREL for the first halve were used to identify cross-loadings, items with high cross-loadings were then omitted. The reduced model was then tested on the second halve of the data set. All in all, 13 items stayed in the reduced item-set, seven InS and six C items. The fit indices provided good values for both the sub-sample and the full data set (see Table 47).

Table 47. CFA for the LBDQ data (Study 1)

Model	**N**	χ^2	**df**	**GFI**	**AGFI**	**CFI**	**NNFI**	**RMSEA**
(1)	1311	19277.36	435	.25	.20	.00	.00	.182
(2)	1311	4625.82	404	.78	.74	.78	.76	.089
(3)	673	171.83	64	.96	.95	.96	.96	.050
(4)	1311	353.09	64	.96	.94	.95	.94	.059

Note: (1) Null-Model; (2) full model with two correlated factors – 30 items; (3) reduced model with two correlated factors - 13 items for one sub-sample; (4) reduced model with two correlated factors for the full sample; GFI = goodness of fit index; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation.

The two factors correlate with r = .24 (p < .01), reliabilities are α = .88 for C and α = .77 for InS. The selected items and their resp. factor loadings for the reduced LBDQ model are presented in table 48.

Table 48. Factor loadings of the LBDQ items on their resp. factors (N_1 = 1311), completely standardized solution

	Initiating Structure	**Consideration**
Item 9	.45	
Item 11	.52	
Item 13	.52	
Item 14	.76	
Item 18	.60	

Table 48 continued		
	Initiating Structure	**Consideration**
Item 22	.50	
Item 28	.66	
Item 1		.78
Item 3		.76
Item 7		.71
Item 12		.76
Item 17		.73
Item 25		.73

3.2.6. Results

The correlations of the MLQ and LBDQ scales show higher correlations for C and the MLQ scales than for InS (see Table 49). Only MbEa correlates higher with InS than with C. For the transformational scales, all differences in correlations with the transformational scales between InS and C are significant. IC does not show the highest correlation with C (r = .70), IIA shows a higher correlation (r = .72). However, this difference is not significant. With all the other transformational scales, C has lower correlations.

Table 49. Scale intercorrelations for MLQ and LBDQ scales (N_I = 1311)

	IIA	**IIB**	**IM**	**IS**	**IC**	**CR**	**MbEa**	**MbEp**	**LF**
Initiating Structure	.32	.41	.35	.38	.32	.45	.33	-.20	-.26
Consideration	.72	.55	.55	.65	.70	.64	-.06	-.49	-.60

Note: IIA = idealized influence attributed; IIB = idealized influence behavior; IM = inspirational motivation; IS = intellectual stimulation; IC= individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; r > |.05|: p < .05; r > |.19|: p < .01.

Taking a look at the transactional scales, CR reveals a significantly higher correlation (p < .01) with C than with InS. As already mentioned, MbEa is the only scale that correlates higher with InS than with C. However, the correlation of InS with MbEa is not significantly different from the correlations of InS with the transformational scales. Both LBDQ scales show negative correlations with the passive scales of the MLQ,

whereby InS shows weaker correlations than C. Assumption 11 is therefore sustained, assumption 12 is rejected.

A factorial analysis of the IC and C items (PAA, eigenvalue > 1) reveals one factor that accounts for 51.3% of the variance. All items show high factor loadings on this factor (see Table 50).

Table 50. Factor loadings of the MLQ individual consideration and LBDQ consideration items (N_1 = 1311), PAA, eigenvalue > 1

	Factor I
MLQ_15	.65
MLQ_19	.66
MLQ_29	.73
MLQ_31	.75
LBDQ_1	.75
LBDQ_3	.76
LBDQ_7	.68
LBDQ_12	.73
LBDQ_17	.75
LBDQ_25	.70

An imposed two-factorial structure explains 58.2% of the variance. The two factors differentiate between the C and the IC items (see Table 51), however, the items show high cross-loadings.

Comparable to the procedure chosen in order to examine the factorial affiliation of the LF items, here, as well, three models were tested using CFA in order to further examine the factor affiliation of C and IC. The three tested models were one model with two method factors (MLQ and LBDQ) and one trait factor (considering), one model with two method factors and one model with only one trait factor. Only the method factors were allowed to be correlated with each other.

Table 51. Factor loadings of the MLQ individual consideration and LBDQ consideration items (N_1 = 1311), PAA, two-factor solution, varimax rotation

	Factor I	Factor II
MLQ_15	.34	**.60**
MLQ_19	.44	**.50**
MLQ_29	.36	**.72**
MLQ_31	.28	**.87**
LBDQ_1	**.70**	.34
LBDQ_3	**.65**	.40
LBDQ_7	**.68**	.26
LBDQ_12	**.68**	.34
LBDQ_17	**.61**	.43
LBDQ_25	**.71**	.26

The results show that the model with the two separate method factors and the trait factor has the best fit, the model with only one trait factor does not reach the criteria for a good fit (see Table 52).

Table 52. CFA for the MLQ individual consideration and LBDQ consideration items (N_1 = 1311)

Model	χ^2	*df*	GFI	AGFI	CFI	NNFI	RMSEA	$\Delta\chi^2$
(1)	227.80	25	.97	.93	.97	.95	.079	
(2)	342.29	35	.95	.91	.96	.94	.082	114.49 ($p < .01$)
(3)	796.65	35	.88	.81	.89	.86	.129	568.85 ($p < .01$)

Note: Model 1: two correlated method factors and one trait factor, model 2: two correlated method factors, model 3: one trait factor; GFI = goodness of fit index; AGFI = adjusted GFI; CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root mean square error of approximation; $\Delta\chi^2$ was computed by the subtraction of the χ^2 value of the first model from the χ^2-value of the other tested models

Furthermore, the χ^2 difference value indicates a significant fit improvement from the model with two correlated method factors to the model with two correlated method factors and one trait factor. All in all, these results support the idea that C and IC are two individual factors.

The MLQ scales have, as already shown, a problem with the distinct allocation of the items. Therefore, in order to differentiate the MLQ and LBDQ consideration scales, it

seems necessary to take all MLQ and LBDQ items into account in an EFA. The PAA (eigenvalue > 1) reveals a solution with seven factors that explains approximately 51% of the variance. The transformational scales of the MLQ form one factor with some of the CR and LF items, MbEa, MbEp, C and InS form their own factors (see Table 53). Therefore, when taking the whole framework of the MLQ into account, C and IC do not load on one common factor. IC, however, does not build its own individual factor. However, C has high cross-loadings on the factor that contains the IC items. Taking the results of all the factorial analyses together, assumption 13 is retained.

