

Marc Esser-Noethlichs

# **Sensitivity towards strangeness (STS)**

Development of a concept-based measuring instrument in the context of Intercultural Movement Education

DISSERTATION FROM THE NORWEGIAN SCHOOL OF SPORT SCIENCES • 2011

ISBN nr 978-82-502-0456-0



*„[...] denn als fremd wird der Fremde nur in der Fremde wahrgenommen und jeder der sich fremd fühlt ist nur so lange ein Fremder, bis er sich nicht mehr fremd fühlt, denn dann ist er kein Fremder mehr. [...]“ (Valentin, 1984:488)*

Karl Valentin, cabaret artist and author from Munich (1882-1948).



To Sarah and Lina



---

## Acknowledgements

I want to thank the Norwegian School of Sport Sciences for giving me the opportunity of a three years scholarship. This opportunity was an unequalled privilege.

This dissertation would not have been the same result without a number of individuals supporting and inspiring me through this scientific journey. I would like to thank Prof. Ralf Erdmann for giving me the opportunity to participate in his project “Intercultural Movement Education”. I would like to thank all participants in my studies. Special thanks to Prof. Petra Gieß-Stüber and her team in Germany for co-operation, constructive feedback, and data collections in her EU-project.

Furthermore, I would also like to express my gratitude to the Chancellor of the Norwegian School of Sport Science Prof. Sigmund Loland, the head of the department of coaching and psychology Associate Prof. Nicolas Lemyre and the Associate Prof. Lars Tore Rongland. Thank you for your support.

I experienced great support during the final stage of this dissertation at the Oslo University College. I would like to thank Prof. Sigmund Ongstad for critical reading and supportive feedback. He has been my advisor during the final stage of my PhD project. Thanks to Associate Prof. Rolf Fasting for constructive feedback on my methods and helpful discussions. I would like to thank Prof. Harald Jarning and Head of Studies Unni Hagen for the help I needed to finish my thesis besides all other obligations at the Oslo University College. I am very grateful to Calle Bergsjø the head of the department of sport and physical education at the Oslo University College. Without your understanding and efficient all-round support I might still be finishing my thesis. Thanks to all of my colleges of this department.

Special thanks to Turid Pellfoch for proof-reading my thesis and good discussions. I appreciate your direct comments and constructive feedback ☺

Thanks to Dr. Michael Lausberg for feedback, discussions and friendship.

I would like to express my love and gratefulness to my wife Sarah and my daughter Lina. You are the backbone I needed. Thank you for your patience, constructive discussions, support in difficult periods and for all your understanding through this long and challenging journey.





## Table of Contents

<b>ACKNOWLEDGEMENTS</b> .....	<b>VII</b>
<b>TABLE OF CONTENTS</b> .....	<b>IX</b>
<b>LIST OF TABLES</b> .....	<b>XIII</b>
<b>LIST OF FIGURES</b> .....	<b>XVII</b>
<b>LIST OF ABBREVIATIONS</b> .....	<b>XIX</b>
<b>ABSTRACT</b> .....	<b>XXI</b>
<b>PART I: THEORETICAL FOUNDATIONS</b> .....	<b>1</b>
<b>1. INTRODUCTION</b> .....	<b>3</b>
<b>2. REVIEW OF RELATED LITERATURE</b> .....	<b>13</b>
2.1. CONCEPTIONS OF STRANGENESS.....	13
2.2. STRANGENESS IN THE REALM OF SPORT AND PHYSICAL EDUCATION .....	16
2.3. RELATED OPERATIONAL MODELS AND MEASURING INSTRUMENTS.....	19
<b>3. THEORETICAL STARTING POINT: CULTURAL DIFFERENCE, STRANGENESS AND FACETS OF IDENTITY</b> .....	<b>25</b>
3.1. THE DIFFERENCE: CENTRAL CATEGORY WITHIN INTERCULTURAL EDUCATION .....	25
3.2. REACTION PATTERNS TOWARDS DIFFERENCES AND STRANGENESS .....	27
3.3. FACETS OF IDENTITY.....	31
<b>4. FURTHER THEORETICAL CONSIDERATIONS AND A CONCEPTUALIZATION OF “SENSITIVITY TOWARDS STRANGENESS (STS)”</b> .....	<b>39</b>
4.1. CONNOTATION OF STRANGENESS .....	40
4.2. FACETS OF STRANGENESS .....	41
4.2.1. <i>Strangeness, a social construction</i> .....	42

4.2.2.	<i>Strangeness, a specific relationship</i> .....	44
4.2.3.	<i>Strangeness, a concept of uncertainty</i> .....	45
4.3.	DEALING WITH DIFFERENCE AND STRANGENESS .....	47
4.4.	MEANINGFUL EXPERIENCES OF STRANGENESS .....	52
4.5.	SENSITIVITY TOWARDS STRANGENESS (STS) .....	54
<b>5.</b>	<b>PRACTICAL IMPLICATIONS OF SENSITIVITY TOWARDS STRANGENESS IN THE CONTEXT OF INTERCULTURAL MOVEMENT EDUCATION</b> .....	<b>61</b>
<b>PART II: EMPIRICAL INVESTIGATIONS</b> .....		<b>73</b>
<b>6.</b>	<b>METHODOLOGICAL PROCEDURE OF DEVELOPING THE STSQ</b> .....	<b>75</b>
6.1.	METHODOLOGICAL BACKGROUND IDEA OF DEVELOPING A CONCEPT-BASED MEASURING INSTRUMENT .....	76
6.2.	DESIGN AND STATISTICAL PROCEDURES FOR THE DEVELOPMENT OF THE STSQ-ITEMS AND STSQ-SCORES .....	80
6.3.	OPERATIONAL MODEL OF STS .....	88
6.3.1.	<i>Emotional meaning (EM)</i> .....	90
6.3.2.	<i>Cognitive meaning (CM)</i> .....	91
6.4.	FIRST STRUCTURED ITEM POOL: SENSITIVITY TOWARDS STRANGENESS QUESTIONNAIRE (STSQ) .....	93
6.4.1.	<i>STSQ, part I: measuring emotional sensitivity toward strangeness</i> .....	94
6.4.2.	<i>STSQ, part II: awareness of people's argumentation strategies and attributions styles with respect to relevant situations</i> .....	103
6.4.3.	<i>STSQ, part III: relevant attitudes, and understandings</i> .....	110
6.5.	PRE-STUDIES: HELPING PROCEDURES FOR ITEM POOL COLLECTION .....	115
6.5.1.	<i>Design</i> .....	115
6.5.2.	<i>Method</i> .....	116
6.5.3.	<i>Results and discussion</i> .....	119
6.6.	PILOT STUDY: INITIAL ITEM ANALYSES AND EXPLORATION OF RELIABILITY AND VALIDITY HINTS .....	122
6.6.1.	<i>Aim</i> .....	122
6.6.2.	<i>Pre-considerations of the research design and methods</i> .....	123
6.6.3.	<i>Expectations</i> .....	124
6.6.4.	<i>Procedure</i> .....	126
6.6.5.	<i>Results</i> .....	127

---

6.6.6.	<i>Discussion and conclusions</i> .....	138
6.7.	REPEATED MEASUREMENT DESIGN .....	144
6.7.1.	<i>Pre-considerations of the research design and methods</i> .....	144
6.7.2.	<i>Expectations</i> .....	148
6.7.3.	<i>Procedure</i> .....	148
6.7.4.	<i>Results</i> .....	151
6.7.5.	<i>Discussion and conclusions</i> .....	159
6.8.	APPLICATION OF THE STSQ IN A NORWEGIAN HIGH SCHOOL .....	166
6.8.1.	<i>Considerations of the research design</i> .....	166
6.8.2.	<i>Expectations</i> .....	168
6.8.3.	<i>Procedure</i> .....	169
6.8.4.	<i>Results</i> .....	169
6.8.5.	<i>Discussion and conclusions</i> .....	172
6.9.	APPLICATION OF THE STSQ WITHIN THE COMENIUS PROJECT 2.1 “THE DEVELOPMENT OF INTERCULTURAL COMPETENCE THROUGH SPORTS IN AN EXPANDING EUROPEAN UNION “ .....	177
6.9.1.	<i>Considerations of the given research design</i> .....	179
6.9.2.	<i>Measuring instruments</i> .....	181
6.9.3.	<i>Expectations</i> .....	189
6.9.4.	<i>Procedure</i> .....	190
6.9.5.	<i>Results</i> .....	192
6.9.6.	<i>Discussion and conclusions</i> .....	201
<b>7.</b>	<b>DISCUSSION AND CONCLUSIONS</b> .....	<b>209</b>
<b>8.</b>	<b>FURTHER RESEARCH</b> .....	<b>219</b>
<b>9.</b>	<b>APPENDIX</b> .....	<b>221</b>
9.1.	DESCRIPTIVE ITEM/SCALE STATISTICS OF THE STABILITY DATA COLLECTION .....	221
9.2.	MEAN SCORE RESULTS OF THE FIRST DATA COLLECTION IN 2005 WITHIN THE COMENIUS 2.1-PROJECT .....	222
9.3.	STSQ MEAN SCORE RESULTS OF DIFFERENT SAMPLE TYPES .....	225
9.4.	STSQ INTER-SCORE CORRELATIONS OF DIFFERENT SAMPLE TYPES.....	226
9.5.	DIFFICULTY INDEX OF THE ITEMS OF PART II ACCORDING TO GUTTMAN (1950).....	228

---

9.6.	CRONBACH'S ALPHA OF THE STSQ SCORES SEPARATED BY HIGH SCHOOL YEARS.....	231
9.7.	MEAN SCORE RESULTS OF THE SECOND DATA COLLECTION IN 2006 WITHIN THE COMENIUS 2.1-PROJECT .....	232
9.8.	INITIAL AND LAST (ADAPTED) VERSION OF THE STSQ .....	233
9.8.1.	<i>Initial version of the STSQ</i> .....	233
9.8.2.	<i>Last (adapted) version of the STSQ</i> .....	241
9.9.	PRE-STUDY 1.....	249
9.10.	PRE-STUDY 2.....	253
9.11.	SUMMARY OF ITEM CHANGES AND ADAPTATIONS DURING THE DEVELOPMENT AND EMPIRICAL TESTING OF THE STSQ.....	267
<b>10.</b>	<b>REFERENCE LIST.....</b>	<b>271</b>

## List of tables

Table 1:	Pedagogical facets for the development of a STS.....	60
Table 2:	Empirical steps of developing the STSQ successively .....	86
Table 3:	Image used in part I of the initial version of the STSQ .....	100
Table 4:	Items linked to situation 1, part II of the STSQ .....	106
Table 5:	Items linked to situation 2, part II of the STSQ .....	107
Table 6:	Items linked to situation 3, part II of the STSQ .....	108
Table 7:	Items linked to situation 4, part II of the STSQ .....	109
Table 8:	Items linked to situation 5, part II of the STSQ .....	110
Table 9:	Initial item pool, STSQ, part III (EC: Egocentrism, O: Openness, NFS: Need for security, LOC: Loss of control) .....	111
Table 10:	Pre-study .....	115
Table 11:	Key-terms for “Google” picture search.....	118
Table 12:	Demographic data (pilot study).....	123
Table 13:	Results of scale analysis (internal consistency), part I, STSQ v. 1, pilot study.....	127
Table 14:	Correlations (Spearman’s rho) between closeness, dissimilarity and sympathy scale, N=21, German sample, good picture qualities .....	128
Table 15:	Spearman correlation results of situation 1 part II of the STSQ (N=86, German sample).....	129
Table 16:	Spearman correlation results of situation 2 part II of the STSQ (N=86, German sample).....	129
Table 17:	Spearman correlation results of situation 3 part II of the STSQ (N=86, German sample).....	130
Table 18:	Spearman correlation results of situation 4 part II of the STSQ (N=86, German sample).....	130
Table 19:	Spearman correlation results of situation 5 part II of the STSQ (N=86, German sample).....	131
Table 20:	Results of correlations (Spearman’s rho) of RA1, RA2 and RA3 scores between each situation, part II of the unedited/initial version of the STSQ, N=86 (German sample).....	132
Table 21:	Factor matrix of the RA sum scores between the situations, N=86 (German sample).....	133
Table 22:	Resulting factor loadings for the item pool analysis of part III, STSQ v.1 .....	135
Table 23:	Item-Total Statistics of factor 1 ( $\alpha_{\text{Cronbach}} = 0,62$ ), N=86, German sample .....	136

Table 24:	Item-Total statistics of factor 2 ( $\alpha_{\text{Cronbach}} = 0,57$ ), German sample, $N=86_{\text{Total}}$ .....	137
Table 25:	STSQ mean scores separated by the two samples: German sample (GER) $N=86$ and international sample (Int.) $N=24$ .....	138
Table 26:	Revision of items according to the resulting knowledge of this pilot study (STSQ v.2, part III).....	142
Table 27:	Sample characteristics .....	145
Table 28:	Item pool of part II applied in both measurements.....	146
Table 29:	Item pool of part III applied in both measurements .....	147
Table 30:	Stability results of the first measurement and second measurement (Spearman's rho) .....	151
Table 31:	Scale analysis part I, STSQ v.2 .....	152
Table 32:	Results of factor analysis, part I of the STSQ .....	154
Table 33:	Results of factor analysis part III (sample of the first measurement of the stability study, $N=69$ ).....	158
Table 34:	Applied item pool between different data collections. Revisions are illustrated with an item example, STSQ, part I (EM).....	161
Table 35:	The table shows revisions of items of part II of the STSQ according to the result of the parallel conducted intervention study. The revisions are only applied to the second measurement of this stability study.....	162
Table 36:	Revision of items according to the resulting knowledge of this stability study, STSQ v.2, part III.....	164
Table 37:	Demographic frequencies of the high school sample .....	167
Table 38:	Expectations of (extreme) group analysis.....	168
Table 39:	Significant and marginal significant mean score results of the comparisons between (a) first and last high school year, (b) abroad experiences (yes or no), and (c) immigration background (yes or no).....	170
Table 40:	Significant STSQ score inter correlations separated according to high school levels .....	171
Table 41:	Item revisions STSQ, part I .....	173
Table 42:	Further item adaptations, part II of the STSQ .....	174
Table 43:	Further item adaptations, part III of the STSQ.....	175
Table 44:	Sample structure and size of the intervention design in Germany 2005 and in The Czech Republic 2006.....	181
Table 45:	Applied item pool of the data material from the EU-project in 2005 (STSQ, part I) .....	181
Table 46:	Item pool of part II; the item order is changed according to difficulty index (cf. appendix 9.5).....	183

Table 47:	Applied item pool of part III .....	184
Table 48:	Item sub-scores (O, NFS, and LOC score) .....	184
Table 49:	Item scores based on the results of the last study.....	185
Table 50:	Items of an additional questionnaire .....	185
Table 51:	RA1 score items (nested within and between situations).....	188
Table 52:	RA2 score items (nested within and between situations).....	188
Table 53:	O score items.....	189
Table 54:	LOC score items.....	189
Table 55:	NFS score items .....	189
Table 56:	Statistical methods used for the last analyses .....	191
Table 57:	Significant and marginal significant group differences of the 2005 sample .....	192
Table 58:	Correlation results (Spearman's rho) of the eSTS response pattern scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=11) .....	194
Table 59:	Correlation results (Spearman's rho) of the RA1 mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=10) .....	194
Table 60:	Correlation results (Spearman's rho) of the RA2 mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=33; m2: post measurement, N=30; m3: follow up measurement, N=11) .....	195
Table 61:	Correlation results (Spearman's rho) of the O mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=27; m3: follow up measurement, N=10) .....	195
Table 62:	Correlation results (Spearman's rho) of the NFS mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=11) .....	195
Table 63:	Correlation results (Spearman's rho) of the LOC mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=11) .....	196
Table 64:	Correlation results (Spearman's rho) of the O+LOC mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=30; m2: post measurement, N=25; m3: follow up measurement, N=10) .....	196
Table 65:	Correlations between STSQ scores.....	197

Table 66:	Significant and marginal significant group differences of the 2006 sample.....	200
Table 67:	Suggested eSTS pattern score.....	215
Table 68:	Suggestions for further development of the STSQ.....	217
Table 69:	Descriptive item statistics of part I, STSQ, sample of measurement 1, N=69.....	221
Table 70:	The mean score results (Pre, post, follow up and control group [CG ] measurement) .....	222
Table 71:	The German mean score results (Pre, post, follow up and control group [CG ] measurement) .....	222
Table 72:	The French mean score results (Pre, post, follow up and control group [CG ] measurement) .....	223
Table 73:	The Czech Republic mean score results (Pre, post, follow up and control group [CG ] measurement).....	223
Table 74:	The Polish mean score results (Pre, post, follow up and control group [CG ] measurement) .....	224
Table 75:	STSQ scores of different sample types.....	225
Table 76:	STSQ inter score correlation (Spearman's rho), Sport Gymnasium (high school students), N=224 .....	226
Table 77:	STSQ inter score correlation (Spearman's rho), EU project 2006 (intervention sample, post measurement), N=21 .....	226
Table 78:	STSQ inter score correlation (Spearman's rho), master students of social work, N=10 .....	226
Table 79:	STSQ inter score correlation (Spearman's rho), physic students, N=41 .....	226
Table 80:	STSQ inter score correlation (Spearman's rho), master students of multicultural and international studies (international sample), N=10 .....	227
Table 81:	STSQ inter score correlation (Spearman's rho), master students of information technology, N=22 .....	227
Table 82:	STSQ inter score correlation (Spearman's rho), international summer school Oslo, N=24 .....	227
Table 83:	Cronbach's alpha of the STSQ scores separated by high school years.....	231
Table 84:	Mean score results for the whole sample and country specified .....	232



## List of figures

Figure 1:	Theoretical framework of Intercultural Movement Education from the perspective of this thesis.....	5
Figure 2:	Sequence model of developing the theory-based questionnaire (STSQ) .....	11
Figure 3:	Framework model of intercultural learning according to Gieß-Stüber (2000b:12).....	30
Figure 4:	Triad of skill-related identity (according to Haußer, 1997:128) .....	36
Figure 5:	Sensitivity towards strangeness (STS) something between xenophobia and exoticism? .....	41
Figure 6:	Difference vs. strangeness.....	43
Figure 7:	Norms (e.g. type of greeting) may vary across cultures (Passer and Smith 2004:602, fig. 15.10).....	44
Figure 8:	Dimensions of strangeness .....	47
Figure 9:	A “mat-race” requires co-operation and constructive considerations of differences (in height, strength etc.) in order to successfully cope with the task .....	64
Figure 10:	A Judo specific greeting ritual creates a polite and concentrated learning atmosphere .....	66
Figure 11:	To throw each other requires responsibility towards the partner .....	67
Figure 12:	“The task orientation and fun made it possible to responsibly “fight” together despite of differences .....	67
Figure 13:	A clever implementation of one’s own body-weight can become an advantage in Judo .....	68
Figure 14:	Climbing (co-operation, trust and responsibility) .....	70
Figure 15:	Free fall from a Vaulting Box (trust and coping).....	70
Figure 16:	Taking a risk - jumping over a “glacier rift” .....	70
Figure 17:	Take a risk and jump!.....	71
Figure 18:	Operational model of the STS (STS: Sensitivity towards Strangeness, EM: Emotional Meaning, CM: Cognitive/rational meaning, C: Closeness, D: Difference/Dissimilarity, S: Sympathy, RA I-III: Rational Attribution, EC: Ego centrism, O: Openness, NFS: Need for Security, LOC: Loss of control .....	89
Figure 19:	Item example of the initial version of the STSQ, part I .....	96
Figure 20:	eSTS pattern score = .....	97
Figure 21:	“xenophobia” pattern score = .....	98
Figure 22:	“Pity” pattern score = (Dissimilarity = 4 OR Dissimilarity = 5) & (Sympathy = 5).....	98