Table 53. Factor loadings of the MLQ and LBDQ items (N_1 = 1311), PAA, varimax rotation

	Factor	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**
Item	Scale							
MLQ_31	IC	**.78**	.21	.23	.16			
MLQ_25	IIA	**.71**	.36	.30	.16			
MLQ_29	IC	**.70**	.31	.19	.15		-.11	
MLQ_30	IS	**.70**	.17	.26	.15		.10	
MLQ_10	IIA	**.67**	.34	.21	.10	.21		.14
MLQ_32	IS	**.65**	.11	.30	.21		.16	
MLQ_15	IC	**.63**	.20	.25	.19	.25	.16	.19
MLQ_21	IIA	**.62**	.30	.27	.21			
MLQ_16	CR	**.61**	.15	.15	.27	.30	.11	.12
MLQ_18	IIA	**.57**	.34					.14
MLQ_23	IIB	**.57**	.23			.13	.15	
MLQ_34	IIB	**.56**	.25	.13	.18	.35		-.11
MLQ_8	IS	**.54**	.34	.35	.12			.26
MLQ_19	IC	**.53**	.38				-.19	
MLQ_26	IM	**.53**	.14	.21	.15	.35	.12	
MLQ_35	CR	**.53**	.33		.14	.34		-.15
MLQ_1	CR	**.52**	.36	.30				.32
MLQ_2	IS	**.50**	.19	.29	.25	.13		.34
MLQ_36	IM	**.50**	.36	.11	.12	.28		-.15
MLQ_33	LF	**-.49**	-.31	-.46	-.19		-.18	
MLQ_28	LF	**-.46**	-.20	-.37	-.23		-.21	
MLQ_7	LF	**-.46**	-.46	-.38				-.21
MLQ_9	IM	**.39**	.20	.24		.28		.13
MLQ_6	IIB	**.35**	.15		.10	.20		.22

Table 53 continued

	Factor	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**
Item	Scale							
LBDQ_25	C	.33	**.65**	.15			-.13	
LBDQ_7	C	.28	**.65**	.20				
LBDQ_12	C	.37	**.64**	.12				.15
LBDQ_1	C	.38	**.61**	.16				.13
LBDQ_3	C	.44	**.56**	.16		.12		
LBDQ_17	C	.39	**.55**	.35				.11
MLQ_20	MbEp	-.22	-.14	**-.67**				
MLQ_12	MbEp	-.33	-.20	**-.67**	-.17			
MLQ_3	MbEp	-.35	-.25	**-.57**	-.11			
MLQ_17	MbEp	-.12		**-.46**				
LBDQ_14	InS	.13	.12	.14	**.69**	.13		
LBDQ_28	InS	.14		.17	**.60**	.20	.13	
LBDQ_11	InS				**.58**			
LBDQ_22	InS				**.54**			
LBDQ_13	InS	.11			**.52**			
LBDQ_18	InS		.16	.20	**.51**		.20	
LBDQ_9	InS	.11	.10		**.41**			
MLQ_11	CR	.29			**.39**	.25		
MLQ_13	IM	.27		.12	.23	**.67**	.13	.12
MLQ_14	IIB	.33		.11	.28	**.63**	.25	.13
MLQ_24	MbEa	.17		.11	.27	.11	**.65**	
MLQ_22	MbEa			-.21			**.53**	
MLQ_4	MbEa	-.26	-.23	-.31	.14		**.43**	
MLQ_27	MbEa	.31		.15	.24	.14	**.39**	
MLQ_5	LF							**-.21**

Note: IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; C = consideration; InS = initiating structure.

In order to examine if C and InS differentiate between high and low transformational leaders, the mean of the transformational scales for all subjects, i.e. a transformational leadership variable, was computed. Then, three equal groups were built with low, medium and high values in transformational leadership. For high and low transformational leaders, significant differences were found for the C and InS

frequencies (see Table 54). For example, for C the mean value for low transformational leaders is $M = 2.78$, for high transformational leaders $M = .4.14$ ($p < .01$).

Table 54. Consideration and initiating structure means for high and low transformational leadership

Transformational leadership	N	Initiating Structure	Consideration
low	421	2.78	2.78
high	432	3.33	4.14

Furthermore, the results show that low transformational leaders show equal values in InS and C, high transformational leaders show differential values. T-tests for dependant samples show no significant differences between C and InS in the group of the low transformationally estimated leaders, whereas the difference for the high transformational leaders is significant ($p < .01$). Hence, assumption 14 has to be rejected.

Short digression No. 5: Taking the three-factorial model as a basis

As already shown, the MLQ nine-factorial model shows an unsatisfactory fit in Study 1, whereas the three-factor model yields good fit indices ($\chi^2 = 263$, $df = 62$, $p < 0.01$, $N_1 = 1311$, $AGFI = 0.91$, $CFI = 0.93$, $RMR = 0.061$). Taking the three factor model as a basis the following assumptions can be made for the relation with the LBDQ scales:

15. Charismatic goal orientation shows high correlations with the C and weaker correlations with InS, passive-avoidant leadership will show negative correlations with the LBDQ scales.

16. Management by exception will show higher correlations with InS than with C.

Results

The correlations show the expected pattern (see Table 55). All correlations are significant. Charismatic goal orientation shows a significantly higher correlation with C than with InS ($p < .01$), whereas Management by Exception shows the reverse pattern of

correlations. Passive-avoidant leadership correlates negatively with both LBDQ scales. Assumptions 15 and 16 are maintained.

Table 55. Scale intercorrelations of the three-factor model of the MLQ and the LBDQ scales (N_I = 1311)

LBDQ MLQ	Consideration	Initiating Structure
Charismatic goal orientation	.58	.40
Management by exception	.06	.42
Passive-avoidant leadership	-.50	-.20

r > |.05|: p < .05, r> |.19|: p < .01

3.2.7. Discussion

In order to examine the construct validity of an instrument, several aspects have to be taken into account. Concerning Cronbach and Meehl (1955) one part is the correlation with tests of a similar claim of validity.

The TLI also claims to measure transformational leadership and therefore seems useful when validating the MLQ. However, based on only the three reported studies, the evidence on the validity of the TLI is still quite limited. Nevertheless, all three American studies as well as the results obtained here support the proposed structure. The criterion-related validity is also supported in the few studies to date. Moreover, as the construction of the TLI is based on the theoretical views of several authors on transformational leadership, the contents of the TLI do tap what is commonly understood by a transforming leader.

The correlations with the MLQ show high correlations between the transformational scales and the core transformational scales of the TLI. Identifying a vision, providing an appropriate model and fostering the acceptance of group goals are contents that are also part of the transformational scales of the MLQ. Therefore, these results support the convergent validity of the MLQ. The other TLI facets of transformational leadership, however, show weaker correlations with the MLQ transformational scales. Hereby, PIS and IS_2 have relatively high correlations with the transformational scales of the MLQ, whereas HPE shows low correlations. Performance expectations are not part of Bass'

(1985) theory of transformational leadership; therefore the weak correlations are not surprising. Nevertheless, several authors do consider performance expectations to be part of transformational leadership. Therefore, a full range of leadership theory should contain this facet as well. PIS and IS_2 are scales that measure behavioral facets that are also measured by two transformational scales of the MLQ. However, the correlations of the corresponding scales do not show higher values than these scales show with other transformational scales. IS correlates higher with IAV or PAM than with PIS, the corresponding TLI scale. Similar results are obtained for IS_2, meaning that the scales of the MLQ might not be differentially valid. This is furthermore supported by the results of the factorial structure of the MLQ. The factorial analysis of the MLQ and TLI scales showed that whereas, with exception of the core transformational facets, most of the TLI scales built up their own factor, the MLQ scales collapsed into combined factors. Only IIA loaded on one factor with the TLI scales, the other TLI and MLQ items, except one IAV item, were not combined in any factor. IIA therefore seems to tap the same concept that is measured by the core transformational items of the TLI. IC and PIS, as well as IS and IS_2, did not build combined factors, hence they do not measure the same concept. CR also correlated higher with some of the transformational scales of the TLI than with CR_2, and, in the factorial analysis, bound with the transformational scales of the MLQ rather than with CR_2. The CR_2 items, however, built up their own factor and did not collapse into a factor with any of the transformational items. MbEa showed weak correlations with all TLI scales except HPE. The correlation between these two facets, however, is moderate. Nevertheless, it indicates that performance expectations do contain the active monitoring of standards. Bass (1985) argues that transactional leadership is the basis of transformational leadership, and that the latter is not possible without this basis. If HPE is considered as transformational behavior, the correlation between HPE and MbEa can be of support for this assumption.

The examination of the three-factorial model of the MLQ and its relation with the TLI scales showed the expected pattern. High correlations were obtained for all TLI scales and charismatic goal orientation, moderate correlations for the TLI scales and management by exception. It has to be noted, though, that the correlations of management by exception and the TLI scales are mostly higher then the correlations of MbEa and the TLI scales. The omission of one MbEa item therefore seems to

distinctively change the scale. Passive-avoidant leadership shows, as expected, negative correlations with the TLI scales.

However, due to the quantitatively weak empirical support of the TLI, all the results obtained here have to be interpreted with precaution.

Although Bass (1999b) states that the transformational/transactional leadership approach is conceptually independent of the LBDQ scales, both questionnaires measure leadership behavior. Therefore, taking one of the most popular leadership questionnaires in order to validate the MLQ, to date probably the most popular instrument, seemed to be self-evident. As, in my opinion, the LBDQ (although probably conceptually distinct) is construct-near to the MLQ as both measure leadership behavior. This has already been mentioned and this part of the investigation also counts for the examination of tests with similar claims of validity.

As expected, C shows higher correlations with the transformational scales than InS. Taking a look at all of the MLQ scales, only MbEa shows a higher correlation with InS than with C. Furthermore, the correlations of InS with the scales of the MLQ are comparable with the transformational scales and CR or MbEa. Therefore, it can not be confirmed that the LBDQ scales are conceptually nearer to the transactional scales of the MLQ as Bass (1995) states. As already mentioned, C seems to be closer to the concept of transformational leadership than InS, the two concepts share approximately 41% of variance. The closest proximity of C exists with the scales IIA and IC, the correlations with these scales are $r = .72$ and $r = .70$ respectively. Hence, it seems as if the considerate leader also gains the respect of his subordinates and instills pride in them. On the other hand, the two scales that share nearly the same name still seem to build their own factors, although being highly related. Taking a look at the item contents, C emphasizes behaviors as making the subordinates feel at ease, listening to them and supporting them in what they want to achieve. IC, on the other hand, describes a leader who wants to cultivate his subordinates and tries to achieve this by presenting new things to them or developing their strengths. The personal resp. individual component of these two scales might be the strongest link between them and therefore lead to high correlations. All in all the reported correlations support earlier findings by Keller (1992) and Seltzer and Bass (1990).

The examined three-factorial model of the MLQ shows a more clear-cut pattern of correlations with the scales of the MLQ. The transformational facet shows high positive, the passive-avoidant scale high negative correlations with C and InS, whereby C has higher correlations with these scales than InS. MbE, the transactional facet shows, as expected, a higher correlation with InS than with C. The more clear-cut pattern is the result of the aggregation of CR and transformational facets into charismatic goal orientation. CR as a transactional scale shows a correlational pattern with the LBDQ scales that is not distinguishable from that of the transformational scales, and highly correlates with these.

All in all, the MLQ seems to be convergent with the LBDQ scales, the old and the new leadership approaches therefore share common contents. The only weak correlation exists for MbEa and C. Although being considerate as well as actively monitoring mistakes and trying to prevent these are both leadership tasks, the two scales only correlate with r = .06. Apart from this, the correlations support the here chosen initial thought that the MLQ and the LBDQ are instruments to measure aspects of a similar claim of validity.

3.3 Part 3: Relation with external criteria and individual differences

According to Cronbach and Meehl (1955), the correlation of an instrument with several external criteria is an important part of the enquiry of its construct validation. Furthermore, it should be factorially analyzed with these external criteria. Therefore, the next chapter will analyze the relation of the MLQ scales with several external criteria and hence examine parts of the convergent and discriminant validity of the questionnaire. Moreover, individual differences in the questionnaire results will also be examined.

As already mentioned, the effectiveness of transformational leadership is acknowledged throughout the literature despite the unsatisfying results when dealing with its structure. Empirical studies including meta analyses show that the transformational facets have a stronger relation to positive success criteria than the transactional scales do (Judge & Piccolo, 2004; Lowe et al., 1996). In several studies, hierarchical regression analyses demonstrate that the addition of the transformational scales leads to a significantly

higher explanation of variance in diverse criteria. This is called the augmentation effect. The examined success criteria tap a wide variety, comprising subjective criteria as satisfaction, commitment, or OCB, as well as objective criteria (e.g. the financial outcome of banks; Geyer & Steyrer, 1998b). Generally, the hierarchy of correlations can be found for positive external criteria. The correlations reported for satisfaction, however, seem to vary if satisfaction is divided into satisfaction with the leader and satisfaction with the job (Judge & Piccolo, 2004). The transformational scales show higher correlations with satisfaction with the leader, whereas CR shows higher correlations for satisfaction with the job. Generally, the three MLQ internal criteria (satisfaction with the leader - SAT, effectiveness of the leader - EFF, and extra effort - EEF) show higher correlations with the MLQ scales than any other positive criterion.

In this study, several different criteria were chosen for the construct validation of the MLQ. As already mentioned, all studies were part of an organizational analysis of the respective administration offices. However, not all of the collected data seemed suitable for this analysis.

The criteria chosen here can be divided into two categories: variables used to examine the convergent validity and variables used to examine the discriminant validity.

In order to examine the convergent validity the MLQ internal criteria are taken into account. Furthermore, the satisfaction with the job, the affective commitment to the organization (Meyer et al., 2002), and the communication-based support by the focal leader are part of this investigation. As communication is sometimes mentioned as one of the main tasks of a leader (Yukl, 2002) and transformational leadership is based on a high communication rate, the MLQ scales should be related to this facet (Flauto, 1999). Furthermore, the communication-based support by the leader taps the leadership functions set up by the administration offices.

On the side of the discriminant validity, the continuance commitment was examined. As the content of this facet of commitment can be described as ‚I stay in this organization because I have no other choice’, it should not be influenced by leadership behavior (Bycio et al., 1995; Felfe et al., 2004). Moreover, the handling of customer contact as well as the perceived equality of opportunities for women in the administration office should also not be influenced by the leader’s behavior. The handling of customer contacts deals with the aggressiveness of the customer contact. The perceived equality

of opportunities for women is, in administrative offices, influenced more heavily by bureaucratic routines than by the leader. Reform processes in the administration should be perceived as being uncoupled from the focal leader. Therefore, no relation should occur between this variable and transformational/transactional leadership.

The following assumptions can be made:

17. The transformational scales will show high positive correlations with the criteria described as convergent to leadership behavior, CR and MbEa will show lower correlations with these scales, and MbEp and LF will show negative correlations.

Due to the form of presentation of the MLQ internal criteria right after the leadership items within the same section and the same item style, it can be assumed that:

18. The correlations of the MLQ leadership scales with the MLQ internal criteria will be higher than the correlations with the criteria that are not part of the MLQ.

Furthermore:

19. The MLQ scales will show no significant correlations with continuance commitment, handling of customer contact, and the perceived equality of opportunities for women.

20. Respondents who judge their focal leader as showing a high level of transformational leadership will be more affectively committed to their organization and show higher levels of job satisfaction than respondents who judge their focal leader as showing low levels of transformational leadership.

As already mentioned, the augmentation effect, the additional explanation of variance in the external criteria through the transformational scales, has often been confirmed (Bass & Avolio, 1995a; Vandenberghe et al., 2002; Waldman et al., 1987). The effect seems to be culturally universal and has also been confirmed in the few German studies (Felfe et al., 2004; Geyer & Steyrer, 1998a). Therefore, for this investigation, the confirmation of the augmentation effect is also assumed:

21. The transformational scales will explain a significant amount of variance in addition to the transactional scales in the external criteria that are part of the investigation of the convergent validity of the MLQ.