Figure 23:	“Exoticism” pattern score =.....	99
Figure 24:	Item example of part II of the initial version of the STSQ.....	104
Figure 25:	Item example of the third part of the STSQ (initial version).....	111
Figure 26:	Task 1 of the writing tasks in the pre-studies .....	117
Figure 27:	Scree plot for the item pool analysis of part III, STSQ v.1 .....	134
Figure 28:	Changes and adaptations of the items construction of part I of the STSQ. On the left side, an item of the STSQ version 1; on the right side, the item after revisions according to the results of this pilot study. ....	140
Figure 29:	Repeated measurement design.....	144
Figure 30:	Calculation of the STS response pattern score .....	149
Figure 31:	Calculation of the xenophobia response pattern score .....	149
Figure 32:	Calculation of the pity response pattern score.....	150
Figure 33:	Calculation of the exoticism response pattern score.....	150
Figure 34:	Scree plot part I of the STSQ (N=69).....	153
Figure 35:	Frequencies of STS response pattern score of item I.1 – I.11 .....	155
Figure 36:	Frequencies of response pattern analysis in comparison to alternative response pattern .....	157
Figure 37:	Histogram familiarity and sympathy scale if item I.16* (STSQ, part I) .....	171
Figure 38:	Measurement design of the STSQ data collection within the EU Comenius project 2.1 .....	180
Figure 39:	STS response pattern score.....	186
Figure 40:	Item example part II.....	187
Figure 41:	eSTS mean score differences of the intervention group in comparison to control group mean score (2005 sample).....	193
Figure 42:	Bar chart of STS mean scores whole 2006 sample and country specified.....	198
Figure 43:	STSQ scores of all measurements of the 2006 sample (overall and country specified).....	199
Figure 44:	More ambiguous item in comparison to the initial item I.5 .....	214
Figure 45:	Item order of situation 1 according to the item’s difficulty.....	228
Figure 46:	Item order of situation 2 according to the item’s difficulty.....	229
Figure 47:	Item order of situation 3 according to the item’s difficulty.....	229
Figure 48:	Item order of situation 4 according to the item’s difficulty.....	230
Figure 49:	Item order of situation 5 according to the item’s difficulty.....	230
Figure 50:	According to Andersen (Andersen & Bali, 2000; Andersen & Bali, 2002) .....	251

**List of abbreviations**

IME:	Intercultural Movement Education
STS:	Sensitivity towards strangeness
STSQ:	Sensitivity towards strangeness questionnaire
eSTS:	Emotional sensitivity towards strangeness
Part I:	First main part of the STSQ (EM, STS response pattern)
Part II:	Second main part of the STSQ (CM, attribution style)
Part III:	Third main part of the STSQ (CM, fundamental understandings and attitudes regarding the STS)
EM:	Emotional Meaning
CM:	Cognitive meaning
C:	Closeness
D:	Difference/Dissimilarity
S:	Sympathy
RA 1-3:	Rational level of attribution
EC:	Ego centrism
O:	Openness
NFS:	Need for Security
LOC:	Loss of control
EFA:	Explorative factor analysis
CFA:	Confirmative factor analysis
PAF:	Principal axis factoring
SEM:	Structural equation modeling
ISS:	International summer school
IG:	Intervention group
CG:	Control group
m:	Measurement
DMIS:	Developmental model of intercultural sensitivity
CQ:	Cultural Intelligence
MPQ:	The Multicultural Personality Questionnaire



---

## Abstract

Strangeness is a central theoretical concept within a larger research project of “Intercultural Movement Education” (Erdmann 1999). The aim of this thesis was the development of a concept-based device in order to measure sensitivity towards strangeness (STS). This aim required among others that strangeness needed to be conceptualized with respect to its intended operationalization and on the basis of Intercultural Movement Education. In order to meet central ideas of Intercultural Movement Education, sensitivity towards strangeness was first conceptualized as a multi-faceted construct. These theoretical considerations were the first step to construct the “sensitivity towards strangeness questionnaire” (STSQ). The STSQ was supposed to measure a band-width of relevant facets of strangeness. Each facet is represented by a number of items. The developed items of each facet are not parallel items which are supposed to measure exactly the same factor. The items are more understood as complementary whereby each item is supposed to measure a different aspect of one facet (e.g. emotional STS, awareness over different attributions towards strangers, openness towards strangers). The facets are therefore kind of categories structuring a more heterogeneous pool of items. The deductively constructed item pool was developed further with help of smaller empirical studies which helped to improve single items and develop score-keys of the STSQ.

The STSQ was designed for research purposes only. The STSQ was developed in the first place deductively. This concept-based construction of the STSQ allowed making predictions of empirical results from the theory which required lower scale qualities and did not oblige general standardization or norms of the measuring instrument (cf. Erdmann, 1988). The simplest way the STSQ is supposed to differentiate is to check if a criteria (item) is achieved or not (nominal scale level). In this sense, applications of the STSQ are supposed to enable for differentiations between defined populations and can be used as a screening tool on the baseline of the underlying theoretical concept.

However, the empirical investigations of the STSQ provided hints on validity and reliability of the STSQ but more systematical and advanced validity and reliability analyses of the whole instrument are required in the future.



## **PART I: THEORETICAL FOUNDATIONS**





## 1. Introduction

My PhD-project was originally part of a broader research project called “Intercultural Movement Education” (cf. Erdmann, 1999b). Intercultural Movement Education (IME) is a scientific approach which tried to deal with problems related to western multicultural societies from the perspective of sport pedagogy and didactics. The initial impetus for this program stemmed from racial and ethnic conflicts that had taken place in Germany in the late nineties (Erdmann, 1999b). IME is aiming at a better mutual understanding and a more rational management of conflicts within multicultural societies. Based on theoretical groundwork, field studies and critical reflections within the field of sport, IME is further aiming at the development of theoretically founded guidelines for practice in the field of sport and physical education in order to meet problems and conflicts related to a multicultural society.

The specific approach of IME required measuring instruments in order to measure outcomes of the project, especially the efficiency of IME and its applications. The development of a concept-based device which measures central facets of the underlying theoretical concept of IME became the goal of my thesis. My theoretical considerations around IME led me to a conceptualization of “sensitivity towards strangeness” (STS). Based on this concept I derived an instrument which was supposed to measure STS.

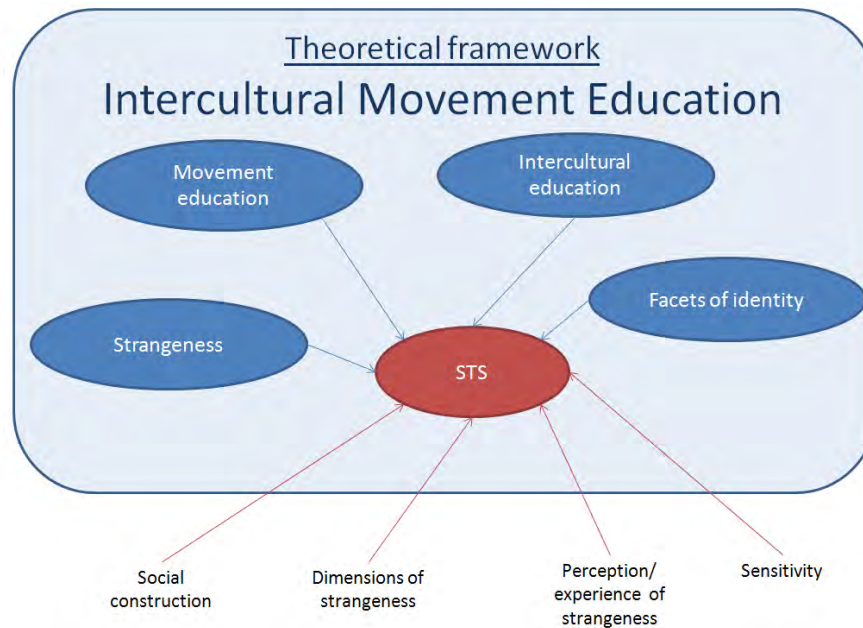
Strangeness is a central category within social science. It is understood as a social construction in the sense of constructing a borderline between strangeness and familiarity; natives and immigrants; insiders and outsiders (Hahn, 2000; Simmel, 1992). Strangeness is related to problems such as xenophobia, discrimination and hate against strangers because it is related to power struggles and the “stranger” is often perceived as a threat to one’s own identity (Faulenbach, Hesse, & Klaeren, 2001; Rommelspacher 1998). Facets of identity theory and conceptions of strangeness are two central theoretical perspectives within IME (cf. Erdmann, 1999a; Gieß-Stüber 1999).

The term “stranger” is usually associated with a foreigner or a person abroad. But strangeness is a more general concept. Differences between men and women, punk and citizen, disabled and “non-disabled” or even between separate fractions in a physical education class can also lead to strangeness (Gieß-Stüber 1999; Gieß-Stüber 2005). A common starting point for IME was therefore the difference – similar to other concepts of intercultural education (cf. Auernheimer, 2003; Nieke, 2000; Holzbrecher, 1997). The construction of difference enables

us to distinguish between our own person or social affiliation and other persons or social groups. Such categorizations help to structure the world around us which is needed to reduce the complexity of our world in way that we are able to deal with it. In this sense the construction of difference is a structuring process as can be illustrated in a simplified way with a differentiation between apples and pears. If one has eaten both fruits one is familiar with the differences between them. Differentiations, demarcations or categorizations are needed because they make the complexities of our world around manageable. In this sense, categorizations function as orientation. But the construction of difference can also lead to exclusion as the distinction between insiders and outsiders illustrate. In addition, the construction of difference and the ways people deal with difference and strangeness is influenced by emotions, underlying valuations and is defined by power relations.

Prerequisite for experiences of strangeness is a perception or attribution of difference. But when the difference is difficult to understand or a difference appears unexpected, differences are realized as unusual, weird or strange. The difficulty to understand or to predict an outcome of situation, a behavior of the stranger or to predict what intentions of the stranger might have are indicators for the implicit uncertainty in the context of strangeness. Uncertainty is also related to identity constructions. The ways people deal with uncertainty is learned and depends to a large extend on the stability of our own identity. Instabilities during identity development can provoke that the stranger and the implicit uncertainty is perceived as a threat to one's own identity. Perceiving strangeness as threatening can provoke rigid defense mechanism such as denying difference and strangeness. As will be explained more comprehensively later in this thesis, a balanced relationship between perceived cohesion of the own identity (something like stability as will be explained later) and the openness for uncertainty is understood as an optimal prerequisite for dealing constructively with uncertainty and strangeness. But as indicated, the implicit unpredictability of strangeness in relation to our identity construction implies the difficulty to deal with difference and strangeness constructively and can lead to conflicts between natives and strangers.

Conceptions of strangeness can be found in different fields of social sciences. Yet, a comprehensive theory of strangeness seems still to be missing. I tried to focus my theoretical considerations on more general structures and mechanism of strangeness because I intend to develop a measuring instrument which measures a band-width of relevant facets of strangeness in line with the underlying theoretical framework of IME. The following figure (1) illustrates the theoretical concepts basically underlying this thesis:



**Figure 1: Theoretical framework of Intercultural Movement Education from the perspective of this thesis**

A survey of literature in the field of sport showed disappointing results with respect to improving an intercultural competence (Erdmann, 1999b). In particular, there were a lack of theory-driven research programs, intervention projects, and longitudinal studies in the field of sports and physical education (ibid, 1999b). Numerous initiatives in the field of sports were conceived mainly without a theoretical conception (Erdmann, 1999b). An inquiry<sup>1</sup> among 104 German sport federations supported the perception that most of the sport organizations lack awareness of problems linked to a multicultural society (Michels and Schulz, 1999; Sonnenschein, 1999). Michels and Schulz (ibid. 1999) showed further that most of the sport organizations' understandings were rather based on the assimilation than the integration perspective (Sonnenschein, 1999; Erdmann, 1999b)<sup>2</sup>. This perception led to the initiation of the project "IME" at the German Sport University, which intended to fill the aforementioned gap. It was assumed that mere engagement is not sufficient to result in any substantial change (Erdmann, 1999b). The essential cause for the noticed setbacks was perceived to be the lack of a conceptual basis. Sport organizations, however, are left to their own devices to solve this problem,

<sup>1</sup> which were conducted in the realm of the mother-project "intercultural movement education"

<sup>2</sup> Walseth and Fasting indicated a similar perception in Norway and a number of European countries (Walseth 2004; Walseth and Fasting 2004).

with the clear absence of the scientific community (Michels and Schulz, 1999; Noethlichs, 2001). As a result, “IME” focused on efforts to develop a theoretical concept offering perspectives and guidelines for practical applications (Erdmann, 1999a; Gieß-Stüber, 1999; Gieß-Stüber, 2003a; Gieß-Stüber, 2008; Noethlichs, 2005b; Noethlichs, 2003; Sonnenschein, 1999).

In this context, the term *intercultural* is used on purpose in distinction to *multicultural*. In pedagogical efforts intercultural is preferred in order to point out that contact and meaningful interaction is required for an initiation of intercultural learning between different cultural groups which then may lead to new (cultural) constellations (cf. Pihl, 2000, Erdmann 1999b; Auernheimer 1996; Auernheimer, 2003). The term *multicultural* refers more to different cultural groups living side by side. The main focus is directed to initiate and support learning between different persons and cultural groups. Intercultural education is referring to intercultural learning within multicultural societies (cf. Nieke, 2000).

Culture<sup>3</sup> in this context is understood differently from conservative concepts which follow premises of a supposed homogeneous and separatist understanding of culture (cf. Blecking, 2008). This traditional understanding of culture does not reflect the complexity of modern cultures. Modern cultures are characterized by pluralistic identities and the borderlines between cultures are diffuse and complex. However, it seems that most of western societies still stick to the traditional concept of culture (Welsch, 1999). The idea of transculturality promotes that we have to focus beyond demarcations between our own and the strange culture (ibid., 1999; cf. also Nieke, 2000). This idea appears reasonable in the sense of avoiding exclusion but it implicates a dilemma because we need to construct differences as natural process of identity construction on the one hand and on the other hand demarcation can lead to conflicts and exclusion. It seems obvious that we need to create differences, but it is necessary to reflect critically on how we chose to deal with difference and strangeness as well. We often overestimate our own, familiar cultural goods, norms and standards to others. Such self-centered perspectives lead to problematical ways of dealing with difference and strangeness and maintain the aforementioned traditional or static understanding of culture.

---

<sup>3</sup> The term culture is also referring to sub-cultural systems within social communities. Culture is understood as a collection of human products, expressions, and values that are supposed to be meaningful to a social group. Furthermore, culture has to be seen as a dynamic dimension which means that culture is continuously changing and developing. Additionally, it seems reasonable to differentiate between “objective” and “subjective” culture where the objective part summarizes all cultural products, and “subjective” culture means how people deal with their cultural goods. Culture is characterized by similarity, symbolic meaning, and it is always referring to a specific time and space context. The main function of culture is orientation (Auernheimer, 1995).

---

The point is that culture, in the context of western industrialized countries, is a dynamic dimension which changes continuously. But on the other hand, common traditions, symbols, goods, habits etc. makes culture definable and creates more or less coherent pictures of culture. Both aspects define the dynamic understanding of culture which creates identity, membership and affiliation and allows change or development.

Intercultural learning focuses on learning how to deal with differences in constructive ways (Auerheimer, 2003; Grimminger, 2009). Respect and acceptance/recognition for otherness are required attitudes for an intercultural understanding and dialogue (Auerheimer, 2003). The terms respect and acceptance are preferred terms instead to *tolerance* because tolerance implicates irrelevance and a power difference (cf. Habermas, 2002). The dominating population is *tolerant* towards immigrants as long they are not perceived as a threat to dominating power structures. On the other hand, if an immigrant would be tolerant with the dominating population, such an attitude would probably be interpreted as arrogance (Auerheimer, 2003). Therefore, the terms acceptance, recognition and respect are considered as more appropriate terms in the realm of intercultural learning.

IME elaborates critically possibilities and limitations of an intercultural orientated education through and with the medium of sport and “movement”. The term “movement” is used on purpose within IME. The idea is to demarcate Movement Education from competitive associated and purely discipline orientated understanding of sport. Competition means a comparison by commonly accepted rules which all participants need to follow. Competition often implies rivalry and the idea of “we against the others”. The dominating goal is to win or beat the opponent and competition consequently focuses on the result (cf. Erdmann, 2008). Competition in this sense appears to be problematical for integration purposes and do not allow much space for intercultural learning. In addition, local sport often reflects the norms and rules of the respective dominating part of the society (cf. Gieß-Stüber, 2000a). This may cause similar difficulties and integration barriers of minorities into sport clubs as can be seen in other social constellation (cf. Walseth, 2004).

However, anthropological considerations of the body and movement suggest investigating the educational potentials of movement and sport (Grupe & Krüger, 2002; Grupe & Krüger, 1996; Meinberg, 1996). The body is understood as a mediator between the person and the world around (Merleau-Ponty, 1996). Grupe & Krüger (2002) summarize the meaning of the body and movement for learning and development as the following (cf. *ibid*, 2002:208-209):

1. The instrumental meaning:
  - a. we can achieve, produce, express with our body and movement, and
  - b. we can experience, try to, and change with and through our body and movement
2. The perceptive-experiencing meaning:
  - a. through our movements we *explore* and *experience* our own body, the material world around, the nature, and other people
3. The social meaning
  - a. we create relationships to other people through *interaction* and *communication*
  - b. we can express our emotions through movements
  - c. we can express ourselves ritually through our movements (socially defined meanings of movements)
4. Personal meaning
  - a. we can experience ourselves with and through movements
  - b. Our body and movements are mirror and projection of our personality

These dimensions are closely related with each other and they shall point out the pedagogical potential of body and movement. They indicate the baseline for possible potentials of body and movement as a kind of mediator for intercultural learning purposes. The meaning of sport, movement and body is also considered as importation tool for a mediation of STS. I will point out this meaning more specifically in chapter 5.

My thesis is structured in two main parts. Part I deals with theoretical foundations and conceptualization of the topic. The second part (II) represents the methodological section presenting the procedure of developing the STSQ. The sequence model shown in figure 2 illustrates the steps of development and the basic structure of this thesis.

In chapter 2, I start with a literature review on related concepts and approaches. In the review I try to focus on three different aspects: a) more general theoretical conceptions of

---

strangeness, b) strangeness in the context of sport and movement education, c) related concepts of models and measuring instruments.

Then, the three theoretical perspectives of IME will be outlined in chapter 3. Cultural difference as a central category of intercultural education, strangeness as a social construction and facets of identity are these three theoretical domains which are related with each other under the “umbrella” of IME” (cf. Figure 1). I try to view and extend these theoretical ideas according to the goal of my thesis.

Having presented the theoretical outlines of IME, I will continue with the conceptualization of sensitivity towards strangeness in chapter 4. I start my considerations with specifying the connotation of strangeness. Then, I will focus on three different facets of strangeness which were considered important theoretical aspects regarding an operationalization. Before finally presenting the concept of STS, I will focus on the process of perceiving difference and strangeness from a classical social psychological point of view.

In chapter 5 I will introduce some practical implications of my project because these ideas represent the field for applications of STS in the future. The practical implications within this chapter are related to IME but also based on my own theoretical considerations and personal experiences in teaching pupils and physical education students. The described examples in this chapter shall illustrate the pedagogical potential of sport and physical education for improving sensitivity towards strangeness.

The next main part (II) starts with chapter 6 and deals with the methodological procedure of developing the sensitivity towards strangeness questionnaire (STSQ). I will start this chapter with a more theoretical argumentation for the followed procedure of developing a theory-based measuring instrument with a number of smaller field studies. I will also show how the methodological idea of the framework program of IME was applied to the goal of my project.

Out of the theoretical considerations presented in part I of this thesis, I generated an operational model shown in chapter 6.3. This model represents the link between part I (theoretical ideas) and the developed (observable) indicators of the STSQ. The operational model is the theoretical guideline for the intended item structure of the STSQ and indicates the assumed relations between the different facets of STS.

The first structured version of the STSQ is presented in the methodological part (chapter 6.4) because it is more or less derived from the theoretical groundwork of part I of this thesis. The operational model represents these theoretical perspectives which are assumed to be rele-

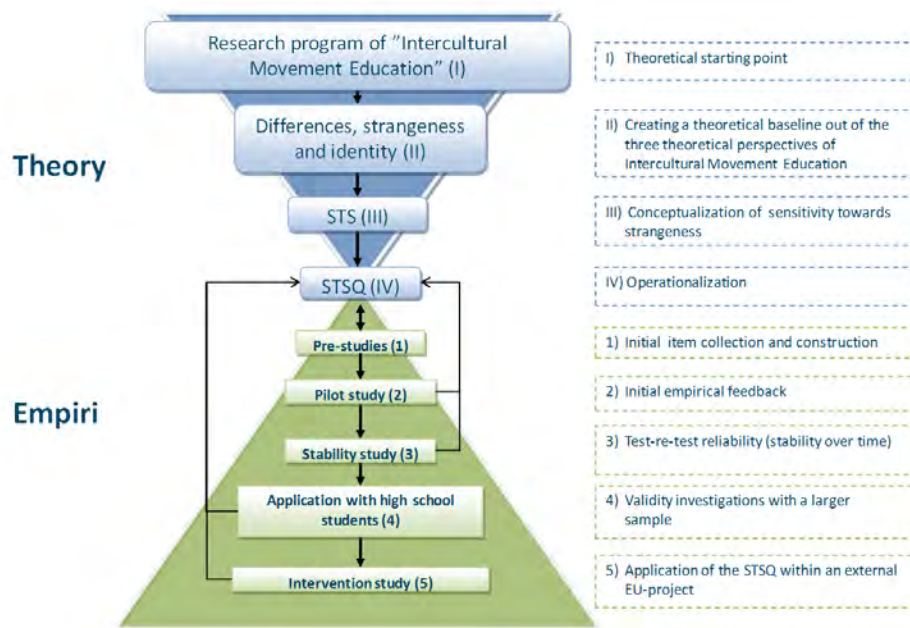
vant for the operationalization. The pre-studies in chapter 6.5 helped to construct and formulate the items.

The following sub-chapters (chapter 6.6 – 6.9) represent the empirical steps of developing the items of the initial version of the STSQ further and develop the score keys. The attachments of this thesis document the adaptations of single items and score keys in line with the conducted studies. The attachments also include the complete initial and final version of the STSQ and a number of statistical references. The attachment also includes some marginal results which were not discussed in detail. But some of these results such as correlations of the STSQ scores with different samples or calculated difficulty indexes of the items of part II of the STSQ may function as an initial reference for potential applications of the STSQ.

In chapter 7 I will discuss the results of my entire project. I finish with some conclusions and suggestions for further research in chapter 8.

The theory-based procedure of developing the STSQ is illustrated in Figure 2. The blue triangle represent the theoretical work and the green one the empirical stages of development whereby the STSQ placed in-between because it represents the transition between theory and empirical representation. The blue triangle narrows to the bottom. This illustrates that I started with quite complex theoretical ideas that was gradually reduced and my first structured version of the STSQ is the result of this reduction process and the connection to the empirical part. Then, the empirical investigations led to more and more empirical knowledge about the STSQ illustrated in Figure 2 inverted green triangle. The arrows illustrate that the increase of knowledge during the empirical investigations led to adaptations of the STSQ items.





**Figure 2:** Sequence model of developing the theory-based questionnaire (STSQ)

The need of measuring devices within "IME" (Erdmann 1999b) promoted the aim of this thesis. My participation within this larger research project required on the one hand that the aims, contents and methodological procedures within this thesis were based on the same fundamental premises of the mother project. This was needed in order to keep the link between my project and broader framework concept. The way of dealing with complex research projects through a number of smaller and theoretically linked research projects should make it easier to handle the complexity of the larger project. On the other hand, this thesis is to be understood as the result of an autonomously conducted research process. Strangeness is a general construct within social sciences and the results of this project can possibly be applied in other contexts as well.



---

## 2. Review of related literature

Intercultural Movement Education was the starting point for the development of my measuring instrument and will therefore be presented in a separate chapter (3). A look through the literature reveals that it is impossible to deal with the amount of scientific contributions around strangeness and related topics in the realm of this thesis. Therefore, my literature review focuses on selected literature with the purpose to indicate and discuss the relation of other concepts and projects to STS, point out the relevance of STS and demarcate my project clearer from others. A more comprehensive discussion of related literature could easily become a thesis of its own.

### 2.1. Conceptions of strangeness

I will mention two comprehensive works though. Julia Reuter dealt theoretically and more comprehensively with “otherness<sup>4</sup>” and “strangeness”. She provided an overview about classical theoretical conceptions of otherness and strangeness within the field of sociology (cf. Reuter, 2002). Her main focus was to analyze what the sociology of strangeness tells us about our own person or social group. In this sense strangeness is viewed as a kind of mirror of our own identity.

Another comprehensive theoretical work can be found in Yoshiro Nakamura’s (2000) thesis “Xenosophie. Bausteine für eine Theorie der Fremdheit”. His aim was to supply elements towards the development of a theory of strangeness. Xenography, xenosophy and xenology are his central concepts theorizing strangeness in line with classical phenomenological approaches according to Husserl and Waldenfels<sup>5</sup>.

A classical starting point of dealing analytical with the phenomenon of strangeness can be found in Georg Simmel’s (1992<sup>6</sup>) “Exkurs über den Fremden”, Margret Wood (1934) “The stranger. A study in social relationships”, and Alfred Schütz (1974) “Grundzüge einer Theorie des Fremdverstehens”. Strangeness is viewed as a specific social relationship. The different

---

<sup>4</sup> In my conceptualization of STS I prefer using the term “difference” instead of “otherness”. This should point out that strangeness refers to differences which can appear on different levels: individual, social and cultural level. Otherness refers more to social differences. But both terms are understood as quite compatible.

<sup>5</sup> cf. Waldenfels, B. (1997). *Topographie des Fremden. Studien zur Phänomenologie des Fremden 1*. Frankfurt am Main: Suhrkamp.

cf. Waldenfels, B. (2006). *Grundmotive einer Phänomenologie des Fremden*. Frankfurt am Main: Suhrkamp.

<sup>6</sup> First printed in 1908

concepts of “the stranger” refer to the wanderer who comes today and will not leave tomorrow according to Simmel’s essay about the stranger. **Georg Simmel** points out two crucial aspects about “the stranger”: the first aspect refers to the assumptions that the stranger is only a guest. It is therefore the stranger is expected to leave without substantial consequences or changes for the majority. However, the stranger is often a guest who comes today but stays tomorrow – in contradiction to the expectations of the majority<sup>7</sup>. In such cases, the dominating society is not prepared for integration which can lead to problems and conflicts. Simmel’s dimension of social closeness and distance implies two perspectives. Firstly, he refers to a more superficial geographical demarcation between foreign and distant areas and close and more familiar (geographical) areas. Secondly, Simmel’s dimension refers to strangeness as a social relationship to the “other” person or social group. Operating with Simmel’s terminologies in a figurative way, the “stranger” can be described as a “distant” (strange) person who became “close” in the sense of a meaningful relationship. I will follow up this idea within my conceptualization of sensitivity towards strangeness in chapter (4).

**Margret Wood** introduces the concept of the stranger as a person who enters a group for the first time, outside the system of relationships which unite the group. If the stranger will be included it is required to extend the existing relationship of the respective group. Inclusion of the stranger depends on the flexibility of the system’s relationships, the personal qualities of the interacting individuals, and the presence or absence of extraneous factors which might tend to hasten or retard the process of inclusion (Wood, 1934). Wood’s focus is the initial (first face-to-face meeting) formation of a new relationship between the native and stranger under the prerequisite that the stranger is a person who remains within the new group.

**Alfred Schütz** (1974) views the stranger from a hermeneutic perspective, as a problem of understanding. He differentiates between the own or the other person or social group and each of them represents different meaning systems, knowledge and understandings. According to Schütz, the stranger or immigrant clashes with the knowledge and meaning systems of a cultural community because the stranger represents different understanding and knowledge. Schütz points out a kind of dilemma. We need our experience and background knowledge to understand the stranger or immigrant. But the native – stranger constellation implies two dif-

---

<sup>7</sup> The term majority and minority members are not referring to quantitative differences but more to power differences between groups within a society whereby the majority represents the dominating part of the respective society.

---

ferent systems which are not necessarily compatible with each other. Consequently, Schütz points out that our understanding is limited. This idea has crucial consequences for pedagogical endeavors. It is important to promote an understanding that sometimes we have to accept and respect difference even though we are not able to understand the difference as for instance the idea of freedom of religion illustrates. But this is not always an easy task because those differences are related to fundamental values and influenced by power struggles as conflicts in the Middle East illustrate.

People deal with strangeness differently, depending on their personal experiences. Xenophobia and exoticism are concepts related to specific ways of dealing with strangeness. Both concepts are understood as defensive mechanisms against the implicit uncertainty of strangeness. **Mario Erdheim** conceptualized both terms from the perspective of psychoanalytic theory (Erdheim, 1988). According to his approach, the basis for developing a meaning of strangeness and the ways of dealing with strangeness through one's lifetime leads back to the attachment towards the mother or caregiver in early childhood. In this context, the stranger is understood as the opposite of mother or "non-mother". This is what Erdheim called representatives of strangeness. Especially xenophobic tendencies are developed through experiences in early childhood. The separation from the mother is viewed as a crucial event in relation to the development of anxieties or xenophobic tendencies. But even though exoticism and xenophobia may appear as opposite dimensions, they are similar in the sense that they are strategies of avoidance. Xenophobia leads to avoid strangeness because it is perceived as threatening. Exoticism attracts one's attention to strange cultures or countries in order to avoid changes "at home" (Erdheim, 1988).

Strangeness represents the limits of understanding. By dealing with strangeness constructively, it can be possible, at least to some extent, to deconstruct strangeness into familiarity. But strangeness is characterized by aforementioned difficulties in understanding. To gain complete understanding might be impossible or an illusion with reference to **Waldensfels** (1997; 2006). He pointed out that we can only follow tracks or hints in understanding of the "other". On the other hand, people may believe to understand the other because they feel a need of constructing familiar and plausible pictures of the world around.

The idea of understanding leads back to the scientific theory of hermeneutics. **Gadamer** (1975) as a representative of hermeneutics points out a further aspect. The knowledge and understanding we gain when dealing with strangeness is just temporary. This aspect leads to an important consequence for intercultural learning. If the understanding of the other is uncer-

tain and temporary, we need to be open for alternative possibilities. In terms of Auernheimer (2003), we have to become aware of the fact that alternative meanings can be hidden behind the already interpreted information of the “other”.

**Schäffter** (1991) describes strangeness as a relationship. He uses, similar to Simmel, the social dimension of closeness and distance. He further views the function of strangeness as a structuring concept (*Ordnungskonzept*). We construct and reproduce differences and patterns to structure the world. Constructed structures and patterns make the world understandable, predicable and manageable. But it is not the characteristics of the other which lead to strangeness; it is the relation to our own identity. Strangeness is therefore a result of a “*selbstvergesene Ordnungsleistung*” (Schäffter, 1991:14). Our ways of structuring the world around us are influenced by personal interests who can easily conflict with other world views and interests. The implicit conflict potential increases to a problematical extension when we proclaim our ways of structuring the world (*Ordnungsleistung*) and our personal interests as universal and absolute (Schäffter, 1991).

The construction of difference shows its relation to the gender construct. Like the concept of strangeness, **gender** is also a social construction. It represents a social relationship, is used as differentiation, it is a structuring category (cf. Voss, 2003), and it is related to identity conceptions. Gieß-Stüber (2000a) points out that strangeness is in the first place not a category of a specific content, but it is a formal category which can be applied into different fields. Gender can to some extent be viewed as such field. Gender refers to (socially constructed) differences between boys and girls or men and women. Gender differences can lead to similar structural and power related issues as described within other contexts such as ethnic differences. *Women in men's domain* illustrate constructed differences between men and women. Empirical studies show that women working in “men’s professions” overcompensate their own efforts and underplay their own success in order to adapt and subordinate to the men’s standards (Gieß-Stüber, 2000a). This stabilizes and amplifies the construction of stereotypes (ibid., 2000a).

## 2.2. Strangeness in the realm of sport and physical education

It is difficult to find conceptualizations of strangeness in the context of sport and in particular within the field of sport pedagogy even though the phenomenon is present in the context of sport.

However, the concept of Bodily Strangeness (*Körperliche Fremdheit*) (Bröskamp, 1994; 2008) is one conceptualization which has its roots in the context of sport. **Bröskamp** deals

specifically with the meaning of our body in relation to the phenomenon of strangeness. The following quote illustrates the theoretical perspective behind his concept:

“Es gibt eine ganz besondere, in den Intelligenztheorien häufig vernachlässigte Art und Weise des Verstehens: Verstehen mittels des eigenen Körpers. Eine Unmenge von Dingen verstehen wir nur mittels unseres Körpers, jenseits des Bewußtseins, ohne über die Wörter zu verfügen, es auszudrücken“ (Pierre Bourdieu quoted in Bröskamp, 1994:V)

Bröskamp views the phenomenon of strangeness and the meaning of the body from the theoretical perspective of Pierre Bourdieu. Based on a number of empirical studies, Bröskamp critically points out that sport does not automatically lead to a better integration of minority groups. Sport often provokes the opposite and sometimes in more extreme ways as in other social contexts (cf. Bröskamp, 1994:10). Bröskamp examines how strangeness can be experienced with the own body in sport. He points out that the phenomenon of strangeness is difficult to grasp because it is related to repressed processes. He indicates further that it is still quite uncertain how we develop individual and collective strategies of dealing with strangeness and how this effects “ethnic group relations” (Bröskamp, 1994:13).

**Walseth and Fasting** (2004) attach importance to a similar critical perspective of sport and its often superficially assigned integration potential. They point out that the organized sport institutions do not seem to be prepared for the diversity within multicultural societies. Existing conceptions under the label of “Sport and Integration” are usually based on the majority’s premises that force minority groups to assimilate to the dominating sport culture. Walseth (2006) showed in her studies of young Muslim women that participation in sport has impact on the individual’s identity work and value adjustment and can contribute in some aspects of integration. But she also showed that the women participating in sport are different from those who do not participate in sport. Women who are active have a closer relationship to the majority population than to those who are not active. In addition, women active in sport challenge the values and norms in their own cultural (minority) group more than women who are not active.

**Gieß-Stüber<sup>8</sup> and Blecking** (2008) provided an anthology in the context of their EU financed project<sup>9</sup>. This project was initiated by the Institute for Sport and Sport Sciences at the University of Freiburg in Germany. Under the guidance of theoretical modules, the interna-

---

<sup>8</sup> Prof. Petra Gieß-Stüber was one of the initiators of Intercultural Movement Education

<sup>9</sup> EU project number: 119019-CP-1-2004-1-DE-COMENIUS-C21

---

tional research group tried to consider critically the difficulties in the contexts of universities, physical education, and organized sport related to migration processes within the increasing European Union. Theoretical perspectives were structured in five modules for teacher education: 1) Sport and Immigration, 2) Sport and Integration, 3) Sport and Regional Movement Cultures, 4) Sport, Ethnicity and Strangeness, and 5) Sport, Ethnicity and Gender. The book represents an actual debate and scientific contributions from different experts in the field of the mentioned modules and basically from the participating countries Germany, France, Poland and The Czech Republic. This project provided some empirical data collected with my developed questionnaire, and I will present some crucial ideas of it later in my thesis.

Grimminger (2009), also a member of the aforementioned EU-project provides a review on empirical studies conducted in the context of intercultural learning and intercultural competence. She also provides an overview on actual debates and scientific positions on intercultural education in the field of sport pedagogy in Germany. Her review on empirical studies illustrate clearly that teachers find it difficult to cope with the challenges related to cultural diversity in schools. Her studies point out that there is lack of empirical studies testing the efficiency of teacher education and on-the-job-training for teachers.

Her PhD thesis tries to fill this gap. She has developed and evaluated a teacher training concept which is aiming at promoting an intercultural competence in sport. Her crucial theoretical assumptions are founded within Intercultural Movement Education (Erdmann, 1999b) and Intercultural Education in and through Sports (Gieß-Stüber, 2005b). Her evaluation design is characterized by a number of smaller designed studies which are theoretically related with each other in order to investigate her research approach successively. Furthermore, her evaluation design includes context-, input-, product-, process-, and transfer- evaluation of her developed didactical conception. This comprehensive evaluation required using qualitative and quantitative methods (multi-method approach). Her studies show that an intercultural competence does not require developing complete new pedagogical competencies of teachers, but more an extension of specific contents.

Grimminger (2009) summarizes intercultural competence of physical education teachers with two general aspects: 1) the ability to deal constructively with cultural diversity and strangeness, and 2) the ability to initiate and critically reflect intercultural learning processes with purpose to promote an intercultural competence of children and juveniles (Grimminger, 2009:148). Her results also point out that dealing with diversity is related to the own identity in particular when intercultural situations are not only perceived as different but also experienced as strange. Her results indicate a relationship between the individual's perceived self-



---

efficacy and the ways people deal with strangeness. Lower self-efficacy leads easier to uncertainty and resignation (cf. also Bender-Szymanski, 2004, 2006). Consequently, teachers with lower self-efficacy showed a lower willingness to consider new and unfamiliar content required for intercultural education (Grimminger, 2009).

### 2.3. Related operational models and measuring instruments

The literature on related measuring instruments helped developing an initial item pool; even though the measuring instruments discovered differ more or less in their theoretical assumptions, they serve, at least, as helpful inspirations for item pool construction. In the following I will more explicitly discuss the relevance of related models and measuring instruments.

“**The Intercultural Development Inventory**” (Hammer, Bennet, & Wiseman, 2003) is based on the conceptualization of the dimensions of intercultural competence (Bennet, 1993). Bennet’s **developmental model of intercultural sensitivity** (DMIS) constitutes a progression in orientations towards cultural differences that include the potential for increasingly more sophisticated intercultural experiences (ibid., 1993:34). The Intercultural Development Inventory (IDI) was constructed by the authors to measure the orientation towards cultural differences in the DMIS (both ethnocentric orientations - denial, defense, minimization - and ethnocentric orientations - acceptance, adaptation, integration). This measuring instrument seemed to return satisfactory results in confirmative factor analysis, reliability analysis, and construct validity tests (cf. Hammer et al., 2003). However, the instrument appeared to be developed for commercial use, was not (economically) possible to obtain the whole questionnaire without participating in a training program.