Apart from the examined convergence and similarities of the C/InS and the transformational/transactional paradigm, the question arose as to which of those

concepts was more useful in predicting success criteria. Furthermore, it should be examined if the TLI and the MLQ are comparable in that same question. As it is always argued that the augmentation effect is due to the transformational behavior, the transformational scales of the MLQ and the TLI should explain comparable amounts of variance. The augmenting effect above the transactional scales of the MLQ should also not differ. Concerning the LBDQ, Seltzer and Bass (1990) argue that "although initiation and consideration have been shown to be related with performance, there are additional behaviors displayed by transformational leaders who may achieve even higher levels of subordinate performance and satisfaction" (p. 964). As already mentioned, Bass (1995) states that "LBDQ initiation and consideration conceptually may substitute for transactional, but not for transformational leadership, for much additional variance in effectiveness is accounted for by transformational leadership when multiple regression analyses are applied" (p. 473).

Therefore, the following assumptions can be made regarding the external criteria SAT-job and COM-emo:

22. The transformational scales of the MLQ and the TLI will explain comparable amounts of variance in the external criteria and the augmenting effect of the transformational scales of the MLQ and the TLI above the transactional scales of the MLQ do not differ.

23. The transactional scales and the LBDQ scales explain comparable amounts of variance in the external criteria. Furthermore, the variance explained by the LBDQ scales will not augment the transactional scales.

24. Transformational leadership, as measured with the MLQ scales, will augment the LBDQ scales.

Furthermore, transformational leadership is supposed to be effective on all organizational levels (Bass & Avolio, 1993a). In accordance with this assumption, Fuller et al. (1996) found in their meta-analysis that the level of the focal leader did not moderate the relation between charisma and performance as well as effectiveness criteria. Therefore:

25. There is no difference in the correlations of transformational leadership and the external criteria for different levels of the focal leader.

26. For all hierarchical levels, respondents who judge their focal leader as showing a high level of transformational leadership will be more affectively committed to their organization and show higher levels of job satisfaction than respondents who judge their focal leader as showing low levels of transformational leadership.

3.3.1. Method

Sample

Basis of this part of the investigation are the samples of Study 1 and Study 4 that were already described in the previous chapters. In Study 1 1311 respondents and in Study 4 404 respondents are taken into account.

Measures

As in the studies before, the MLQ translation of Felfe and Goihl (2002b) was used. Furthermore, the LBDQ (Bryant, 2002) and the TLI (Podsakoff et al., 1996; Podsakoff et al., 1990) were also used for the following analyses.

The MLQ internal criteria SAT, EFF, and EFF were used for the examination of the convergent validity of the MLQ. Furthermore, affective (COM-emo) and continuance commitment (COM-con) were surveyed with the 'Fragebogen zur Erfassung von affektivem, kalkulatorischem und normativem Commitment gegenüber der Organisation, dem Beruf/der Tätigkeit und der Beschäftigungsform' (Questionnaire for the measurement of affective, continuance, and normative commitment to the organization, the occupation and the form of employment, COBB; (Felfe, Six, Schmook, & Knorz, 2003). Satisfaction with the job (SAT-job) was measured as a deficit value that expresses the difference between wanted and obtained conditions (Felfe, Resetka, & Liepmann, 1994). Negative values stand for contents that are more wanted than fulfilled, positive values mark over fulfilled conditions. Several facets of satisfaction are assessed with this questionnaire. For this investigation, however, an overall index for SAT-job is computed. Communication-based support by the leader (COMMU) was also assessed with four items developed for Study 1 in order to tap the leadership tasks formulated by German administration offices (see Appendix C).

On the side of the criteria for the examination of the discriminant validity of the MLQ, the handling of customer contact (CONT) as well as the perceived equality of

opportunities for women in the administration office (EQUAL) were used besides the continuance commitment. The two criteria were both assessed with items developed by Felfe et al. (1994).

All scales range from 1 to 5. For the commitment scales, CONT, and EQUAL the range varied from 1 = *is not the case* to 5 = *is the case*, for COMM the range varied from 1 = *is not true* to 5 = *is true*. For SAT-job, the part where the respondents had to mark how important certain conditions were for them varied from 1 = *not important at all* to 5 = *very important*, the part where the fulfillment of conditions was rated varied from 1 = *not fulfilled* to 5 = *completely fulfilled*. For the criterion SAT-job, the importance of the conditions was subtracted from the fulfillment of the resp. conditions and the mean of these difference values was then computed. Therefore, SAT-job can range from -4 to +4, whereby positive values stand for a mean over fulfillment of all conditions and therefore assumed job satisfaction, negative values mark a mismatch between the importance and the fulfillment of the conditions and show dissatisfaction.

The reliabilities for the criteria chosen here are all satisfying - for both samples (see Table 56). The three MLQ-internal criteria show high correlations among each other.

Table 56. Means, standard deviations, and internal consistencies of the external criteria

		Study 1 (N_1 = 1311)			Study 4 (N_4 = 404)		
Scale	**No. of items**	***M***	***SD***	***α***	***M***	***SD***	***α***
SAT	2	2.99	1.12	.88	3.03	1.07	.90
EFF	4	2.80	.98	.88	2.92	.97	.89
EEF	3	2.64	1.02	.90	2.81	.92	.87
SUC	9	2.79	.96	.95	2.90	.92	.95
COM-emo	5	3.18	.97	.83	3.12	.98	.86
SAT-job	18	-1.30	.70	.89	-1.27	.76	.90
COMMU	4	3.85	.94	.86	-	-	-
COM-con	4	3.67	.99	.70	3.47	.97	.75
CONT	1	3.12	1.09	-	3.18	1.05	-
EQUAL	4	2.73	.69	.70	2.90	.74	.70

Note: SAT = satisfaction with the leader; EFF = effectiveness; EEF = extra effort; SUC = success (combined success criterion of the MLQ internal criteria); COM-emo = affective commitment; SAT-job = satisfaction with the job; COMMU = communication-based support by the leader; COM-con = continuance commitment; CONT = customer contact; EQUAL = perceived equality for women.

The examination of their intercorrelation in these two samples reveals a range between $r = .76$ and $r = .81$ for Study 1 and a range from $r = .81$ to $r = .84$. A PCA reveals one factor for these criteria in both studies. In order to assure comparability with results of earlier studies, the results for both, the three MLQ criteria as well as a combined criterion called *success* (SUC), are reported.

3.3.2. Results

The correlation pattern of the MLQ scales with the external criteria reveals high correlations with the MLQ internal criteria and COMMU, smaller correlations with COM-emo and SAT-job and small to zero correlations with COM-con, CONT and EQUAL (see Table 57). For the MLQ internal criteria and SAT-job, COM-emo and COMMU, the already mentioned hierarchy of correlations appears. The transformational scales and CR show the highest positive correlations with these criteria, followed by MbEa that shows small to zero correlations. MbEp and LF both correlate negatively with these criteria. Different to assumption 17, however, CR shows correlations with these criteria that are comparable to the transformational scales. In some cases, the correlations of CR are even higher. For example, in Study 1 CR correlates with EEF with $r = .70$, IIb correlates with EEF with $r = .63$, in Study 4 the correlations are CR with EEF: $r = .76$, IIb with EEF: $r = .72$. For both studies CR shows a significantly higher correlation with the criterion than the transformational scale ($p < .01$). Taking a look at the correlation pattern with all the external criteria, CR can not be distinguished from the transformational scales. Therefore, assumption 19 is not accepted.

The correlations of the transformational scales with the internal criteria are significantly higher than the correlations with SAT-job and COM-emo. COMMU shows significantly lower correlations with the transformational scales, except for IS, where the difference between the correlations with COMMU and SAT is not significant. Furthermore, COMMU shows higher negative correlations with MbEp and LF than the MLQ internal criteria. For SAT-job and COM-emo, this relation is the other way around.

MbEa shows an inconsistent pattern of correlations for the two examined samples. In Study 1, MbEa shows the highest correlations with SAT, EFF and SUC, the correlations with SAT-job and COM-emo are weak and not significant. On the other hand, the correlation with COMMU ($r = .08$) does not differ from MbEa's correlation with EFF,

but is significantly lower than the correlation with SAT. EEF shows no correlation with MbEa. In Study 4, COM-emo shows a significantly higher correlation with MbEa than this scale shows with any of the MLQ internal criteria. LF and MbEp show significantly higher negative correlations with the MLQ internal criteria than with COM-emo and SAT-job. COMMU correlates significantly higher with MbEp and LF than SAT does. The finding that CR correlates higher with job-related satisfaction and the transformational scales correlate higher with satisfaction with the leader can not be confirmed in this study (e.g. Study 1: CR with SAT-job: r = .24, with SAT: r = .68, mean correlation of the transformational scales with SAT-job: r = .32, with SAT: r = .70). Although the correlations for the transformational scales and CR with the criteria COM-emo and SAT-job are lower than those with the MLQ internal criteria, the pattern for COMMU or the other MLQ scales is not as clear. Therefore, assumption 18 has to be rejected.

Table 57. Scale intercorrelations of the MLQ and external criteria (N_1 = 1311; N_4 = 404)

	SAT	EFF	EEF	SUC	COM-emo	SAT-job	COMMU	COM-con	CONT	EQUAL
MLQ 5x										
IIa	.73 (.83)	.82 (.77)	.84 (.79)	.85 (.84)	.25 (.25)	.35 (.40)	.67	.00 (.03)	-.06 (.00)	.12 (.10)
IIb	.66 (.65)	.69 (.69)	.63 (.72)	.71 (.74)	.26 (.33)	.25 (.37)	.53	.03 (.05)	-.05 (-.04)	.10 (.09)
IM	.63 (.60)	.68 (.64)	.64 (.67)	.69 (.68)	.29 (.34)	.30 (.41)	.50	.03 (-.01)	-.09 (-.07)	.09 (.13)
IS	.72 (.75)	.78 (.78)	.74 (.79)	.80 (.83)	.20 (.26)	.32 (.44)	.70	-.01 (-.04)	-.10 (-.04)	.10 (.08)
IC	.74 (.80)	.80 (.82)	.80 (.79)	.84 (.85)	.26 (.26)	.39 (.50)	.67	-.05 (-.04)	-.07 (-.06)	.10 (.12)
CR	.68 (.72)	.74 (.77)	.70 (.76)	.76 (.80)	.24 (.28)	.29 (.45)	.62	.04 (-.05)	-.05 (-.05)	.14 (.09)
MbEa	.15 (.05)	.10 (-.05)	.00 (.07)	.10 (.06)	.02 (.15)	-.03 (-.03)	.08	.09 (.11)	.03 (-.05)	.02 (.03)
MbEp	-.45 (-.67)	-.57 (-.67)	-.59 (-.63)	-.56 (-.69)	-.12 (-.26)	-.25 (-.46)	-.54	.07 (-.01)	.13 (.06)	.02 (-.10)
LF	-.57 (-.69)	-.67 (-.67)	-.66 (-.65)	-.68 (-.71)	-.13 (-.23)	-.28 (-.41)	-.68	.05 (-.01)	.10 (.04)	.01 (-.08)

Note: The correlations of study 4 are given in brackets; IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; SAT = satisfaction with the leader; EFF = effectiveness; EEF = extra effort; SUC = success (combined success criterion of the MLQ internal criteria); COM-emo = affective commitment; SAT-job = satisfaction with the job; COMMU = communication-based support by the leader; COM-con = continuance commitment; CONT = customer contact; EQUAL = perceived equality for women.
Study 1: $r > |.05|$: $p < .05$, $r > |.06|$: $p < .01$; Study 4: $r > |.07|$: $p < .05$, $r > |.12|$: $p < .01$

The correlations with the criteria that are used in order to examine the discriminant validity also show no clear-cut pattern. For Study 1, the transformational scales, CR, and LF show no significant correlations with COM-con, but some of the transformational scales and CR show significant correlations with CONT and EQUAL. The transactional scales MbEa and MbEp show significant correlations with all three scales. In Study 4, the transformational scales, CR, and LF show no significant correlations with COM-con and CONT, but they correlate significantly with EQUAL. MbEa shows a significant correlation with COM-con, but not the other two criteria, MbEp correlates negatively with EQUAL ($p < .05$) but shows no significant correlation with COM-con and CONT. All in all, assumption 19 has to be rejected.

Factorial analysis (PAA, eigenvalue > 1, oblique rotation) shows that the leadership items and all of the MLQ internal criteria load on a common factor in both studies. In Study 1 the COMMU items also load on this factor. The other criteria do not load on this factor, apart from a few SAT-job items. Hence, the MLQ internal criteria and COMMU can not be distinguished from leadership and are therefore excluded from the following analyses.

According to assumption 20, respondents who judged their leaders as being highly transformational are more affectively committed to their organization than respondents who judged their leaders as showing a low frequency of transformational leadership (see Table 58). Furthermore, the respondents with the high transformational leaders experience fewer differences between the wanted and obtained criteria of job satisfaction. For both studies and criteria, the differences between the low and high transformational leaders are significant ($p < .01$).

Table 58. Job satisfaction and affective commitment means for high and low transformational leadership

	Study 1			Study 4		
	N	**SAT-job**	**COM-emo**	***N***	**SAT-job**	**COM-emo**
Transformational leadership						
low	421	-1.59	2.87	142	-1.73	2.77
high	432	-1.05	3.49	140	-.86	3.45

Note: SAT-job = satisfaction with the job, COM-emo = affective commitment

As mentioned above, the augmentation effect is confirmed in several studies even though the factor model is not. Study 1 and Study 4 also confirm this effect (see Table 59). In this study, the structural equation approach to hierarchical regression as described by Bentler and Satorra (2000) was used. The results show that the transformational scales provide a significant additional explanation of variance in the examined criteria COM-emo and SAT-job that amounts up to 8% of variance for SAT-job in Study 4. Assumption 21 is therefore retained.

Table59. Hierarchical regression analysis for the MLQ scales on COM-emo and SAT-job

Version	**Number of Items**	**Criterion**	**Model**	R^2	**Change in** R^2
Study 1: N_1 = 1311	36	COM-emo	1	.02	.02**
			2	.06	.04**
			3	.09	.03**
		SAT-job	1	.08	.08**
			2	.11	.03**
			3	.15	.04**
Study 4: N_4 = 404	36	COM-emo	1	.05	.05**
			2	.09	.04**
			3	.14	.05**
		SAT-job	1	.17	.17**
			2	.25	.08**
			3	.33	.08**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: laissez-faire; Model 2: contingent reward, management by exception active, management by exception passive, laissez-faire; Model 3: idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized consideration, contingent reward, management by exception active, management by exception passive, laissez-faire.*

In order to examine which of the MLQ scales are important for the prediction of COM-emo and SAT-job, another regression analysis was performed whereby all scales were entered at the same time. IS and CR provide no significant β-weights for both criteria (see Table 60). IIA, MbEa, MbEp, and LF provide one significant β-weight each. The most important scale for the prediction of COM-emo and SAT-job seems to be IM as it has significant β-weights for both criteria in both studies. IC does not provide a

significant β-weight for COM-emo in Study 4, but has the second highest β-weight for this criterion in Study 1.