Even though this inventory seems to represent a relevant item pool with respect to measuring sensitivity towards strangeness, the only item examples available are those included in the literature. These examples, however, are an useful inspiration for the item development of the STSQ. In addition, crucial conceptual differences between STS and DMIS made a direct transfer of some dimensions of the IDI unfeasible. For instance, the DMIS seemed to be developed more for emigration purposes (cf. Paige, 1993); whereas the concept of strangeness is meant to be applicable to migration in general and, consequently, include both the perspectives of the dominant group members, as well as the minority group members. In other words, strangeness includes immigration and emigration perspectives respectively.

STS also refers to more general structures than emigration perspectives or ethnic differences. A further difference was revealed when comparing the DMIS with the STS. The DMIS

claims that people need to go through developmental stages progressively in order to achieve a more sophisticated intercultural experience. The STS, on the other hand, assumes that the principle of “learning through insight” will probably lead to more sophisticated intercultural experiences without necessarily passing through all of Bennet’s stages. In STS, it is assumed that sports and physical education can create opportunities which can provoke fundamental insights through personal experiences with one’s own body as the reference point (Noethlichs, 2005b; Noethlichs, 2003). Finally, the underlying concept of STS refers to a more differentiated understanding of culture than that which is shown in Bennet’s.

The “**Intercultural Sensitivity Scale**” (Chen & Starosta, 2000; Chen & Starosta, 2002; Chen & Starosta, 1997) appeared to be a relevant scale and, consequently, the most influential with respect to the development of the STSQ. The Chen and Starosta’s concept of sensitivity (Chen et al., 2000; Chen et al., 2002) is placed within the field of communication studies. They developed an interesting and highly relevant concept of intercultural sensitivity. A crucial discovery of Chen and Starosta was the apparent consistency between different theoretical approaches in the realm of intercultural communication. Consequently, they tried to synthesize three different theoretical conceptions. They furthermore developed “The Intercultural Sensitivity Scale” based on their attempt to create a more comprehensive and more consistent model which would include Bennet’s conceptualization as well. The resulting model included emotional (Intercultural Sensitivity), cognitive (Intercultural Awareness), and behavioural elements (Intercultural Adroitness). The resulting Intercultural Sensitivity Scale includes a total of six dimensions: self-esteem, self-monitoring, open-mindedness, empathy, interaction involvement, and non-judgment. These dimensions and the resulting indicators were substantially beneficial with respect to the development of the conceptualization of STS and the STSQ as a measuring instrument.

Similar to Bennet’s DMIS, however, it seems that Chen and Starosta’s approach was more directed towards training an intercultural communicative competence and, therefore, more focused on the emigration perspective. It required a component of manager training intercultural competence in order to prepare managers to work abroad.

Nevertheless, both discussed approaches (both the DMIS and The Intercultural Sensitivity Scale) were interesting and helpful for item pool development, even though their main focus was not discovering mechanisms and structures which lead to conflicts between immigrants and natives. Such analyses are needed in order to develop training programs directed towards more rational conflict management (cf. Erdmann, 1999b). These approaches refer exclusively

---

to ethnic differences as opposed to the more general concept of strangeness which refers to differences between sub-cultures in the broadest sense.

The conceptualization of STS refers to an emotional dimension of sensitivity. In order to measure emotional sensitivity towards strangeness, a semi-projective measuring method is required to avoid measuring errors, such as political correctness which is assumed to be a typical error influence within this field.

The **Implicit Association Test** (IAT) developed by Greenwald, McGhee and Schartz (1998) is a projective or indirect measuring method. This device was originally developed for exploring the unconscious roots of thinking and feeling<sup>10</sup>. “The IAT seeks to measure implicit attitudes by measuring their underlying automatic evaluation” (ibid., 1998:1464). In particular the online version of the IAT is designed to measure specific associations, such as “implicit racism”. Differential associations of two target concepts with an attribute are used to generate measurement results in the IAT. The concepts appear visually in a 2-choice task, with the associations in the first task (e.g. flower vs. insect names) and the attribute in the second task (e.g. pleasant vs. unpleasant words for an evaluation attribute). The underlying assumption is that when the instructions include highly associated categories (e.g. flower + pleasant) with a shared response key, performance is faster than when less associated categories (e.g. insect + pleasant) share a key. In their article, the authors show that this performance difference implicitly measures differential associations of two concepts with the attribute (ibid., 1998).

In relation to the measurement of emotional sensitivity as included in the STS, some picture items of the IAT are used in the first part of the STSQ because they represent relevant categories with respect to strangeness (cf. chapter 6.3.1). The IAT and the STSQ (Part I) both apply an indirect measuring method whereby the pictures function as visual stimuli. The measuring instruments (IAT vs. STSQ, Part I), however, differ in their theoretical assumptions (cf. Gawronski & Conrey, 2004). The IAT uses reaction time between two concepts and attributes as an indicator of implicit attitudes; whereas in the STSQ, specific response patterns are supposed to indicate an attributed emotional meaning towards perceived strangeness. The pictures illustrated in the STSQ, Part I are functioning as stimuli to provoke perceptions of differences. The instruction to write down associations when looking at the picture is meant to initiate additional reflected associations for the respondents as regards the supposed meaning of the picture or illustration (cf. chapter 6.3.1).

---

<sup>10</sup> <https://implicit.harvard.edu/implicit/demo>

The following concept of **argumentation integrity** was introduced by Groeben, Schreier and Christmann (1993). Their concept of argumentation integrity was then applied by Bender-Szymanski (2004) in a classroom research project called “Die schwierige Toleranz”. Argumentation integrity outlines three argumentation strategies in the realm of intercultural settings:

1. *Assimilation*: Minorities shall adapt to the norms and rules of the country of residence.
2. *Segregation*: Minorities still keep their cultural orientations without meaningful interactions with the host society
3. *Integration*: Implies mutual interactions of both parties with the opportunity to keep existing cultural orientations and a willingness to develop them further.

Bender-Szymanski developed a role-play which was intended to mediate basic democratic values, further openness for different cultural and religious convictions, and improve solidarity. In other words, a role-play meant to mediate intercultural competence. The role-play is understood as a learning sequence in which to achieve a more rational approach to the management of religious conflicts and conflicting worldviews. Part II of the STSQ was designed to structure items according to these different classes of argumentation and reasoning styles. The intended structure is inspired by the role-play sequence designed by Bender-Szymanski (ibid., 2004). The STSQ, Part II is intended to measure the awareness of different argumentations and reasoning styles with respect relevant situations. One item of the STSQ is closely related to the role-play, and consequently followed precisely the intended item structure of different argumentation strategies (cf. chapter 6.4.2).

There are further concepts such as **Cultural Intelligence, CQ** (Earley and Ang, 2003) or **the Multicultural Personality Questionnaire, MPQ** (Van der Zee and Van Oudenhoven, 2001) which seem to show some relevance in comparison to STS. But these concepts are similar to Bennet’s DMIS focusing more on the emigration perspective. The MPQ was developed in the context of global marketing and managing businesses. Consequently, applications are directed to prepare managers for working efficiently overseas.

“The Multicultural Personality Questionnaire (MPQ) was developed as a questionnaire aimed at measuring multicultural effectiveness. The Multicultural Personality Questionnaire has scales for Cultural Empathy, Openmindedness, Emotional Stability, Orientation to Action, Adventurousness/Curiosity, Flexibility, and Extraversion (Van der Zee & Van Oudenhoven, 2000). Cultural Empathy is probably the most frequently mentioned dimension of cultural effectiveness.” (Van der Zee and Van Oudenhoven, 2001:279)

The focus is clearly directed at effectiveness, international success, and understanding of other cultural background based on this (probably marketing) purpose.

The CQ concept focuses in a similar direction:

“Our focus is to provide an understanding for the age-old problem of the sojourner: Why is it that some people adjust relatively easily, quickly, and thoroughly to new cultures but others cannot seem able to do so. [...] some managers appear lacking in social skills yet adjust effectively to new cultures.” (Earley and Ang, 2003:4)

“What enables some people to function as cultural chameleons while others flounder?” (Earley and Ang, 2003:12)

STS focuses on a mutual learning process where the stranger helps to restructure and develop unquestioned patterns of one’s personal, social and/or cultural background. STS is consequently related to identity construction as will be pointed out later in this thesis. MPQ and CQ provide relatable dimensions, but the different intentions lead to different claims regarding operationalization, application and even interpretation of empirical findings. CI and the MPQ focus more on the assimilative perspective and points out efficiency in communication as the two quotations above indicate. Assimilation does not allow much space for one’s own cultural peculiarities. Assimilation is a one-sided process of trying to adapt and understand the other culture. STS is a pedagogical concept which focuses on integration and not assimilation. Integration is a mutual and dynamic process which is related to identity development for both the natives and the newcomers. This does not necessarily imply developing strategies of dealing with each other in a most efficient way but more in a more rational, respectful and responsible way. This requires critical reflections of often unquestioned own, social and/or cultural norms and standards. STS is a pedagogical approach which is supposed to provide developing strategies to deal with strangeness in a more constructive way. Besides some similar dimensions, the backgrounds and intentions of MPQ and CQ are different from STS. The resulting measurements are consequently also different even though they provide inspiring ideas for my item pool development.



### **3. Theoretical starting point: cultural difference, strangeness and facets of identity**

As indicated in the introduction, sensitivity towards strangeness has its theoretical roots in the larger theoretical research program of IME (cf. Erdmann, 1999b; Gieß-Stüber, 2003b; Gieß-Stüber, 2005; Gieß-Stüber and Blecking, 2008). The theoretical components of IME are focusing on three main areas: cultural differences, strangeness and facets of identity. The initiators of IME focused each on a different component but all of them have their theoretical starting point within general approaches of intercultural education basically according to Auernheimer (1995; 2003), and Nieke (2000). The theoretical approach of IME is directed to improve pedagogical practice in the field of sport and physical education.

I try in this chapter to show why the mentioned three theoretical components of the concept are particularly interesting for applications within the field of physical education and a theoretical guideline for the development of the STSQ.

#### **3.1. The difference: central category within intercultural education**

Increasing global migration has led to multicultural societies in almost all Western industrialized nations. This provokes social changes and problems which societies have attempted to address with different pedagogical approaches. As a result, more “deficit orientated concepts” have gradually been replaced by a more differentiated and integrative approach under the guise of intercultural education and pedagogy (cf. Auernheimer, 1995; Auernheimer 2003).

“Multicultural” means in this context, that a society consists of several cultural groups within a demarcated area such as a country. The term implies further that the different groups of immigrants are and/or will not assimilate and adapt unconditionally (Reviere, 1998). Such constellations can lead to challenges for both groups the dominating group and immigrant groups – specifically, when different value systems “crash”, conflicts become unavoidable. Differences in values probably are one of the major challenges within intercultural education. It is not self-evident that the values underlying this thesis are considered as meaningful in other cultural groups. In how far can we expect that other people are open to new cultural facets? How can we respect other values when they appear as meaningless or probably contradictory to our own values? These questions illustrate the complexity of this topic.

In the context of education, and through the indicated meaning of “multicultural”, *intercultural* education focuses on a mutual learning process which is directed at creating a common future. The perspective has changed from focusing exclusively on minority groups to including critical reflections of majority groups as well (Auernheimer, 1995; Gieß-Stüber, 1999; Nieke, 1995). Intercultural education concentrates on *responsible interactions* between the “majority” and the “minority”<sup>11</sup> (Nieke, 1995). Problems arise when the responsibility of the dominating group or majority is underestimated. Majority members often assign main responsibility towards minorities. It can often be heard that “*they (immigrants) have to „integrate’ into the new society*”. As this example indicates, the term “integration” is reduced to the meaning of assimilation. But integration actually implies to open up existing structures, norms and value systems for adaptations and changes in the way that it becomes possible for the immigrant to integrate and not only adapt to the “new” system. Integration can only be successful if the conditions for integration are present. Consequently, intercultural efforts require interventions which focus on extending and restructuring the as yet unquestioned patterns of native and immigrant constellations (Gieß-Stüber, 2000a).

The school system represents a central arena for intercultural meetings. On the one hand, this can be understood as a chance for intercultural learning but on the other hand, it needs also to deal with problems related to meetings between different cultures. But people seem permanently to avoid the phenomenon of cultural differences (Auernheimer, 1995; *ibid.* 2003). Teachers do not seem to be prepared to deal with related problems and they are mainly acting from an assimilative perspective. Similar difficulties can be met in the context of sport. Sport organizations expect migrants to adopt dominating cultural (sport) patterns at the expense of their own culture of origin (Sonnenschein, 1999; cf. also Walseth, 2004).

Auernheimer points out “cultural difference” as the central category of intercultural education. Learning to deal constructively with differences becomes a general goal for intercultural learning (Auernheimer, 2003; Holzbrecher, 2004). The concept of cultural difference demarcates intercultural learning from related ideas of social learning. *Social learning* focuses more on similarities and common sense. Co-operation, empathy, and “fair-play” are for instance central components within social learning. These components are to some extent prerequisites for intercultural learning as well. But as Auernheimer observed, there is a tendency to overlook or deny cultural differences (Auernheimer, 2003). Furthermore, the impetus of re-

---

<sup>11</sup> The terms “majority” and “minority” are to be understood from a power perspective. Consequently, the majority is referring to dominating social group, and “minority members” are the dominated group members.



---

lated conflicts is not really the difference in itself but more the way ways people deal with difference (Erdmann, 1999b).

A quick look through the daily news often reveals that difference often seems to be the crucial starting point for conflicts like racist motivated infringements on people with different skin-color. There is obviously a need to learn how to deal constructively with these perceived differences. Intercultural learning, in distinction to social learning, establishes difference and the manner of dealing with difference as the starting point for intercultural endeavors (Auernheimer, 2003). Intercultural learning focuses on an improvement of skills in learning how to deal with diversity in a constructive sense. One essential insight is to experience the difference as natural, as we are more or less different from each other no matter where we may find ourselves. Experiences of differences and strangeness are natural as they are inevitable companions of the development of our own personality. Dealing with differences is necessary within any learning processes. Instead of focusing on differences as being solely social constructs, people should develop an awareness of them as an inevitable part of social interactions, interactions which can potentially lead to enrichment in the sense of an incentive for further development. In these terms, the perception of differences takes on a constructive meaning as a baseline for intercultural learning approaches.

The concept of strangeness can in this connection be viewed as an extension of cultural differences. It refers to constructed differences which are difficult to understand or which appear as unexpected. Strangeness is therefore also a central element within intercultural conceptions.

### 3.2. Reaction patterns towards differences and strangeness

In the last chapter I introduced the concept of cultural difference as a central category for intercultural education. The concept of strangeness is closely related to the cultural difference. For analytical purposes the concept of uncertainty helps to demarcate difference and strangeness. The more uncertainty is linked to the difference, the more relevant becomes the concept of strangeness. The implicit uncertainty makes it difficult to deal with strangeness. People try to construct a plausible pictures of the world around them (cf. Heider 1958) but strangeness appears as a kind of opposite to plausibility.

In addition, the complexity in modern, multicultural societies is extreme. Without reduction of this complexity, it is difficult to manage our daily life. We need to classify and structure the world around us otherwise we may lose orientation. Having orientation creates a feel-

---

ing of control and certainty. Making a distinction between the own and “others” is part of this structuring process and creates identity – the stranger symbolizes what we are not.

The symbolic function of strangeness can lead to problematical ways of dealing with it. Strangeness seems often to be „manifested’ to differences in religion, language, the pigment of one’s skin, country of origin, specific disease, lifestyle or profession. People point out the difference in a way that gives simplified and devaluating meanings to it (stigmatization) – probably lacking an awareness that we are in almost every context in one way or another strangers and different to the rest. Such a problematical way of dealing with difference is also way of making the world understandable to us. Stigmatization of otherness can appear in a superficial way as plausible and illustrates that the search for plausible explanations does not automatically imply a differentiated view of otherness and strangeness. Without critical rethinking of such unquestioned pattern, stereotypes and stigmatization can appear plausible to us without realizing the simplified and devaluating character. This classifies stereotyping and stigmatization as destructive ways of dealing with difference and strangeness.

Dealing with difference is also related to power difference between strangers and natives. Power difference can lead to conflicts. Conflicts are often due to generalizing, emphasizing and devaluating specific characteristics of the minorities (Rommelspacher, 1997). Revealed power structures demonstrate that the reasons for conflicts are not the difference itself but more the way people deal with difference and strangeness. Disparaging otherness is a kind of defense mechanism in which one perceives the other person as a threat. By disparaging characteristics of the other person or social group, initially perceived threats become neutralized. In this ways dealing with difference and strangeness can lead to exclusion, stigmatization or even hatred towards strangers (cf. Gieß-Stüber, 1999).

A further problem is related to the construction of national identity of the dominating group (majority). Majority members often believe that their “country of origin” (or their respective nation) is a homogeneous unit defined by common understanding, norms and standards. In addition, the majority usually claims universal validity of their norms and values. Integration of strangers into the society is consequently based on considerations whether the presence of foreigners is functional to the stabilization of the system or not (Gieß-Stüber, 1999). Such one-sided considerations are guided by pragmatic, political and economic calculations of the majority (Baumann, 1995). In addition, the political space with its rules, laws and values is defined by the majority, which moreover has the power to determine who will be included or not. This constitutes status-differences which are supposed to secure the claims of the majority. Consequently, there is no equivalent access to power resources for the minori-

---

ties. Following the pointed mechanism, minorities can assimilate<sup>12</sup> the existing norms and standards of the dominating group or they can react with retreat which finally leads to exclusion from majority resources and networks (cf. *ibid.*, 1999). Such destructive mechanisms can lead to segregation like the development of “ghettos” or in the context of sport to ethnic football clubs etc.

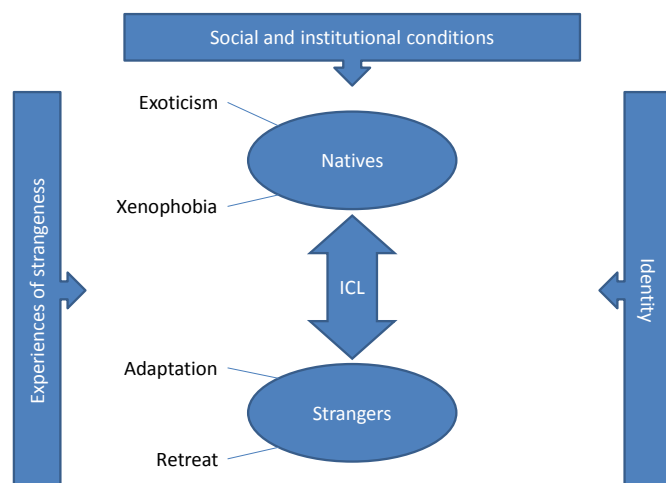
The intention of intercultural endeavors is to achieve respectful and responsible interactions between both natives and strangers whereby the dominating group gains the most responsibility because of their power advantage. Integration is a mutual process which requires that specifically the dominating group need to reconsider existing structures critically with respect to open up the system for integration. But the dominating group’s attention is more often directed towards members of the minority who are forced to give up their personal norms and standards and assimilate or adapt exclusively to the dominating norms and standards of the majority culture (assimilation). Possible differences and strangeness then becomes “neutralized” and will not be perceived as a threat to the majority’s proclaimed identity (Rommelspacher, 1997; Rommelspacher, 1998).

The above mentioned power relations seem to indicate a kind of paradox. On the one hand, identification and stability within a social system (such as a country) is needed in order to function as orientation for the members of the system and thereby ensure security. On the other hand, social systems change and develop. Development always implies a risk because the outcome is uncertain. Yet, the apparent paradox may be better understood as a type of dualism instead of as a paradox. Stability is needed in order to tolerate changes without the collapse of the whole system. The balance between stability and openness for change seems to determine the constructive character of development in general. Following this idea, the stranger can be a “healthy” influence for development because he/she may help to question so far unquestioned pattern and structures of one’s own society.

Intercultural learning including constructive ways of dealing with differences and strangeness is a complex and challenging effort. Gieß-Stüber (2000b:12) illustrates this complexity with the following model:

---

<sup>12</sup> Assimilation is understood as to expect minority groups to adapt one-sidedly to the dominant culture patterns at the expense of their own culture of origin.



**Figure 3: Framework model of intercultural learning according to Gieß-Stüber (2000b:12)**

As illustrated in Figure 3 interactions between natives and strangers are influenced by different factors: experiences with strangeness, social and institutional conditions and identity conceptions. The different points of view between strangers and natives lead to different ways of reacting towards strangeness. The native's extreme poles of reactions can vary between exoticism and xenophobia. The stranger on the other hand can react with adaptation or retreat. In contrast to these destructive extreme types of reactions towards strangeness, IME tries to initiate and support intercultural learning (ICL) between natives and strangeness.

Not being capable of dealing with strangeness constructively can lead to the before mentioned problematical reactions. These reaction patterns are basically directed at "neutralizing" the assumed danger the stranger seems to represent: The stranger is excluded from the majority and assigned to ghettos or the assumed threat becomes neutralized by assigning the stranger inappropriate and disadvantaging labels. These labels are often irrelevant details whose meaning is blown up to gain a symbolic meaning, point out the difference and demonstrate the majority's dominance (cf. Guttadin, 1993).

The different ways of dealing with strangeness can also be illustrated within the field of elite-sport. The "foreigner" or "stranger" is allowed to play within a sport team as long he/she contributes to its success. But when he/she is perceived as disturbance and is not able to func-

tion according to the expectations of the “natives”, the “stranger” can be excluded, devaluated or abused (Gieß-Stüber, 2003).

In addition, conflicts between natives and strangers also arise in a competition for resources and supposed short goods where the stranger can be perceived as a threat.

Destructive reactions towards difference and strangeness are understood as the lack of acceptance and, even rejection of “otherness”. The purpose is to maintain existing power structures and keeping the privileges of the dominant authorities. The “stranger” may also reveal to us what we do not represent or know, and this could consequently put our own knowledge into doubt (Erdmann, 2005). As a result, integration and STS require the involvement of both the majority and the minority. A large number of problems, however, seem to be found on the majority’s side, which suggests that integration can only be obtained, if the society is in principle open and ready for integration (Auernheimer, 1995; Gieß-Stüber, 1999).

### 3.3. Facets of identity

The concepts of difference and strangeness are related to identity constructions. What we perceive as strange depend to some extent on our personal experiences. In this sense represents the stranger what we have not experienced in the past. Our personal experiences define our identity. Consequently, what we perceive as strange may therefore reveal something about one’s own personality (Auernheimer, 2003; Reuter, 2002).

On the other hand, the stranger shows us what we are not and therefore may question our identity. Dependent on our experiences with strangeness and our own identity conceptions we can perceive the stranger as a threat. In a more constructive way, dealing with strangeness can challenge our identity and promote its development in similar ways as identity crises can promote identity<sup>13</sup>.

In spite of the inflationary use of the term “identity” and the variations in connotations, there do remain some common, basic assumptions of the construct. These are particularly helpful for applications within intercultural education and in the realm of physical education. Identity theories also had a strong impact for the development of my measuring instrument.

---

<sup>13</sup> Identity crises are perceived as extremely overwhelming and threatening and they are in addition accompanied by reduced self-worth. If a person managed to solve a crisis more or less autonomously, the crisis was an incentive for entering a new state of elaborated identity. But people also deny crises and leave them unsolved or they deal with the situation by applying imposed coping strategies. According to Hauffer (1995:107) every situation can in principle be evaluated as a threat, harm or challenge to one’s own identity.

In this context, the term “identity” is to be distinguished from the term “personality”. Personality is understood as all collected psychological characteristics of an individual. Identity refers to the individual and its awareness of itself, how the subject constructs and senses its own identity (Haußer, 1995). The term “identity” indicates that there exists a distinction between different units such as between the own and other persons which means that an individual has specific, unique characteristics.

The difference such as you and I is a key-element in identity development (Keupp, 1997). In early childhood, the infant learns gradually to make a distinction between its own body and everything else. It learns that some things such as the own body are always there and some things are there only sometimes such as the objects and people around. The child discovers that boundaries exist between what its “me” and what is “not me”. This early stage in identity development points out the meaning of one’s own body as a central reference. The reference to the own body in identity constructions makes the theoretical construct of identity particular interesting for movement education as will be pointed out later. However, the comparison between the own body and that of others indicates that identity is a relational term. The other person shows us something we do not represent self and through the mirror of the others we construct our identity (“looking-glass self”). Perceived differences show us what we are not but at the same time defines us as individuals. In later periods of development the social comparison becomes more and more important for the individual’s identity work. Evaluations of oneself are made in comparison with a reference group.

The idea of “looking-glass self” (Cooley, 1902) illustrates that we construct and value our individual identity through the mirror of other individuals or social groups. We ask ourselves how do we appear to the others and how are we being evaluated by *significant others*<sup>14</sup>? Having a **relationship** to family, friends, other groups and the society and receiving **acknowledgement** from those significant others are stabilizing elements for individual’s identity. But a closer look at the relationship between the individual and social dimension of identity constructions reveals problematical mechanisms. An overestimated social identity in combination with a weakened and unstable individual identity is seen as problematical. In this case, the identification with a social group or nation is used as kind of compensation for the probably “insufficient” effort regarding the own identity work. An unconsidered or even blind

---

<sup>14</sup> The term significant other refers to persons who are considered as meaningful to the individual’s life such as family members or “close” friends.

identification with the ideas a social group seems to suggest a security which is used as compensation for own individual insecurities (cf. Erdmann, 2005).

The difference is similar to strangeness a relational dimension. Interactions with our social environment are a crucial element in our identity construction (Leontjew, 1977 in Haußer, 1995:8). Through these interactions we learn something about the other person and about ourselves. Perceived similarities serve as social identification, and the differences promote the uniqueness of one's own identity. But not all differences are relevant for the own identity. Basically meaningful "objects" are relevant for identity construction; whereby "objects" is used in the following argumentation as a summarizing term for the environment including persons, relationships, things, events and situations. The *subjective meaningfulness* is a concept within identity theory and is defined as the perceived importance an object has on an individual (cf. Haußer, 1995; Haußer, 2007). In other contexts, the subjective meaningfulness is also called *centrality*, *valence* or *incentive* (*Aufforderungscharakter*). The related term *incentive* points out that the subjective meaningfulness is related to one's personal needs and interest. The implementation of one's personal needs and interests influences how we deal with objects (ibid. 1995). The respective objects gain a specific "valance" or subjective meaningfulness.

The subjective meaningfulness is influenced by the individual's experiences and expectations. The perceived characteristics of the objects determine the *intensity* of meaningfulness or how attracted we are to the object. Variations of the perceived *intensity* have probable a close impact on how people behave or deal with an object. When the cues I receive from an object appear as *attractive* to me, the whole object is considered interesting or fascinating. The other extreme can be illustrated in a way that cues I perceive are considered and valued as not attractive or even "repulsive" and I'd prefer to avoid. Receiving no cues of interest indicates indifference. In this sense, the subjective meaningfulness is related to motivation.

The subjective meaningfulness functions as orientation. We select, control and steer our actions dependent on subjective meaningfulness. A person's meaningfulness can be profiled by various subjective meaning themes e.g. job orientation, political involvement, family, sport etc. (Haußer, 2007). These content related themes represent the cognitive dimension of subjective meaningfulness. The aforementioned attractiveness is the emotional dimension of subjective meaningfulness. The *intensity* represents the perceived degree of involvement whether one is more attracted, more repulsive or just indifferent with respect to the object of perception.

---

I will define this concept of intensity later during my conceptualization of STS as sensibility (in distinction to sensitivity). The *emotions* tell how the perceived intensity is related to the object in categories like more attractive or more repulsive. The *subjective meaning themes* define the content of the subjective meaningfulness. All dimensions together may provide for profiling individual subjective meaningfulness. As will be shown in chapter 4 and 6.4 I tried to apply this idea in the context of strangeness particularly to measure an emotional STS. The basic idea of the concept of subjective meaningfulness is related to Simmel's sociological dimension of closeness/ distance and the Bogardus scale (Bogardus, 1925)<sup>15</sup>. Therefore, these related theoretical conceptions also influenced the operationalization of STS.

The relation of subjective meaningfulness to motivation theory points out that subjective meaningfulness determines to some extent how we react in a situation or how we deal with difference and strangeness. The subjective meaningfulness refers to our past experiences. Strangeness, however, represents something unexpected and in some way beyond our experiences. It is something more or less surprising and new. But past experiences in the context of strangeness are related to what kind of experiences we had with strangeness. What strategies did we develop in order to deal with strangeness? Even though uncertainty defines strangeness and it is difficult to understand such a situation, we may anyway try to attribute a meaning to it. But these constructed meanings can vary in their quality. Some meanings are often based on prejudice and stereotypes, others might be more differentiated and open for alternative interpretations; and that was exactly one crucial starting point that I later in this thesis defines as sensitivity towards strangeness.

Modern theoretical concepts conceive identity as a mental construct of the individual (Keupp, 1997). Its construction results from a continuous process, in which the individual develops and modifies its identity. But identity is not understood as a consistent unity. Also the own identity can be perceived as different and sometimes contradictory, dependent on the situation. The term "**plural identities**" indicates that rather than to look at identity as a stable, consistent unit, it is to be seen as a "**patchwork**" construct consisting of somewhat different, yet **coherent identities** according to the respective social situation (Strauss and Höfer, 1997; Keupp et al., 2008).

---

<sup>15</sup> Bogardus (1925) and Simmel's dimension of social distance and closeness in his essay about the stranger is a kind of sociological analogy of the psychological concept of subjective meaningfulness in the context of identity (cf. Haüßer, 1995; 2007)



---

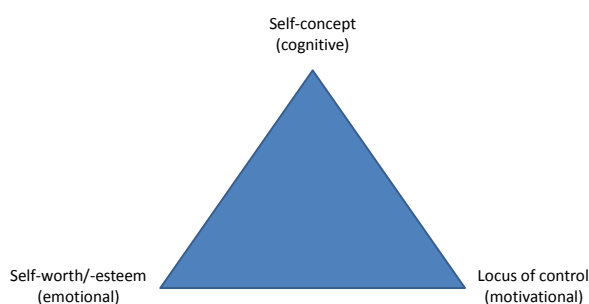
How do we construct this coherence? Subjective meaningfulness works as a filter for situational experiences. We perceive the other person through such a filter. We select, organize, value received information from the others and through assimilation and accommodation we stabilize and construct identity. Both, assimilation and accommodation are processes related to Piaget's concept of development. Identity work is understood as a continuous and interactive process between people's existing and new identity. Identity assimilation means that perceptions and experiences are interpreted on the basis of a person's existing identity. But the process of identity accommodation refers to perceptions and experiences which lead to a change of identity. In this sense assimilation stabilizes identity whereby identity accommodation is an adaptive process which constructs new identity (cf. Haußer, 1997).

The so called core identity (Kernidentität) is cumulated and generalized from experiences and situational perceptions. Experiences can support or question identity. As a result, the identity of an individual is a product of the interaction between the elaborated (and continuously evaluated) "statement" of a person with its social environment (Keupp, 1997). Resulting from preliminary learning processes, the subjective knowledge and expectations determining identity focus on three areas (cf. Haußer, 1995; Frey and Haußer, 1987):

1. **Self-concept:** the cumulated knowledge about oneself
2. **Self-worth:** emotional valuation of the own person
3. **Locus of control** (German: *Kontrollüberzeugung*)<sup>16</sup>: the subjectively perceived ability of a person to explain and/or predict and/or influence events and outcomes in life

---

<sup>16</sup> The "locus of control" concept originates from the control theory of Rotter (1966). It represents the principal need of a person to influence or control events and outcomes of their environment. The initial one-dimensional concept according to Rotter was extended with the three dimensions according to Frey et al. (1977): attribution (German: Erklärbarkeit), anticipation (German: Vorhersehbarkeit), and influence (German: Beeinflussbarkeit).



**Figure 4: Triad of skill-related identity (according to Haußer, 1997:128)**

The three concepts can be illustrated with an example from physical education class. A P.E. student asking him/herself “*How good am I at sports?*” refers to the cognitive dimension (self-concept regarding physical skills). The question “*How do I feel about the fact that I am not good at sports?*” illustrates the emotional component (self-worth/-esteem regarding physical skills). “*What can I do to improve myself in sports?*” refers to the motivational dimension (Locus of control regarding the belief of being able to control or change one’s own skills).

Haußer (1997)<sup>17</sup> points out some interesting findings which are important for measuring the three concepts shown in figure 4. At first, highly generalized self-concepts of an individual are relative stable concepts regarding intervention in comparison to more specific self-concepts and situational self-perceptions. Secondly, the emotional side (self-worth) appears to be more stable in comparison to the cognitive dimension (self-concept). This means that it is more difficult to influence self-related emotions than self-concepts. Finally, the relationship between an individual’s attitude and behavior becomes more consistent with increasing subjective meaningfulness of the respective identity related object or aspect. In addition, with increasing meaningfulness the locus of control increases as well. This means that the belief of being able to influence and control things in life increases with the subjective meaningfulness. Haußer (1997) also points out that it is more difficult to measure the emotional and motivational perspectives than the cognitive dimension.

I have tried to demonstrate that an individual’s identity influences ways of dealing with difference and strangeness. Dealing with difference and strangeness are crucial parts within daily identity work. But the stranger can be perceived as threatening. Supporting identity development may help to improve constructive ways of dealing with difference and strangeness.

<sup>17</sup> These findings are based on studies conducted by Haußer, 1995; Rheinberg 1993; Epstein 1993; Haußer and Kreuzer 1994; Frey and Haußer 1987 (quoted in Haußer, 1997:129)

---

In order to tolerate the implicit uncertainty of strangeness, efforts should be directed to both stabilizing identity and elaborating identity. Since the body is the individual's personal reference for identity constructions, movement education may offer special opportunities for identity development because the body and movement are central. In addition, identity is constructed by the individual (individual identity) and in relation to others (social identity). Sport and physical education are social arrangements and depend on relating to each other. These relationships are characterized by co-operation, inclusion and fellowship but also related to challenges to one's own identity development and conflict management. Therefore, in P.E. sensitive pedagogical treatment is needed in order to arrange opportunities which can help to stabilize identity and allow constructing new identity. Experiencing **acknowledgement** and **social affiliation** are central pedagogical guidelines in the context of physical education. Situations can be constructed in order to mediate constructive ways to deal with difference and strangeness (cf. chapter 5).



---

#### **4. Further theoretical considerations and a conceptualization of “Sensitivity towards Strangeness (STS)”**

The goal of this thesis is an operationalization of strangeness as being part of the underlying framework concept of IME. Strangeness is understood in this thesis as a multi-faceted construct because it represents ideas of the broader concept of IME.

Strangeness is a theoretical construct and is not directly observable. An operationalization means deriving observable indicators out of theoretical ideas. This requires a focus on facets and dimensions of strangeness which direct the attention towards observable indicators. Considerations around *how people perceive and deal with strangeness* seem to be relevant aspects to generate measurable information. Identity conceptions and past *experiences* have an influence on people’s behavior and an operationalization in the context of strangeness may consequently generate information about how people deal with difference and strangeness potentially.

Measuring in social sciences means basically a comparison of individuals or different groups according to defined criteria. The conceptualization of *sensitivity towards strangeness (STS)* is needed in order to support such criteria theoretically for the intended measuring instrument. In this sense, the more general theoretical background and more concrete conceptualization of STS are the rationale behind the even more concrete operationalization (STSQ). The criteria for group comparisons are founded in the theoretical background of my thesis. On the other hand, the following conceptualization is inevitable needed for interpretations of measuring results.

I start the conceptualization of STS with some basic considerations about connotations of strangeness. A basic clarification of the terminology and facets of strangeness revealed the first hints for the intended operationalization. Some specific insights into human perception from a social psychological point of view is expected to support an understanding of perceiving differences and strangeness as a more natural part of a person’s development than a threat to his/her identity. But not all differences lead to meaningful experiences of strangeness and only meeting the “stranger” does not automatically lead to a better mutual understanding. The meeting needs to be considered as meaningful and needs to gain a certain degree of subjective meaningfulness. Therefore, some conditions leading to meaningful experiences of strangeness will be pointed out in the next chapter (4.4). The basic components of “sensitivity towards strangeness” (STS) will be described in chapter 4.5.

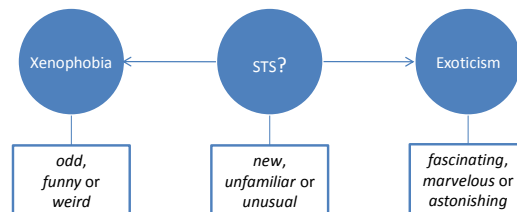
#### 4.1. Connotation of strangeness

Having a closer look at the everyday use of the term “strange” already reveals relevant dimensions of the construct. The adjective “strange” originates from Anglo-French “*estranger*”, from Latin “*extraneus*”. Literally, it stems from external, from “extra” outside. In the German language, the term strange is translated as “*fremd*” and originates from the Old High German (althochdeutsch) term “*fram*” which means forward, further, from and away but it is also interpreted as distant, unknown and unfamiliar (Der Duden, 2001:235). Specifically the connotation of “forward” and “further” indicates an idea of move on, progress or development. The German substantive “*Fremde*” demarcates this understanding of strangeness further. “*Die Fremde*” is initially associated with abroad or a distant country from home (Land fern der Heimat” [ibid. 2001]) which points out the dimension of closeness and distance according to Simmel’s essay about the stranger (Simmel, 1992). Together with the understanding of “*die Fremde*” as a distant and foreign country, terms such as separation and enemy or hatred became associated with the German term “*Fremde*” as well. By drawing a borderline between one’s own and foreign country, the stranger is separated and beyond this borderline. The stranger becomes an enemy when he/she crosses such a borderline and is perceived as an unpleasant invader because he/she may disturb and threaten familiar structures, traditions and habits (cf. Noethlichs 2005a).

Further related terms such as “surprising”, “unexpected” indicate the characteristic uncertainty-facet of strangeness. We usually label situations, events or objects as strange which are realized as different, unexpected, new, unfamiliar and difficult to arrange in a consistent way into our previous knowledge and experiences.