Table 60. Regression weights for the MLQ scales

	Study 1 (N_1 = 1311)		Study 4 (N_4 = 404)	
	COM-emo	**SAT-job**	**COM-emo**	**SAT-job**
MLQ 5x				
IIa				-.21
IIb		-.09	.19	
IM	.19	.13	.21	.17
IS				
IC	.13	.28		.41
CR				
MbEa			.11	
MbEp				-.23
LF	.11			

Note: IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; COM-emo = affective commitment; SAT-job = satisfaction with the job; only significant β-weights are displayed.

In order to compare the explanatory amount of the transformational scales of the MLQ and the TLI, regression analyses were computed with only the transformational scales of either the MLQ or the TLI. The results show that in Study 4 the transformational scales of the MLQ explain 13% of variance in COM-emo and 27% of variance in SAT-job, whereas the transformational scales of the TLI explain 14% in COM-emo and 26% in SAT-job. The amount of variance explained by the transformational scales of the MLQ and TLI does not differ significantly. Furthermore, it was examined if the transformational scales of the TLI add an amount of explained variance in the two criteria to the MLQ transactional scales and LF of the MLQ that is comparable to the additionally explained variance of the transformational scales of the MLQ. The hierarchical regression analyses show that the MLQ transformational scales add 5% of variance in COM-emo, whereas the TLI transformational scales add 11% in this criterion (see Table 61). For SAT-job, on the other hand, the MLQ scales add 8% of variance, the TLI scales add 5%. Hence, the TLI transformational scales explain a

significantly higher amount of variance than the MLQ scales in COM-emo do ($p < .01$). Therefore, assumption 22 has to be rejected.

Table 61. Hierarchical regression analysis for the MLQ and TLI scales on COM-emo and SAT-job (N_4 = 404)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
MLQ	36	COM-emo	1	.05	.05**
			2	.09	.04**
			3	.14	.05**
		SAT-job	1	.17	.17**
			2	.25	.08**
			3	.33	.08**
TLI	38	COM-emo	1	.05	.05**
			2	.09	.04**
			4	.20	.11**
		SAT-job	1	.17	.17**
			2	.25	.08**
			4	.30	.05**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: laissez-faire; Model 2: contingent reward, management by exception active, management by exception passive, laissez-faire; Model 3: idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized consideration, contingent reward, management by exception active, management by exception passive, laissez-faire; Model 4: identifying and articulating a vision; providing an appropriate model; fostering the acceptance of group goals; high performance expectations; providing individualized support; intellectual stimulation, contingent reward, management by exception active, management by exception passive, laissez-faire.*

However, hierarchical regression analyses also show that both the TLI and the MLQ transformational scales augment the resp. other transformational scales significantly (see Table 62). The TLI scales add 6% resp. 7% to the explained variance in the MLQ scales for COM-emo and SAT-job, the MLQ scales add 4% resp. 8% of variance.

Table 62. Hierarchical regression analysis of the MLQ and TLI transformational scales on COM-emo and SAT-job (N_4 = 404)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
MLQ and TLI	42	COM-emo	1	.12	.12**
			2	.18	.06**
		SAT-job	1	.26	.26**
			2	.33	.07**
MLQ and TLI	42	COM-emo	3	.13	.14**
			4	.18	.04**
		SAT-job	3	.25	.25**
			4	.33	.08**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized consideration; Model 2: idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized consideration, identifying and articulating a vision; providing an appropriate model; fostering the acceptance of group goals; high performance expectations; providing individualized support; intellectual stimulation; Model 3: identifying and articulating a vision; providing an appropriate model; fostering the acceptance of group goals; high performance expectations; providing individualized support; intellectual stimulation; Model 4: identifying and articulating a vision; providing an appropriate model; fostering the acceptance of group goals; high performance expectations; providing individualized support; intellectual stimulation, idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized consideration.*

In order to compare the explanatory power of the transactional scales of the MLQ and the LBDQ scales for COM-emo and SAT-job, here, too, regression analyses were computed. The results show that in Study 1 the MLQ scales CR, MbEa, and MbEp explain 6% of variance in COM-emo and 10% of variance in SAT-job, whereas the LBDQ explains 5% in COM-emo and 11% in SAT-job. None of the explained variance parts differs significantly. Hierarchical regression analyses further show that the LBDQ scales add a small but significant amount of variance to the transactional scales in both COM-emo and SAT-job (see Table 63). Consequently, assumption 23 has to be rejected.

Table 63. Hierarchical regression analysis of the transactional MLQ and LBDQ scales on COM-emo and SAT-job (N_I = 1311)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
MLQ and LBDQ	25	COM-emo	1	.06	.06**
			2	.07	.01**
		SAT-job	1	.10	.10**
			2	.13	.03**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: contingent reward, management by exception active, management by exception passive; Model 2: contingent reward, management by exception active, management by exception passive, consideration, initiating structure.*

On the other hand, the transformational scales add a significant amount of variance to the scales of the LBDQ (see Table 64). For COM-emo, transformational leadership as measured with the MLQ adds 4%, for SAT-job it adds 5% of variance. Putting the transformational scales into the analyses first shows that the LBDQ scales do not augment transformational leadership for COM-emo, but add 1% of variance ($p < .05$) to the explanation of SAT-job. Assumption 24 is maintained.

Table 64. Hierarchical regression analysis of the transformational scales of the MLQ and the LBDQ scales on COM-emo and SAT-job (N_I = 1311)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
MLQ and LBDQ	35	COM-emo	1	.06	.06**
			2	.10	.04**
		SAT-job	1	.12	.12**
			2	.17	.05**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: consideration, initiating structure; Model 2: consideration, initiating structure, idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized consideration.*

In accordance with the findings of Fuller et al. (1996) it was further examined if the correlations of the MLQ scales with the external criteria COM-emo and SAT-job vary dependent on the level of the focal leader. In German bureaucratic administrations, the hierarchical level of a person is closely matched by the group of salary. In Study 1, five

levels of salary were distinguished. However, according to the composition of the work group, several levels of salary could report to the same leader. In order to reduce the problem of confounded groups, only the lowest, middle and highest salary groups are taken into account, assuming that the overlapping of focal leaders is therefore minimal. As in Study 4 only three levels of salary were distinguished the following analyses refer only to Study 1.

Table 65. Scale intercorrelations of the MLQ transformational scales and external criteria for three groups of salary (Study 1)

	COM-emo	**SAT-job**
MLQ 5x		
IIa	.37 / .25 / .44	.48 / .38 / .09
IIb	.65 / .26 / .51	.47 / .28 / .02
IM	.43 / .30 / .59	.32 / .34 / .19
IS	.31 / .26 / .50	.46 / .37 / .11
IC	.38 / .29 / .58	.34 / .38 / .23
Transformational leadership	.48 / .30 / .58	.46 / .40 / .14

Note: Correlations for the lowest salary group (N = 15) are given first, followed by the middle salary (N = 641) and the highest salary group (N = 32). Note: IIA = idealized influence attributed; IIB = idealized influence behavior, IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration; CR = contingent reward; MbEa = management by exception active; MbEp = management by exception passive; LF = laissez-faire; COM-emo = affective commitment; SAT-job = satisfaction with the job. Low salary group: $r > .64$, $p < .01$; middle salary group: $r > .25$, $p < .01$; high salary group: $r > .43$, $p < .05$; $r > .50$, $p < .01$.