Synonyms of the term “strange” point out another important dimension of strangeness. Terms such as *new*, *unfamiliar* or *unusual* are neutral or less value-loaded classifications of strangeness. These adjectives indicate the only fact that a person, an object or a situation is perceived as unfamiliar. Terms such as *odd*, *funny* or *weird* implicated a more value-loaded connotation but in a more devaluating way. Synonyms such as *fascinating*, *marvelous* or *astounding* indicate also an emotional involvement (or subjective meaningfulness) also but in an overwhelming positive direction (Mirriam-Webster, 2009; Noethlichs, 2005a; Oxford Dictionary, 2009). Bringing the mentioned synonyms together on a kind of dimension, suggests that the connotation varies between negative, neutral and positive classifications. This carefully indicated dimension refers to different ways of dealing with strangeness. To contrast the ends

of this assumed dimension even more, it seems that dealing with strangeness can vary between xenophobia and exoticism (cf. fig. 5).



**Figure 5: Sensitivity towards strangeness (STS) something between xenophobia and exoticism?**

The exotic perspective shall point out that strangeness is not only associated in negative ways, but it can also be perceived as surprising, interesting or exciting. However, a supposed positive evaluation of the stranger does not automatically mean that this way of dealing with strangeness is unproblematic. Mystifying ravings about foreign (“exotic”) cultures may appear especially in contrast to xenophobic tendencies as a positive interpretation. But such unreflective focus to exclusively positive aspects of foreign cultural goods view the stranger in a reduced and devaluating way. It views the individual in a distorted and one-sided way and only indicates that the stranger is not taken seriously (cf. Gieß-Stüber 1999, 46f.). Therefore, it seems more useful to distinguish between **constructive** and **destructive** ways of dealing with strangeness because those terms allow a clearer conceptual distinction between intended educational and more problematic coping strategies.

#### 4.2. Facets of strangeness

Strangeness is a theoretical construct which refers to social phenomena where people, situations, or objects are perceived as different from the individuals’ own perception. Generally, this involves two perspectives: 1) the stranger’s view, where one could feel strange on a personal level in a specific situation, such as an “outsider” within a group of unknown people, and 2) from the “insiders” perspective, where a person is perceived as strange by others. In both perspectives, the perception of a difference is a crucial prerequisite to experience strangeness. Both perspectives are necessary in emphasizing that the phenomenon of strangeness is specific **social construction** and not a characteristic of a person or social group. Strangeness can be experienced when persons or social groups with different backgrounds

come closer. Strangeness is consequently a **specific relationship** which is intensified by becoming closer to each other. The social closeness is needed because only then can constructed differences gain a subjective meaningfulness. Only meaningful differences can lead to the often problematical conflicts on personal, social, political, economical or cultural levels. But what makes it difficult to deal with strangeness in a rational way? As indicated previously, the stranger can question our identity because he/she represents something different. But the difference is not the only characteristic of strangeness. We have difficulties understanding the supposed clashing backgrounds. Strangeness is therefore a concept of uncertainty.

#### 4.2.1. Strangeness, a social construction

Difference and strangeness are social constructions. These terms are abstract constructions we make, in order to facilitate our comprehension of complex phenomena with many, related facets. These terms are theoretical or latent ideas which are not observable directly. We cannot see, smell, or touch difference or strangeness. But usually we have an idea about those concepts. These ideas are shaped from one's own subjective experiences and interpretations. We can observe the ways people deal with difference and strangeness. These observations may reveal to some extent people's subjective meaning related to difference and strangeness.

The basic function of constructing difference and strangeness is to differentiate between "I" and "you" or "we" and "them". By drawing a borderline between one's own (familiarities) and another person or social group, we construct difference and strangeness. In this sense, the social construction of difference is a kind of *structuring concept* with the function to give orientation. This structuring concept is probably a basic function of social constructs in general such as illustrated in the previously introduced concepts of individual and social identity according to Haußer (1995) or the concept of culture according to Auernheimer (2003; cf. also Erdmann and Schulz, 1999).

We are almost everywhere different from each other and can be perceived as strange or feel strange. The following quote illustrate the social construction of strangeness further:

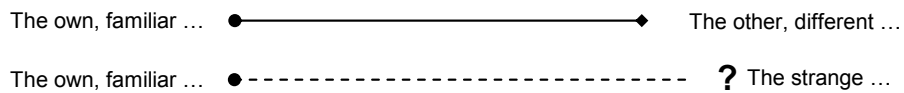
„[...] denn als fremd wird der Fremde nur in der Fremde wahrgenommen und jeder der sich fremd fühlt ist nur so lange ein Fremder, bis er sich nicht mehr fremd fühlt, denn dann ist er kein Fremder mehr. [...]“ (Valentin, 1984:488 f.)

This quote means that only "abroad" the stranger can be perceived as strange. A person who feels strange is a stranger as long as he/she does not feel strange anymore. Then the stranger is no longer a stranger. The phrase "in der Fremde" is translates with the term



“abroad”. But “Fremde” is understood in a broader sense than geographically such as a different country. A different context within the same country or places can lead to perceptions of strangeness as well. At the beginning of new school year for instance a new class of pupils may appear as strange to their new teacher because he/she does not know the students in the beginning and vice versa. After teaching a while, mutual perceptions of strangeness change more and more into familiarity. However, there might always be a chance to experience strangeness anyway even though teacher and students became more and more familiar with each other through daily contact and interactions. Each individual, social group or cultural unit have their own historical background which affects its presence and future (cf. Schäffter, 1991). Becoming closer and meeting each other with diverging histories can in principle create tensions and strangeness in every situation. Consequently, strangeness does not only refer to the natives and immigrant constellation. Historical backgrounds vary in all levels, the individual’s, social, institutional or cultural level.

The construction of difference and strangeness are guided by social norms and standards. In order to illustrate this, a difference can be imagined as the perceived discrepancy of a specific characteristic between the own person or group and/or another person or group.



**Figure 6: Difference vs. strangeness**

The perception of such differences does not automatically lead to experiences of strangeness. Usually, people of the same community seem to have quite clear ideas about how they differ from each other. Consequently, they perceive themselves as different to another individual, but the other person is not necessarily perceived as strange. When people’s associations appear as plausible to them, then, the other person can just be perceived as different or a difference is attributed in order to demarcate oneself from the other person, as is part of the process of identity construction. The difference is perceived as strange when people may not be able to explain the way in which they differ from another person or do not understand the situation. Thus, people realize strangeness emotionally as confusing, funny or irritating.

It seems like perceptions of strangeness often are related to people’s *expectations* in the situation. Comparisons between initial expectations and the way the situation actually is per-

ceived at the moment may result differently and can then be perceived as strange (because it was unexpected). Culturally different welcome-habits can illustrate the social construction and unexpected character of strangeness further. They can be perceived as strange in different contexts and even have an entirely different meaning (cf. Figure 7).



**Figure 7: Norms (e.g. type of greeting) may vary across cultures (Passer and Smith 2004:602, fig. 15.10)**

For instance, in some cultures men kiss each other as a conventional welcoming gesture; whereas in other cultures this could be perceived as a sign of a homosexual relationship. The aforementioned example illustrates yet again, that what we perceive as different and strange depends on perspective, and this again reinforces the idea that difference and strangeness is socially constructed.

#### 4.2.2. Strangeness, a specific relationship

Strangeness as a social condition and refers to specific relationships between people, such as between natives and immigrants (Gieß-Stüber, 1999; Gieß-Stüber, 2000a). In Georg Simmel's social construction of strangeness, he introduces the dimension of social closeness and distance in order to define and describe strangeness as a specific human relationship (Simmel, 1992; cf. Bogardus 1925). Strangeness is a social relationship and becomes even more intensive when coming closer to the stranger (Schäffter, 1991). Thus, the stranger is the person who comes today and stays tomorrow, while the wanderer, on the contrary, comes today and leaves tomorrow. According to this definition, strangeness is referring to the stranger, who may be perceived as very different and though "distant" in Simmel's terms, is actually not leaving "tomorrow" and therefore "closer" than expected.

Being closer to a "stranger" implicates also a more meaningful relationship (cf. Schäffter, 1991). According to Simmel's concept of the stranger, strangeness needs to be viewed in terms of meaningful relationships, an aspect which he illustrates by imagining people on a

---

very distant galaxy. These “aliens” are according to Simmel’s understanding not really strange to us, because, in a real sense, they do not really exist for us at all. These hypothetical people are beyond closeness or distance (ibid., 1992). They are irrelevant to us, and consequently, they are irrelevant for a meaningful experience of strangeness. A relationship between the stranger and the native must be perceived, and evaluated as, meaningful to both parties. Otherwise they are according to Simmel’s illustration, “beyond closeness or distance”, and therefore not relevant in relation to the experience of strangeness.

#### 4.2.3. Strangeness, a concept of uncertainty

The crucial characteristic of strangeness is the unexpected and unknown. This uncertainty is related to characteristic emotions which can be described as funny, surprising, irritating, and threatening. Characteristic emotions related to perceptions of uncertainty are quite distinctive in comparison to the content perceptions of strangeness can be related to. The content of experiences of strangeness is subjective and historically bounded and can vary more or less between individuals, social and cultural groups because it is related to identity. The more distant (in a figurative sense) the object of perception is the more variable may be the content of what leads to an experience of strangeness. But there might on the other hand exist a consistency of strangeness related topics and situations the closer the background of interacting individuals, social or cultural groups. Members of the same (majority) group may share similar perceptions and experiences of certain minority groups and vice versa.

The initial mechanisms leading to perceptions of difference and strangeness are, in principle, similar to such self-conceptualizations where one’s own identity is perceived and constructed by demarcating ourselves from others (Haußer, 1995; Hirschfeld, 1993; Holzbrecher, 1997; Porter, 1973). Furthermore, during specific stages in childhood development, infants are particularly afraid of strangers (Gerrig & Zimbardo, 2002). This is why the strange person is perceived – in this case by an infant – as quite different, and, consequently, these perceptions create feelings of insecurity/uncertainty. As a result, infants experiencing strange or unfamiliar situations need to feel secure in order to tolerate and deal constructively with the given uncertainty. Research also reveals that attachment to parents is quite important in young lives in order later deal constructively with other persons. “Secure attachments to adults who offer dependable social support enables the child to learn a variety of pro-social behaviours, to take risks, [and] to venture into novel situations [...]” (Gerrig & Zimbardo, 2002:344). In other terms, dealing with strangeness challenges our identity and can so be part for construc-

tive identity work. But in order to perceive strangeness in such a way, we need stability or security in the background. This stability is needed in order to tolerate the implicit uncertainty and to not perceive strangeness as overwhelming and possibly threatening (cf. Buhr, K. & Dugas, M. J., 2002). Consequently, a balance between stability and openness may characterize a good starting point for dealing constructively with strangeness.

Research in the learning process has revealed a significant relationship between the different degrees of novelty and problem solving tasks. A basic conclusion of this research was that only in situations where the degree of novelty was moderate, was the intelligent use of past experience proven to be beneficial to problem solving (Raaheim & Brun, 1985). Dealing with strangeness constructively can be related to this problem solving learning approach. The earlier mentioned balances between stability and openness is related to a moderately perceived level of novelty<sup>18</sup>. In other cases the novelty may be perceived as so overwhelming that it is difficult or most likely too difficult to handle in constructive ways.

The concept of uncertainty moreover defines strangeness in comparison to concepts of “otherness”. Regarding strangeness, we cannot predict the outcome of a strange situation, or the stranger’s attitudes and intentions. This unpredictability characterizes the concept of uncertainty (cf. Gudykunst et al. 2003). The subjectively perceived degree of uncertainty varies between certainty or confidence which means that no doubt exist and absolute unpredictability or uncertainty respectively (cf. *ibid.* 2003; cf. Erdmann 2002). The perceived degree of uncertainty is a subjective dimension. Being absolutely confident of knowing the other person does not necessarily mean that the assumptions about the other person are correct.

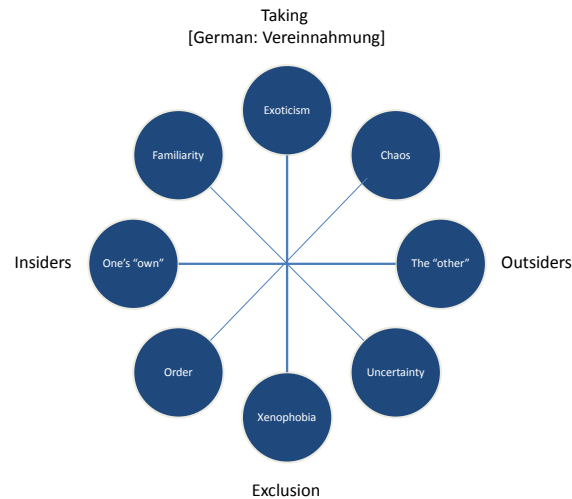
As shown above, strangeness is a complex phenomenon which is related to several constructs and dimensions. Despite its dependence on subjectively influenced perceptions and experiences, strangeness reflects some general patterns, such as:

- a) The perception of difference is the crucial starting point for experiences of strangeness,
- b) The implicit uncertainty defines strangeness further,
- c) Strangeness needs to be considered as meaningful,
- d) Strangeness is related to identity construction, and
- e) Dealing with strangeness varies between xenophobia and exoticism, or better constructive and destructive ways.

---

<sup>18</sup> whereby novelty is specifically associated with the degree of individually perceived uncertainty

The following figure shall illustrate implicit dimensions of the aforementioned considerations about the connotation of strangeness in summary:



**Figure 8: Dimensions of strangeness**

#### 4.3. Dealing with difference and strangeness

The following considerations are related to basic concepts of social perception, classic attribution theory and Piaget's developmental concepts of assimilation and accommodation. These concepts are helpful analogies which can be applied to the phenomenon of strangeness in order to point out problematical/destructive and constructive ways of dealing with differences and strangeness.

Strangeness is related to differences and uncertainty. All differences between individuals imply a degree of uncertain information. In relating to other persons we generate a picture in our mind which represents a kind of theory about the respective person. Through interactions we learn more and more about each other. Our subjective "theory" about the other person becomes "tested" and probably adapted or revised the more information we receive and treat. This is what is meant by deconstructing strangeness. It is a kind of learning process. But it is probably impossible to deconstruct strangeness completely – sometimes we may reach limits of our understanding and a certain degree of uncertainty remains. However, the psychological process of perception and attribution theory may help to point out some basic patterns and mechanisms when dealing with strangeness and implicit uncertainty.

The psychological process of perception indicates that our prior knowledge, experiences and expectations are represented in our mind in an abstract form as cognitive schemas. If the perception process results in inconsistencies between initial expectations and actual perceptions, then a situation is experienced as strange. Accordingly, the perception of strangeness can be explained as experiencing a lack of categories or problems in classifying perceived differences caused by inconsistencies within our existing cognitive schemas<sup>19</sup> (Bennet, 1993).

Even though “schema theory” is accused of not being clearly described and being inaccurate or simplistic specifically when applied to unique individual cases (cf. Fiske and Taylor 1991:173ff.), this theoretical concept<sup>20</sup> helps to explain some rigid but crucial reaction patterns towards strangeness which often become very relevant in actual conflicts between natives and immigrants. In combination with Piaget’s concepts of assimilation and accommodation, schema approaches underpin constructive and destructive ways of dealing theoretically with strangeness.

The goal of this thesis was the development of a measuring instrument which was supposed to measure a band-width of relevant facets of strangeness. Developing a measuring instrument implies reduction of respective complex phenomena. The focus is on central patterns in order to enable for approximate differentiations between different populations. For this purpose, the schema concept appears to be suitable. Since STS and the resulting STSQ are understood as multi-faceted, schema approaches serve as additional theoretical reference for some of the item constructions of the STSQ; the STSQ refers for instance to *perception* and *attribution* patterns towards strangers. Having in mind the STSQ, problematical patterns of dealing with strangers, e.g. stereotyping, stigmatization, or prejudice, can sufficiently be explained by schema approaches. As already indicated, the STSQ is not supposed to be applied for individual diagnostics. Therefore, schema approaches in spite of all the critique function sufficiently in order to point out basic human perception patterns and specific ways of dealing with it<sup>21</sup>.

---

<sup>19</sup> whereby categories are understood as hierarchical structured classification systems and schemas are more abstract cognitive structures

<sup>20</sup> Because of the indicated critiques (for more comprehensive discussion cf. Fiske & Taylor, 1991), I would rather call it a theoretical concept than a theory. A theory would be more of general character and probably more precisely.

<sup>21</sup> Previously, strangeness was viewed from a sociological point of view whereby strangeness was viewed as a structuring concept. Both perspectives deal with schemata or categorizations but they bring light into the phenomenon of strangeness from different perspectives and more or less differentiated.

---

Generally speaking, the basic process of human perception can be summarized as a sequence of the detection, selection, structuring, and interpretation of incoming signals from our environment (Gerrig & Zimbardo, 2002). Our mind, however, is not able to consciously process all incoming signals. Consequently, most of the signals detected by our sensory systems are not converted into recognizable information. People continuously have to filter or select the constant flow of data from the outside world. What we consider as our environment is already noticed and structured information. Reduction is necessary in this so called *bottom-up* or *data-driven* process of perception (Passer et al., 2004). Already existing cognitive schemas, however, serve as a kind of interpretation foil in order to select, structure, and make the incoming data interpretable.

Imagine for instance a person looking on a written page in book, a person who has never learned to read. This person has no compatible schemas for written language. Such a person may only perceive a piece of paper with various signs (letters) on it. These „secret symbols‘ may appear as unstructured marks, and they would make no sense to this person. This example (Kriz, Lück, & Heidbrink, 1996) illustrates a perception of strangeness and that we are dependent on existing schemas<sup>22</sup> in order to structure and interpret incoming data. A parallel can also be drawn to the previously introduced conceptions of identity work. The existing schemas can be imagined as our past experiences which define us as individual and are used as interpretation foil as well.

People, however, differ individually in their perception depending on which type of the above mentioned perception processes is dominant. In interactions with others, people are revealed to be more sensitive with respect to the specific qualities of the other individual when data-driven processes dominate their perception. Fiske and Taylor state that:

“[...] purely schematic theories have, in the extreme cases, portrayed people as blithely glossing over important details, as stubbornly refusing to see the information in front of them, and as maintaining their schemas at any cost. In contrast, data-driven approaches show that people do indeed care about the information given [...]”(Fiske et al., 1991:98-99)

Even so, schemas are essential cognitive tools when making assumptions about other people, ourselves or the situations we perceive. They enable us to make our world reasonable, and

---

<sup>22</sup> “A schema may be defined as cognitive structure that represents knowledge about concept or type of stimulus, including its attributes and relations among those attributes [...] The schemas facilitate what is called top-down, conceptually or theory-driven processes, which simply means processes heavily influenced by one’s by one’s organized prior knowledge, as opposed to processes that are more bottom-up or data-driven [...]” (Fiske et al., 1991:98).

often they are accurate enough even if they sometimes are sadly mistaken (Fiske et al., 1991). If the cognitive schemas are too simply structured or are based on misleading associations, the resulting conclusions can be misleading or wrong even though they are perceived as plausible.

Strangeness is in the first place not perceived as plausible. It is moreover perceived as confusing, and can therefore lead to disorientation. This challenges our ability to function without the usual level of prediction and control by schemas. Prior knowledge, such as a map about an unknown university campus, expectations about a foreign culture from guidebooks, or an introduction by a mutual friend to the stranger we meet, would facilitate each encounter. Nevertheless, our inevitable reliance on such prior knowledge is not perfectly adaptive, as illustrated by relying of wrong assumptions or a too simplified understanding (Fiske et al., 1991). Anyway, we have a need to create consistent pictures of our world around us. The perception of something new as unusual or strange requires an effort in order to adapt them to our existing cognitive structures. This process of “making sense” can be understood as the active reorganization of our cognitive schemas or concepts. As a result, strangeness needs not to remain strange; it can be deconstructed and made more and more familiar.

Misunderstandings and conflicts related to strangeness often seem to be caused by errors in the process of making sense (Heider, 1958). The process of “causal attribution” may underlie a fundamental attribution error<sup>23</sup> (Fiske & Taylor, 1991; Gerrig & Zimbardo, 2002): People have a tendency to attribute the cause of identical behaviour from a person belonging to their own group differently to that of a person perceived as a stranger. A typical consequence is that people tend to make “outsiders” personally responsible for what they perceive as an inappropriate behaviour (internal attribution), whereby people tend to make situational factors responsible (external attribution) when they or one of their own group make an identical mistake. Furthermore, this attribution error leads more often to the exoneration of the person inside the group, while it leads more often to the incrimination of outsider groups or strangers (cf. Thomas, 1992).

These errors in perception and attribution are basic mechanisms for the stereotyping of different groups. Stereotyping is understood as the oversimplified categorization of the other person or social group based on a significant lack of information. The fact that dealing with strangeness is based on lacking information about the other person and in combination with

---

<sup>23</sup> Fundamental attribution error means the dual tendency to overestimate the influence of dispositional factors of a person’s behaviour and to underestimate the impact of situational factors (Gerrig & Zimbardo, 2002:546).



---

the aforementioned need for plausibility, stereotyping can be a typical strategy when dealing with strangeness.

An additional consequence of the fundamental attribution error is that people tend to be more attentive to the perception of similarities within their own group and to stress perceived differences between themselves and strangers more distinctively (cf. Thomas, 1992:239). Stressing the difference means pointing out the difference from a self-centric point of view (cf. Gieß-Stüber, 1999). This increases the distance towards the other person or social group and may inhibit meaningful interactions. Pointing out the perceived differences may also promote stigmatizations because the differences gain a symbolic meaning and different characteristics become attached to the difference.

A consequence of attribution errors in people's perceptions is that people tend to take credit for their successes and deny responsibility for their failures (*self-serving bias*). As psychological experiments have demonstrated, people judging a member of their own group make strong attributions to themselves when their own group has succeeded. When judging a strange group, they tend to make the strangers responsible for the failure (cf. Gerrig & Zimbardo, 2002:547-548). In this sense the attribution error can be understood as self-defense mechanism. In the context of immigration, majority members often tend to make minority members responsible for failed integration. Phrases from majority members like "they [immigrants] have to integrate themselves into society" illustrate this problem. In this context integration seem to be understood as adaptation or assimilation and not as real integration which requires mutual efforts.

Deconstructing strangeness into familiarity can be compared to problem-based learning according to Piaget's concept of accommodation. During a child's cognitive development, accommodation is the process by which new experiences cause existing (cognitive) schemas to change (Passer & Smith, 2004). A crucial consequence of accommodation is that every time a schema is modified, it helps to create a better balance between the environment and the child's understanding of this environment (ibid. 2004).

The analogy to the thinking process and problem solving strategies of Piaget, however, shows also limitations when trying to apply it to strangeness. It needs to be pointed out that it might seldom be possible to level out or understand all perceived differences between distinguishing cultures (Auernheimer, 2003). As a result, it is necessary to learn to accept and moreover respect differences even though one is not capable of completely understanding them.

In summary, theory-driven perception is based on more general and abstract schemas which, to some extent, help us to operate in unexpected situations. When perceiving strangeness, the cognitive schemas lack information (Bennet, 1993). These schemas usually aid orientation, but in relation to the perception of strangeness, they may be unreliable which can lead to feelings of confusion and missing orientation. In order to deal with strangeness constructively, we are more dependent on the data-driven procedure as prerequisite. The use of cognitive schemas is needed in order to manage an uncertain situation because they help to interpret and control a situation. On the one hand, dealing with strangeness constructively requires awareness that our interpretations can be wrong and/or too simplified. An open attitude is therefore needed which implies a willingness to adapt and change our existing schemas.

#### 4.4. Meaningful experiences of strangeness

A STS requires meaningful experiences with difference and strangeness. As indicated in the context of identity theories, a subjective meaningfulness is needed in order to be relevant for identity constructions (individual, social or cultural identity).

Specific differentiation between the “perception” and “experience” of strangeness is beneficial in distinguishing the different levels and qualities of strangeness and defining sensitivity towards strangeness. First of all, perception is understood as a psychological process. It is the process of organizing sensory stimulus input and assigning it meaning as explained previously. In addition, perception implies becoming aware of the object of perception (Fiske et al., 1991). The process of social perception refers specifically to the process by which people perceive other persons and/or social groups and the process by which they deal with each other (Thomas, 1991). A social perception of strangeness, however, refers to the previously described relationship between “natives” and “strangers”. An experience can be understood as a comprehensive term representing knowledge, skills, behavior and/or an insight gained from further reflections around perceptions, understandings and feelings accumulated through one’s life. A perception is in comparison to an experience more related to becoming aware of an actual event or situation, such as a first impression of another person. Past experiences, among other factors, do determine to some degree the outcome of a perception. Past experiences function as a kind of interpretation foil.

The German translation of experience “*Erfahrung*” has a slightly different implication. In comparison to the English term, “*Erfahrung*” is connoting the coherency of life’s experiences. In order to generate an experience of strangeness, one needs to deal with it further which

means that reflections and evaluations about perceptions are unavoidable. The idea to differentiate between perception and experience originates from the relationship between the German terms of “*Erlebnis*” and “*Erfahrung*”, which has its origin in the so called “*Erlebnispädagogik*” (cf. Fischer et al., 2000). Both “*Erlebnis*” and “*Erfahrung*” are usually translated in English with experience. “*Erlebnis*” seems to be closer related to impressions and perceptions related to a situation or an event. “*Erlebnisse*” refer to impressions and perceptions in a holistic way. But situations and events only become an “*Erlebnis*” if they are meaningful to the individual. Then, “*Erlebnisse*” can challenge our personality because they are related to taking risks and adventure. In a pedagogical setting, they can become an impressive starting point or individual references for active identity work and learning experiences.

As pointed out in earlier chapters, subjective meaningfulness is related to our identity. Our past experiences define our identity. In order to create our identity we need to work or deal with our daily perceptions and impressions as illustrated with assimilation and accommodation. These processes require further reflections beyond the respective situations or events. The perceptions and impressions need to be related to the individuals past experiences and assimilated and/or accommodated to the individuals identity. Then, an “*Erlebnis*” can become an integrated experience of a person’s identity. Experiences are accumulated with knowledge and insights on a more abstract level than “*Erlebnisse*”. “*Erlebnisse*” are more on a sensory level and more related to feel and sense the situation. An “*Erlebnis*” refers to situational perceptions which are stronger affected by emotions which result into a kind of diffused awareness. Those feelings can later become a characteristic reference as part of a gained and meaningful experience. Cognitive reflections help to structure and adapt or accommodate impression and perceptions. But further considerations and reflections are not automatically related to “*Erlebnisse*”. Without initiating discussions and reflections, an adventure event may remain as impressive but probably without a learning effect. Strong impressions may remain as strong impressions but regarding experiences of strangeness, it is necessary to initiate a learning process by using situational impressions and perceptions as starting point for further theoretical reflections. Such a qualitative learning approach refers to experience-based learning strategies (Noethlich, 2000; Esser-Noethlich, 2010; cf. Hotz, 1997).

Negative experiences with strangeness can also become constructive elements for development of a STS. Irritations and disappointing experiences can be a starting point for constructive discussion and further considerations. Irritations in the context of strangeness can arise when we for instance become confronted with unusual or even provoking behavior of

the “others” contradicting familiar values or norms. The resulting conflict can be used as starting point for reflections around own and familiar standards. The confrontation and subsequent discussions can either help to strengthen our point of view or lead to consider restructuring our initial point of view (cf. Gieß-Stüber, 2008).

For the development of the STSQ, the cognitive parts of my developed questionnaire refer to different qualities of experiences and insights regarding sensitivity towards strangeness. The sensitivity is therefore an indicator for potential qualities of strangeness related experiences.

#### 4.5. Sensitivity towards strangeness (STS)

In this chapter I conceptualize STS more specifically. This conceptualization is based on previous theoretical perspectives and considerations. Since it is intended that the STSQ measures in band-width, STS is understood as a multi-faceted concept. The starting point is the social construction of difference (cf. chapter 3.1). When the difference is difficult to understand it becomes related to feelings of uncertainty, and perceived as strange. Both concepts- a) the perception of difference as prerequisite and b) strangeness as a specific concept of difference were used as the starting point for the development of observable indicators in constructing my questionnaire (STSQ).

Meeting “the stranger” is prerequisite for an experience of strangeness. Personal experiences with strangeness are supposed to support a better understanding of how it feels being perceived as different and strange. Perceptions of the own cultural background from a self-centric point of view can superficially see the own cultural background as homogeneously structured. Experiences of strangeness can promote an understanding that one’s own culture is only one of various other cultures. Deeper theoretical insights into the social construction of difference and strangeness are supposed to support such a relativistic and pluralistic understanding of culture.

STS means moreover that strangeness is understood as an opportunity for modifications and development of one’s own culture, social group and even one’s own identity. To meet the stranger is prerequisite but in order to create meaningful experiences of strangeness we need to deal further with perceived strangeness. Reacting with stigmatizations and prejudice is one (irrational and problematical) way to deal with strangeness. Dealing with strangeness constructively means deconstructing strangeness. By dealing further with our perceptions, strangeness can be deconstructed and transformed more and more into familiarity. Therefore, dealing with strangeness means essentially to re-construct the “new” or “unusual” into the

“familiar”. In other terms, the deconstruction can be understood as a learning process. The analogy of the process of accommodation in Piaget’s classical learning theory demonstrates the principle behind the idea of deconstructing strangeness and dealing with strangeness constructively.

As indicated previously in the context of identity theory (cf. chapter 3.3), accommodation was understood as the cognitive process where a re-arranging of existing cognitive structures takes place. Assimilation, the first step in Piaget’s learning theories, adapts the perceived outside world into the existing cognitive structures without changing them, “but potentially at the cost of “squeezing” the external perceptions to fit – hence pigeonholing and stereotyping” (Atherton, 2009). New information which fits into the existing structures can, therefore, be assimilated without any “trouble”. On the other hand, accommodation requires adapting new information into existing structures. This requires changing existing structures. This process can be difficult and challenging, similar to problem solving tasks. Perceptions of strangeness and learning how to deal with it, however, appear to proceed similarly to the cognitive perspective of Piaget. The perception of differences might be more related to Piaget’s process of assimilation, while perceptions of strangeness and attempts to deal with strangeness seem to be dominated by accommodation.

When applying attribution theory (Heider, 1958) towards ways of dealing with differences and strangeness, people attribute a particular meaning to perceived differences and strangeness because they try to construct a plausible world around them. Depending on the type or content of the ascribed attribute and the extent of personal relevance/involvement or *subjective meaningfulness* (Haußer, 1995), people deal differently with perceived difference and strangeness. The process of assigning a meaning towards perceived differences and strangeness is often influenced by a kind of cost-benefit analysis, an analysis which moreover can lead to either assimilation or exclusion of the stranger. As long as the stranger seems to be “compatible” to the majority’s political and economical ideas and, consequently, does not disturb the system, he/she is granted access to the dominating system’s resources, if not, access is denied (Gieß-Stüber, 2005a; Gieß-Stüber, 2006).

Existing conceptualizations of *intercultural sensitivity* indicate that sensitivity imply open-mindedness, willingness to deal with strangers, empathy, self-esteem, self-monitoring, open-mindedness, interaction involvement, and non-judgment (Bennet, 1993; Chen & Starosta, 2000; Chen & Starosta, 2002; Chen & Starosta, 1997). Since sensitivity implies *receptivity* and *openness* these concepts are related to each other. Receptivity is a pre-requisite for an

---

open attitude and refers to the indicated dominance of bottom-up process in social perception (cf. chapter 4.3). Openness in the context of strangeness implies receptivity for different and new ideas, behaviors, habits, cultural facets, environments and experiences etc., which are perceived as different from familiar, conventional, traditional or from one's own understandings. Receptivity to new or different ideas implies a subjective meaningfulness which means that an interest needs to be attached to the difference and strangeness and not only towards similarity and familiarity. Yet, noticed familiarities in foreign countries can attract one's attention because it was probably unexpected that such a strange country was not that different and strange from one's own. That insight could consequently be a supportive indicator for a STS. But when pointing out and focusing exclusively on familiarities, the chance of learning something new may become reduced. Therefore, the subjective meaning should also imply a principal interest in difference and strangeness and an understanding that dealing with differences and strangeness is a learning process. This could imply a confidence or optimism that the initial uncertainties can be overcome. Prerequisite of such an understanding is a certain "stability" of one's own identity. Otherwise the uncertainty can easily be perceived as overwhelming and threatening as pointed out earlier. Under such a premise strangeness can become an opportunity to learn something about ourselves and develop our own identity. Perceptions of differences and strangeness can show us new perspectives, meanings and understandings and by dealing with it constructively it can lead to a better understanding of other persons, social groups, and cultures. Conflicting results of interactions with "strangers" can strengthen or possibly weaken our own opinions. When perceiving conflicts in a constructive sense, they can consequently become an incentive to evaluate or re-considering our own point of views (cf. Gieß-Stüber, 2008).

Openness is a supportive condition to learn from each other. Without openness, it becomes difficult to learn something new. Openness is understood as a *liberal* attitude and part of one's own personality. The term *liberal* indicates that an open attitude implies a principal *willingness* to understand and respect other people's behavior, opinions, etc., especially when they appear as strange to one's own or does not make sense in the first place. The *willingness* to meet the "other" open-minded implies a subjective meaningfulness or an interest to hear/listen and consider other people's opinions. This does not mean that it is always possible to understand and accept other people's opinions and behavior because they can also be perceived as conflicting or provocation to one's own norms and standards. But for STS it is important to deal with strangeness further and not only reject or end a potential learning process with prejudice.

Openness is a prerequisite for achieving a better mutual understanding within intercultural contexts. Improvement of a mutual understanding refers to the aforementioned learning process between native and stranger. But “mutual” implies also that we can learn something about ourselves by dealing with strangeness (cf. Reuter, 2002). Dealing with strangeness in an open-minded way can function as a control instance to critically test and re-consider once own, familiar perceptions and so our existing concepts we have about the “others”. Such re-considerations are needed because the concepts about strangers are often based on stereotypes and prejudice. Experiencing alienation can be a key-experience for developing meaningful experiences of strangeness and empathy for how it feels to be a “stranger”. Openness is a prerequisite for achieving meaningful experiences out of an alienation of the supposed “own” familiar arenas<sup>24</sup>. Alienation is related to power differences and refers to the bottom-up perspective where the alienated individual feels at the mercy of dominating power. The idea of alienation as method for the development of a STS is illustrated with an example in chapter 5. However, constructing alienation in teaching approaches can create intensive feelings. Arranging carefully controlled situations which provoke such feelings of uncertainty can help to promote insights of how it feels to be a stranger who perceives him/herself “at the mercy of waves” or not feeling able to control events in life.

Subsequent discussions can help to create awareness that alienation implicates socially constructed power differences. The dominating group defines the standards and rules which are unfamiliar for the newcomer. In a pedagogical setting such as physical education, dealing with alienation is supposed to initiate re-considering (dominating) habits, values, or behavior. This may promote a more differentiated perception of one’s own (social/cultural) background but also that our own cultural background is only one of numerous others. Alienation can be a constructive learning experience in particular for the dominating group members or majority members. Experiencing alienation can help to realize that a possible dominant and self-centric behavior makes it more difficult for the (struggling) new-comer or stranger to get to know the native’s system, structures and resources. The subjectively perceived contrast between the usually perceived certain/familiar grounds and experiencing an uncertain situation by alienat-

---

<sup>24</sup> Alienation is related to identity development and power relations. The German term “Entfremdung” goes back to Karl Marx (Haußer, 1996). Marx refers to the worker’s condition in capitalist society: “the worker is alienated to the extent that the prerogative and means of decision are expropriated by the ruling entrepreneurs.” (Seeman, 1959:784, quoted in Haußer, 1996:112). The psychological meaning is specifically related to Rotter’s locus of control concept and refers to power structures as well (Rotter, 1966). Alienation in a psychological understanding points out the subjective perspective of alienation. Seeman (1959, quoted in Haußer, 1996:112) differentiates between five different meanings of alienation: 1) powerlessness, 2) meaninglessness, 3) normlessness, 4) isolation, and 5) self-estrangement.

ing these familiar grounds is the key element for initiating such an experience. By giving students the task to observe and discuss different coping strategies, constructive ways of handling strangeness can be developed together.

Experiences of alienation are intensive and dominated by overwhelming emotions which easily can lead to frustration and resignation (cf. Noethlichs, 2005a). Such experiences require a rationalization as indicated above. In order to develop or improve STS a broader understanding of social structures and psychological mechanisms which can lead to constructions or perceptions of difference and strangeness are beneficial. Such a rationalization can help to reduce the risk of perceiving strangeness as threatening. A crucial insight for a STS would be that dealing with difference and strangeness is understood as a natural process of personal, social and cultural identity development and consequently seen as a potential learning process. As will be pointed out in chapter 5, it is possible in physical education to promote such insights. Changing tasks and creating situations in physical education can initiate developing meaningful experiences with difference and strangeness. This is a needed prerequisite for creating awareness of different strategies to deal with strangeness and starting point for a training of constructive ways of dealing with differences and strangeness. Consequently, applications of STS are aiming at developing a competence of dealing with differences and strangeness constructively.

Knowledge about sociological structures and psychological mechanisms leading and influencing perceptions of strangeness is supposed to support developing openness and a tolerance to deal with the implicit uncertainty. An awareness of the fact we all underlie the risk of attribution errors when dealing with each other is supposed to help considering other people's behaviour more carefully. Awareness of making attribution errors in the context of strangeness is supposed to support that we consider alternative interpretations of the "other" and being aware of the implicit uncertainty of our interpretations. This awareness may help to prevent pre-judgemental ways of dealing with strangeness and support more differentiated views of the "other".

The implicit uncertainty assigns strangeness a special status. Because of the uncertainty we cannot rely unconditionally on our as universal assumed interpretation background when dealing with strangeness (cf. Schäffter, 1991). Dealing with strangeness constructively means that we reflect more differentiated on our anticipations and interpretations about the stranger and being aware of the fact that our ways of interpretation are not that universal. Sometimes we may realize that unusual or unknown reaction forms can be difficult to handle and accept. But aiming at a better or more rational conflict management one should at least try to consider



---

alternative interpretations and possible circumstances which may have provoked the as strange and possibly provocative perceived behavior. A better conflict management requires therefore a more differentiated view on the circumstances and being less judgmental in the first place (which means more rational). This is a challenging task as actual conflicts between different religious groups illustrate. It is difficult to argue about which belief is right or wrong. Faith is a personal decision and, at least in principle, supposed to be one of free choice. Therefore, one is required to accept and respect other people's faith. Problems linked to different ways of practicing one's religion seem to imply the "real" conflict potential and should be debatable because extreme interpretations can cross the line regarding fundamental human rights (cf. Habermas, 2002).