Although the correlations of the transformational scales with COM-emo are lower for the middle salary group than the two other groups (see Table 65), none of the differences in correlations are significant. For SAT-job, the highest salary group shows the lowest correlations with the transformational scales. However, the correlations with SAT-job also do not differ from the correlations in the other two salary groups. Hence, assumption 25 is retained.

In order to further examine the effectiveness of transformational leadership in all three salary groups, hence hierarchical levels, the means of COM-emo and SAT-job were compared for high and low transformational leaders. The results of the T-test for independent samples show that for the low salary group, the respondents with the high

transformational leaders are more affectively committed and show less dissatisfaction with their job (see Table 66). However, the differences are not significant. For the middle salary group, the means show significant differences in the expected direction ($p < .01$). The high salary group shows a significant difference for COM-emo ($p < .01$), but not for SAT-job. Assumption 26 therefore has to be rejected.

Table 66. Job satisfaction and affective commitment means for high and low transformational leadership in three salary groups (Study 1)

	Low salary (N = 15)		Middle salary (N = 631)		High salary (N = 32)	
	SAT-job	COM-emo	SAT-job	COM-emo	SAT-job	COM-emo
Transformational leadership						
low	-1.19	3.50	-1.70	2.91	-.88	2.41
high	-.87	4.40	-1.15	3.58	-.89	3.39

Note: SAT-job = satisfaction with the job; COM-emo = affective commitment.

Short digression No. 6: Comparison with the three-factorial model

The three-factorial model has, in comparison to the nine-factorial model of the MLQ, less items and not all of the original MLQ scales are part of the reduced questionnaire. For example, the IC items are not part of the three-factorial model, but are an important predictor for the here examined criteria COM-emo and SAT-job (see Table 60). Therefore, the next part will take a look at the possible loss of information of the three-factorial model in comparison to the nine-factorial model concerning the two other examined leadership questionnaires and the two criteria COM-emo and SAT-job.

As some of the transformational facets are missing, it can be assumed that the three-factorial model explains less variance in the two criteria in comparison to the nine-factorial model or the TLI. Furthermore, the hierarchy of correlations should be more clear-cut.

Therefore:

27. Charismatic goal orientation will show high-positive correlations with COM-emo and SAT-job, passive-avoidant leadership will show negative, and MbE will show no

significant correlations with these two criteria. No significant correlations with all three scales will be obtained for COM-con and EQUAL.

28. The nine-factorial model of the MLQ will explain significantly more variance in the two criteria COM-emo and SAT-job than the three-factorial model.

29. The transformational scales of the TLI will augment the three factors of the reduced model of the MLQ for these two criteria. The transformational scale of the three-factorial model will not augment the transformational scales of the TLI.

The augmenting effect of the transformational scale above the LBDQ scales should still be existent.

30. The LBDQ will augment the passive-avoidant and MbE scale of the three-factorial model. The transformational scale of the three-factorial model will augment the LBDQ for the two criteria COM-emo and SAT-job.

Results

The correlational pattern of the three-factorial model shows high positive correlations with COM-emo and SAT-job for charismatic goal orientation, no correlations with management by exception, and negative correlations for passive-avoidant leadership (see Table 67). On the side of the discriminant criteria, however, management by exception and passive-avoidant leadership correlate significantly with COM-con and charismatic goal orientation has a significant correlation with EQUAL. Therefore, assumption 27 is rejected.

Table 67. Scale intercorrelations of the three-factor model of the MLQ with COM-emo and SAT-job (N_1 = 1311)

	COM-emo	**SAT-job**	**COM-con**	**EQUAL**
Charismatic goal orientation	.30	.31	.03	.12
Management by exception	.03	-.04	.07	.00
Passive-avoidant leadership	-.10	-.26	.08	-.02

r> |.06|: *p* < .01.

Juxtaposing the regression analyses for the nine-factorial and the three-factorial model, one can see that there is no significant difference in the portions of explained variance for Study 1 (Table 68). In Study 4, the three-factorial and the nine-factorial model explain comparable amounts of COM-emo, whereby the nine-factorial model explains with 14% a slightly higher portion of variance than the three-factorial model (12%) does. The difference, however, is not significant. For SAT-job, the nine-factorial model explains a noticeably higher amount, i.e. 9% more than the three-factorial model. The difference in explained variance, though, is not significant. Hence, assumption 28 has to be rejected.

Table 68. Regression analysis for the nine- and the three-factorial model of the MLQ on COM-emo and SAT-job (N_1 = 1311, N_4 = 404)

Version	**Criterion**	**Number of Items**	**Model**	R^2
Study 1: N_1 = 1311	COM-emo	36	9	.09
		13	3	.09
	SAT-job	36	9	.15
		13	3	.12
Study 4: N_4 = 404	COM-emo	36	9	.14
		13	3	.12
	SAT-job	36	9	.33
		13	3	.24

Note: SAT-job = satisfaction with the job; COM-emo = affective commitment; Model 9: nine-factorial model of the MLQ; Model 3: three-factorial model of the MLQ.

The augmenting effect of the MLQ transformational scales above the three-factorial model, however, clearly shows that the transformational scales of the MLQ add significant additional variance to the amount of variance explained by the three-factorial model in the two criteria (see Table 69).

Table 69. Hierarchical regression analysis for the transformational scales of the MLQ and the three-factorial model on COM-emo and SAT-job (N_1 = 1311, N_4 = 404)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
Study 1: N_1 = 1311	33	COM-emo	1	.10	.10**
			2	.12	.02**
		SAT-job	1	.09	.09**
			2	.13	.04**
Study 4: N_4 = 404	33	COM-emo	1	.09	.09**
			2	.14	.05**
		SAT-job	1	.15	.15**
			2	.26	.11**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: passive-avoidant leadership, management by exception, charismatic goal orientation; Model 2: passive-avoidant leadership, management by exception, charismatic goal orientation, idealized influence attributed, idealized influence behavior, inspirational motivation, intellectual stimulation, individualized support.*

The transformational scales of the TLI also add additional variance to the amount of variance explained by the three-factorial model (see Table 70). For COM-emo, the TLI scales add 8%, for SAT-job they add 7% of variance. Hence, assumption 34 is accepted. For both, COM-emo and SAT-job, charismatic goal orientation adds only the amount of nigh 1% of variance to the amount explained by the transformational scales of the TLI. As this change in R^2 is not significant, assumption 29 is maintained.

Table 70. Hierarchical regression analysis for the transformational scales of the TLI and the three-factorial model on COM-emo and SAT-job (N_4 = 404)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
Study 4: N_4 = 404	35	COM-emo	1	.09	.09**
			2	.17	.08**
		SAT-job	1	.15	.15**
			2	.22	.07**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: passive-avoidant leadership, management by exception, charismatic goal orientation; Model 2: passive-avoidant leadership, management by exception, charismatic goal orientation, identifying and articulating a vision, providing an appropriate model, fostering the acceptance of group goals, high performance expectations, providing individualized support.*

The LBDQ scales do augment the transactional scales of the nine-factorial model. As can be seen, this is also confirmed for the transactional and passive-avoidant scales of the three-factorial model (see Table 71). The LBDQ scales add an additional amount of 4% variance to the explanation of COM-emo and 5% variance to the explanation of SAT-job.

Table 71. Hierarchical regression analysis of the LBDQ and the three-factorial model of the MLQ on COM-emo and SAT-job (N_1 = 1311)

Version	Number of Items	Criterion	Model	R^2	Change in R^2
MLQ and LBDQ	20	COM-emo	1	.01	.01**
			2	.05	.04**
		SAT-job	1	.07	.07**
			2	.12	.05**
MLQ and LBDQ	20	COM-emo	3	.05	.05**
			4	.09	.04**
		SAT-job	3	.11	.11**
			4	.13	.02**

*** p< 0.01. Note: COM-emo = affective commitment; SAT-job = satisfaction with the job; Variables: Model 1: passive-avoidant leadership, management by exception; Model 2: passive-avoidant leadership, management by exception, consideration, initiating structure; Model 3: consideration, initiating structure; Model 4: consideration, initiating structure, charismatic goal orientation.*

The transformational scale of the three-factorial model, charismatic goal orientation, augments the LBDQ scales by 4% for COM-emo and 2% for SAT-job. Assumption 30 is therefore retained.

3.3.3. Discussion

The effectiveness of transformational leadership when identified by the augmentation effect, as already shown in numerous studies, can also be confirmed in this study. The transformational scales show high correlations with the positive criteria and augment the transactional and LF scales in hierarchical regression analyses. The results of, for example, Felfe (in press) or Waldmann et al. (1990) are therefore replicated. However, the additional amount of explained variance by the transformational scales is lower than the values reported in other studies (see e.g. Felfe, in press). This is due to the fact that

normally the augmentation effect is only examined for the transformational and transactional scales. LF is usually not part of the hierarchical regression analyses. In order to keep the results comparable with those obtained with the three-factorial model, LF was entered into these analyses as a first step. As LF and the transformational scales show high positive correlations with each other, a part of the explained variance in the criteria is already bound by LF as this scale is entered first into the hierarchical regression. On the other hand, the amount of variance explained by all the MLQ items was also of interest. With 9% to 33% of variance, the leadership scales do have a significant impact on the affective commitment and the satisfaction with the job. The effectiveness of transformational leadership is further supported by the results of the t-tests. Followers who judge their leaders as showing a low frequency of transformational behaviors are less affectively committed to their organization and less satisfied with their jobs than their colleagues are who judge their leaders to be highly transformational.

On the side of the discriminant validity, the results are ambiguous. All in all, the MLQ scales show lower correlations with COM-con, CONT, and EQUAL than with the positive criteria. However, significant correlations occur for all MLQ scales for at least one criterion in one of the studies. COM-con seems to be a discriminant criterion for the transformational scales. However, it shows significant but moderate correlations in both studies with MbEa. The other two criteria seem to be more influenced by leadership than expected. Although the promotion of subordinates is a much more standardized process in German administrations than in private organizations, there still can be the feeling that a transformational leader might recommend individuals to a higher or lower degree. This is supported by the analysis of correlations in the male and female subsamples. In Study 1, for example, the transformational scales and CR show significant correlations with EQUAL that range from $r = .13$ to $r = .17$ for the female raters, whereas in the male subsample, none of the MLQ scales have significant correlations with EQUAL. Therefore, the female subordinates might have the feeling that a transformational leader can recommend them for promotion, whereas the male subordinates see the promotion chances for female colleagues as being independent of leadership. A possible explanation for the correlations of the MLQ scales with CONT is that, although the actual aggressiveness of the customer contact is independent of the leadership style of the superior, coping strategies or effective procedures dealing with

the exposure to customers can be provided by transformational leaders. A more transformational leadership style might therefore lead to a less intensely perceived aggressiveness of the customers.

The augmentation effect seems to be a phenomenon that is not limited to the MLQ transformational scales. MLQ and TLI transformational scales both explain comparable amounts of variance in the two criteria, COM-emo and SAT-job. Furthermore, they both augment the transactional scales of the MLQ and LF by 5% to 11%. Although the augmentation effect is comparable for SAT-job, the results show that the TLI transformational scales add significantly more variance to the MLQ scales for COM-emo. The TLI transformational scales therefore seem to contain leadership behaviors that share common variance with COM-emo that are not part of any of the MLQ transformational, transactional or LF scales. It has to be noted, though, that both ways of measuring transformational leadership augment the respective other. Therefore, both questionnaires seem to tap behavior facets of transformational leadership that are not part of the respective other instrument, but are important for SAT-job or COM-emo, in the sense that they share common variance.

Bass' (1995) statement that the LBDQ scales can substitute for the transactional scales of the MLQ is confirmed insofar as these scales explain comparable amounts of variance in COM-emo and SAT-job. However, the hierarchical regression analyses show that the LBDQ scales still augment the transactional scales by a small but rather significant amount. A mere substitution of the transactional scales is therefore not the case. On the other hand, Bass' expectation that the transformational scales add significant amounts of variance to the LBDQ scales is confirmed. Moreover, it can be shown that the LBDQ scales do not augment the transformational scales for COM-emo and only slightly for SAT-job. As a consequence, one can conclude that transformational leadership represents a meaningful extension of leadership behavior facets.

Transformational leadership is supposed to be equally effective on all organizational levels (Fuller et al., 1996). It can be shown that transformational leadership shows no significantly different correlations for the three examined hierarchical levels. However, the results of this study show that, depending on the hierarchical level, satisfaction with the job does not differ either for high or low transformational leaders. Although there is a significant difference for the middle salary group, people in high and low hierarchical

levels seem to be equally satisfied no matter if their leader shows a high or low frequency of transformational behavior. Yet, the hierarchical level is always more or less linked to the age. Clark, Oswald, and Warr (1996) found a U-shaped relationship between age and job satisfaction that is also perceived in this study: for the middle salary group, dissatisfaction is higher than for the other two groups. It might therefore be the case that satisfaction is influenced more by other variables than leadership, respectively that transformational leadership is important for a higher job satisfaction only in age groups that are less satisfied with their jobs. If job satisfaction is on a high level anyway, having a transformational leader will not make a significant difference. These results, however, have to be interpreted with precaution, as the low and high salary groups are, in comparison to the middle salary group, very small. Hence, sample effects can be more noticeable in these two than in the middle salary group.

The regression analyses show that, for the criteria examined in this study, the reduced set of items in the three-factorial model has barely lost information in comparison to the original model. While the total amount of explained variance is lower for the three-factorial model, especially for Study 4, the differences are not significant. Although the most important facets for the prediction of affective commitment and satisfaction with the job for the nine-factorial model seem to be IM and IC, the omission of IC items does not seem to be that important. However, there is some information lost, as the transformational scales of the nine-factorial model significantly augment the three-factorial model of the MLQ. Concerning the comparison of the three-factorial model with the TLI and LBDQ scales, the results show that transformational leadership as measured with the TLI shares a significant additional amount of variance with COM-emo and SAT-job. On the other hand, charismatic goal orientation only adds a small amount of variance to the scales of the TLI. Furthermore, the obtained results for the comparison of the transactional and transformational facets of the nine-factorial model with the LBDQ scales are confirmed with the three-factorial model.

The results of this part of the investigation, however, are restricted in interpretability mainly due to two reasons. First of all, a strong problem of colinearity occurs for the regression analyses. As was shown in chapter 3.1.1, high correlations are found for the transformational scales and CR. Furthermore, the scales of the other two leadership questionnaires also show high correlations with the MLQ scales, especially the transformational scales of the TLI. On the other hand, the effectiveness of

transformational leadership is examined using criteria that are collected at the same time as the leadership ratings using the same method (questionnaires) and the same source. Hence, the here presented data are susceptible to a common-source mono-method bias, resulting in overestimated relations.