A differentiation between "sensibility" and "sensitivity" towards strangeness can help to demarcate STS further (cf. Noethlichs 2005a). The subjective meaningfulness was explained previously in the context of identity theory. Subjective meaningfulness refers to perceived importance an object has for an individual. Sensitivity is supposed to represent the more rational facets of subjective meaningfulness towards an "object" of perception whereby sensibility characterizes more the emotional and unconscious facets. Sensibility is further characterized by overwhelming and uncontrolled reactions towards particular stimuli. People's perception is consequently sensitized towards particular stimuli. Perceived differences in skin-tone for instance can be such a stimulus that evokes hatred or violence. Psychological mechanisms such as attribution errors or self-fulfilling prophecies can function as an amplifier in this context. In addition, prejudiced expectations sensitize the process of perception towards specific characteristics, as well and support an overestimation of the difference (cf. Bruner & Postman, 1949).

Associations linked to perceived differences can vary in their meaning, and can lead to positive and/or negative experiences with strangeness. Negative experiences are often amplified through social influences as people have learned to associate strangeness with fear or other unpleasant feelings (cf. Schwarzer, 2000). They thereby seem to abate the required openness to deal constructively with strangeness. Conversely, people who become downtrodden by their daily lives, often tend to escape from their "daily grind" by aspiring after exotic alternatives (cf. Gieß-Stüber, 1999:46f.). Such exoticism is characterized by a mystic zeal about foreign cultures with focus only on the pleasant facets of these cultures. However, exaggerated positive and negative experiences lead to one-sided, biased images of strange cultures. Exoticism, consequently, is more a related to euphoric sensibility than to thoughtful

sensitivity. Xenophobia represents more to opposite extreme and is related to emotions such as anxiety or hatred against strangers.

Based on these theoretical considerations, the following table 1 shows the mentioned facets of STS in overview. In addition, table 1 indicates some pedagogical implications. These pedagogical implications are understood as guidelines for developing didactical conceptions in the future (cf. also Gieß-Stüber, 2005a).

**Table 1: Pedagogical facets for the development of a STS**

Facet of STS	Pedagogical implication
Perception of differences and strangeness	Meeting strangeness Starting point for meaningful experiences of strangeness Differentiated perceptions of strangers Supportive factor for creating an awareness of various perception and attribution patterns Supportive factor for creating a differentiated awareness of difference and strangeness
Openness	Crucial facet of sensitivity concept Dominance of data-driven perception (receptivity) More careful and more differentiated interpretations of the object of perception Confidence that feelings of uncertainty can be overcome
Awareness	Awareness of the implicit uncertainty when dealing with strangeness Knowledge about the social construction of difference and strangeness and different ways of dealing with it Promotion of a more rational conflict management
Experiences of strangeness	Empathy for being perceived as a “stranger” Willingness to accept the implicit uncertainty and deal with it constructively Appreciation of differences and strangeness Development of meaningful experiences with strangeness Development of constructive strategies of dealing with difference and strangeness
Subjective meaningfulness	Central facet of identity theory (Haußer, 1995; <i>ibid</i> , 2007) Willingness or (intrinsic) interest in dealing with difference and strangeness in more differentiated ways Prerequisite for perception and experiences of difference and strangeness
Identity development	Strangeness as incentive for “identity work” (Keupp et al., 2008) Strengthening identity as prerequisite for tolerating the implicit uncertainty by experiencing acknowledgement and affiliation Accommodation of new ideas, meanings and perspectives
Self-relativism	Learning to accept and respect difference or otherness Critical reflections about judgmental tendencies when dealing with difference and strangeness Support of openness and empathy for “otherness” or for being perceived as different
Alienation of the supposed familiarity	Method to achieve crucial insights and understandings More differentiated view of one’s own and other cultures Relativistic view that one’s own culture is only one of various other cultures. Alienation of the supposed familiar helps to create a self-relativistic point of view – in contradiction to self-centric point of views.

---

## **5. Practical implications of sensitivity towards strangeness in the context of Intercultural Movement Education**

STSQ will be applied in intervention projects in the future. The following ideas are background for this thesis and starting point for further research intentions in the future; since it is still an open question if and in how far sport and physical education can contribute to intercultural learning. In theory, there are some indicators which seem to support movement education as a special learning arena for intercultural learning purposes and development of STS. Personal teaching experiences seem also to support these ideas. But they require further scientific investigations in order to prove the supposed potentials and demarcate systematically under which conditions sport and movement education may contribute to intercultural learning. In addition, this chapter shall also illustrate how differences and strangeness can be experienced and how constructive ways of dealing with it can be promoted in the context of physical education.

Practical experiences in the field of sport and movement education in combination with and theoretical reflections around different aspects of strangeness and Intercultural Movement Education are a starting point for considering how far sensitivity towards strangeness can be mediated within the field of sport and physical education. Yet, doubt exists about an automatic relationship between competitive orientated sport and positive interactions and learning processes (Thiele, 1999). The widespread assumption about competitive sports creating harmony and understanding is a neither confirmed nor refuted myth (Gieß-Stüber, 2000b). Generalizing “promises of integration by sport” obstructs a differentiated view and more pedagogical acting in this as important considered field of sport (Gieß-Stüber, 1999:58). Intercultural learning requires perceptions of difference, whereas the rules of competitive sport inevitably exclude cultural peculiarities and differences. Participation is limited to those, who share the same principles. A general (sports-) code omits the individuality. A person functioning within such a code loses its specific characteristics, which make up for its identity. Consequently, strangeness can hardly be experienced in relevant terms, and differences exist only according to the rules of the system.

However, it needs to be considered what possibilities sport and teaching in sport can offer for mutual respect and a more rational conflict management between different persons or peer-groups. Since the body can be considered to be the primary reference for the individual

---

in developing its identity, the (moving) body is equally serving as signal and screen for projections of the own identity (Haußer, 1995). The experiences with the own body as subjective reference seem to prepare important insights into the field. Consequently, physical activities or better movement education is to be seen as promising field for intervention. In addition to the development of physical skills in physical education classes, sports lessons can be arranged such that they mediate social competence. Cooperation, empathy, affiliation and acknowledgement are fundamental pedagogical principles in both social and intercultural learning as mentioned earlier. The following arguments are supposed to stress the specific qualifications of movement education for intervention purposes in the realm of intercultural learning and STS training:

1. Modification of tasks, assumed to be prerequisite for intended learning processes, can be arranged more easily in sports and physical education than in other fields. In addition, essential feedback does not depend primarily on the mediation of another person. Personal experiences with the own body as reference are assumed to support key-understandings.
2. Anthropological concepts stress the importance of physical experiences for individual development. Thus, the body and movement education is not only understood as a medium to learn physical skills but even as a medium to construct and perceive the world (Grupe & Krüger, 1996; Hotz, 1997; Selter, 1996).
3. In addition, it can be expected that many juveniles might be attracted by the domain of sport and physical activity (Noethlichs, 2005b; Noethlichs, 2003; Noethlichs & Schulz, 2007).
4. The relative independence from spoken language is assumed to be further advantages of this field for the given purpose.
5. Strangeness can be experienced through the whole body as personal reference. Insights, generated through subjective experiences of strangeness are then expected to support a better mutual understanding. To experience how it feels to be a stranger serves as basis for subsequent reflections, crucial understandings and insights. Resulting realizations are at least assumed to support more empathic and differentiated perceptions of other persons such as foreigner or immigrants.
6. The own body is understood as a mediator between the individual and the world around. The reference to the own body in identity constructions makes the theoretical construct of identity particular interesting for movement education. Experienc-

---

ing acknowledgement and affiliation are stabilizing elements in identity development. Stable elements are needed to (at least tolerate) the implicit uncertainty of strangeness related situations.

Sport is understood as part of a culture and reflects to some extent structures, norms, rules and values of a respective society (cf. Bröskamp, 1994). Sport games for instance are social interactions which can easily be changed in a way to experience strangeness as will be illustrated in one of the following examples. A basketball game for instance, can be changed or alienated by changing the usual rules. This creates a new situation for all players. Giving the task to find out about what changes have been made creates a situation where the “bottom-up perception” dominates. In such a way alienating usual sport disciplines or introducing unknown sport disciplines, games or unusual movement tasks, can serve as training method for developing strategies in order to manage such “uncertainty-loaded” tasks constructively. Subsequent discussions are necessary in order to initiate reflections about perceptions and to develop and mediate an awareness of different coping strategies and possibly refine them further. The aim is to develop and promote constructive ways of dealing with difference and strangeness.

Unusual sport disciplines and movement arrangements can help to show the manifoldness of movement in relation to different cultural backgrounds. A comparison of different sport and movement cultures may also reveal that there are similarities, common senses or trans-cultural elements beside all supposed differences.

Identity theories indicate the importance of experiencing acknowledgement and affiliation in order to develop a locus of control feeling during childhood (Haußer, 1995; Keupp et al., 2008; Selter, 1996). The locus of control feeling is understood as a crucial capability of dealing constructively with differences and strangeness, and in particular the implicit uncertainty. Not knowing exactly what the outcome of a strangeness related situation might be implies this uncertainty. A subjective feeling of being in control of a situation may promote tolerance of the implicit uncertainty and reduce the risk of perceiving strangeness as a threat to once on identity. Experiences of acknowledgement and affiliation are needed in order to develop such positive experiences and stabilize or balance identity. Physical education allows construction of situations where pupils can experience affiliation and acknowledgment. But this requires specific pedagogical sensitivity because simply competitive orientated P.E. can provoke the opposite and move the focus to constructed differences which further can lead to exclusion and devaluation. Therefore, sensitive pedagogical treatment is required in order to formulate

instructions and create movement tasks/arrangements which allow experiences of affiliation and acknowledgement. The task has to be in focus and not the result as in winning a competition. Movement tasks which require co-operation illustrate how situations easily can be constructed. It is often not even necessary to change much of the classical content of P.E. classes. It is more the pedagogical focus which needs reflection as the later mentioned examples will illustrate.

A mat-race between two heterogeneous structured teams of a P.E. class may illustrate that co-operation is needed in order to win the race. Each team is asked to carry a large and heavy gymnastic mat across the gym. The mat needs to be carried (in different ways) and as a team across the gym. Each team needs to arrange the mat-transport according to his/her strength and height. Experiences show that the team which co-operates best usually wins the race. Therefore, the focus has to be upon the task and co-operation.



**Figure 9: A “mat-race” requires co-operation and constructive considerations of differences (in height, strength etc.) in order to successfully cope with the task.**

Practical conceptions of sensitivity training focus on pedagogical aspects such as experiencing acknowledgement and affiliation, perceiving differences and learning to respect them, enhancing self-esteem and co-operation, development of trust towards other peers despite differences and promoting a trust or belief in one’s own abilities (cf. locus of control feeling).

The following examples shall illustrate more concretely why sport and physical education is assumed to be a promising arena for intercultural learning purposes. The examples also illustrate how strangeness can be deconstructed through the medium movement education.

The examples illustrate that a little modifications of the tasks and sensible pedagogical treatment in the situations offer promising possibilities for relevant learning processes.

*Example 1: Altering the rules of a sport game such as basketball in order to experience strangeness (Noethlichs, 2003; cf. Neuber 1999)*

The first example is supposed to show how a sport game such as basketball can be re-arranged in order to create situations with respect to the experience of strangeness. The basic intervention is to alter one rule of a basketball game without one or two students knowing about it. At first, the respective students need to be separated from the rest of the class. Then, the class decides together what rule is going to be changed. For instance, it is allowed to dribble once only and after dribbling once, the ball needs to be passed over to a team mate. Afterwards, the respective student(s) joins their team without being informed about the changes. The game starts according to the usual set of rules. Each time the student(s) breaks the new rule, the referee stops the game and the opposite team gets the ball and the game continues.

The modification of the rules intends to induce uncertainty in the selected players. The constructed situation shows students, who are familiar and certain with playing basketball, how it feels to be an outsider. Subsequent discussions and reflections are necessary with respect to an improvement of insight and initiation of the learning processes towards a better mutual understanding. Such discussions should circle around trying to describe arising feelings, how the other group members are perceived, and what kinds of strategies were helpful to deal with the situation in a constructive way.

Even though the selected students know that the situation is constructed, the strength of the arising feeling should not be underestimated when trying out such an “experiment”. Some students may feel frustrated and react with resignation. Thus the competence of the teacher is required once again. They must observe the student’s behavior and reactions sensitively. It has to be pointed out and ensured that the exposure and related exclusion will not continue and that the membership and acknowledgement of the group will be restored.

In addition, teaching in an unusual language, including unfamiliar activities into familiar learning sequences or traditional exercises demonstrates how usual sport discipline easily can be re-arranged in order to create situations where *strangeness* can be experienced.

*Example 2: Judo – a constructive way to deal with difference and strangeness (Noethlichs, 2005b; Noethlichs, 2003)*

Responsibility towards a partner is experienced and demonstrated in judo. In addition, the pedagogical principles of judo training define co-operation and respect towards the training partner as compulsory. A translation of the terms “judo” and “*Ji-ta-kyo-ei*” already accentuates its pedagogical relevance:

<i>Ju</i>	the principle of gentleness, yielding, or giving way
<i>Do</i>	way, path, or principle
<i>Judo</i>	the gentle way
<i>Ji-ta-kyo-ei</i>	mutual benefit and welfare

The high incentive of this activity ensures dealing with different partners and pursues a common aim despite individual differences. Teaching attempts showed impressively that pupils are able to train judo in a responsible manner in spite of former obvious aversion to each other (Noethlichs 2003; 2005b).



**Figure 10: A Judo specific greeting ritual creates a polite and concentrated learning atmosphere**

Traditional rituals and rules insure a formal frame for a respectful and polite learning atmosphere. Each training session and each exercise with a new partner starts and ends with specific salutation forms. When learning a new throwing technique for instance, both training partners are responsible for their learning progression (*Ji-ta-kyo-ei*).



In particular, TORI (the one who throws) is responsible for securing the fall of his partner UKE (the one who is being thrown). This rule is especially important at the beginner's level where students have not learned yet to fall in a correct manner. However, UKE gains responsibility as well. According to the before mentioned principle "Ji-ta-hyo-ei", UKE is jointly responsible of TORI's learning success. Consequently, UKE needs to offer TORI a so called "key-situation". A key-situation is an arranged situation which allows applying a reduced number of optimal techniques. In this sense, judo training is characterized by problem orientated learning principles. As a result, TORI needs first to recognize an offered situation, and further apply (react with) a suitable throwing technique. In addition, UKE intervenes TORI'S act with constructive hints if TORI implements the technique incorrectly. The mentioned example indicates that tasks in Judo can be arranged more and more complex. This requires consequently a co-operation of both partners in order to experience mutual learning success – in the sense of *Ji-ta-kyo-ei*.

Furthermore, Judo training requires frequent change of partners with the purpose of learning how to adapt techniques to distinguishing prerequisites such as different height, weight, specific behavior of the opponent etc. In this sense judo training is problem orientated in two ways: (a) students have to learn how to provoke, or identify key-situations in order to prepare a suitable throwing technique, and (b) they need to learn a flexible use of techniques with respect to individual differences of new partners.



**Figure 11: To throw each other requires responsibility towards the partner**



**Figure 12: "The task orientation and fun made it possible to responsibly "fight" together despite of differences**

Personal teaching experiences with the subject judo showed that students learn to develop strategies to deal with individual differences in a constructive manner. For instance, students learn to perceive that all students differ more or less from each other. This insight should help to perceive differences as more natural, and may prevent having prejudice against perceived differences. Overweight for instance can gain a more positive connotation in the context of judo. A clever implementation of body weight can be an advantage for the success of a conducted technique, especially on the ground.



**Figure 13: A clever implementation of one's own body-weight can become an advantage in Judo**

Nevertheless, an additional aspect makes judo interesting for sensitivity training purposes in the context of strangeness. Identity development is among others needed in order to tolerate uncertainty and moreover learn to deal constructively with uncertainty. As a result, it is assumed that judo skills can improve self-confidence. As a case in point, judo training is focusing on a pallet of different mental and physical skills such as challenging anticipation tasks, flexibility, and body control. In addition to those skills, judo is a “gentle” but effective self-defense sport. In summary it is assumed that judo may affect positive on one's own body image, locus of control, and consequently self-esteem as well. If judo is able to improve such self-concepts of individuals, it could help to tolerate uncertainty better, and moreover support constructive ways of dealing with strangeness.

The distinguishing cultural roots of judo can be used to offer opportunity to learn more about a different culture – from the perspective of a western society. Judo has a long tradition in Japan, and can function as a starting point for learning about cultural facets of Japan.

Judo is still an unusual discipline within physical education classes. But constructing unusual situations and exercising tasks are ways of creating situations which are new or strange to most of the students. By constructing strange situations, the attention is more drawn upon the task instead of focusing on the difference such as different look, style, and physical / mo-

---

tor skills. Consequently, elements of adventure sport and special arrangements within usual physical education are assumed to provoke situation where students get an opportunity to experience acknowledgement and inclusion.

Meeting a stranger, strange habits or strange situations make us usually suspicious because strangeness comes along with feelings of uncertainty which can be perceived as threatening to one's own identity. It is often the easiest way to react with prejudice and stereotyping. Dealing with strangeness constructively means to take a kind of risk because we cannot be certain about the outcome of a situation. But it is important to promote a principal optimistic feeling, confidence or belief that the perceived uncertainty will be temporary and that the situation is manageable. Development of trust or belief into one's own ability to manage the outcome of the situation may help to tolerate the implicit uncertainty and probably makes it easier to deal more constructively with differences and strangeness. Strangeness may then not be perceived as a direct threat to one's own personality.

The belief or trust in one self's ability is one perspective. It is also helpful to develop trust towards the different and strange perceived person. Often, we are not that different as we may assume. Strangeness is a construction and can be deconstructed by dealing with it. By trying to understand the other person, we learn something new or we learn about hidden similarities as well. We often assign differences and strangeness based on prejudice and stereotypes. Without dealing constructively with strangeness, possible misunderstandings cannot be revealed. Promoting a locus of control feeling in combination to a development of a mutual trust despite supposed differences and strangeness, may help to reduce suspiciousness towards the stranger and help to meet the stranger more open-minded.

When teaching heterogeneous P.E. classes (7<sup>th</sup> grade) I could often observe conflicts between boys and girls which were more problematical than between pupils with different cultural backgrounds. In particular the physical differences between boys and girls often provoked that the girls avoided physical activities together with the boys. By changing the focus on coping with the task in contrast to competition and result orientated P.E., the girls became a chance for receiving acknowledgement from the boys because the new and unusual arrangements required different coping strategies in order to manage the tasks. This uncertainty to manage was relevant for both the boys and the girls. In some cases, the girls were more willing to take a risk and try to manage the task compared to the boys (Noethlichs, 2001). The following pictures shall illustrate some the pedagogical ideas such as development of trust;

taking a risk; belief in one's own abilities; task orientation (in contrast to competitive and result orientated P.E.).



**Figure 14: Climbing (co-operation, trust and responsibility)**



**Figure 15: Free fall from a Vaulting Box (trust and coping)**



**Figure 16: Taking a risk - jumping over a "glacier rift"**





**Figure 17: Take a risk and jump!**



## **PART II: EMPIRICAL INVESTIGATIONS**





---

## 6. Methodological procedure of developing the STSQ

Before presenting the methodological steps and empirical results of this thesis, I want to start with some methodological considerations in order to point out the scientific baseline for the underlying methodological idea of my thesis.

I will introduce my methodological considerations with an analogy in the context of strangeness. Heider (1958) introduced his attribution theory with a picture of “the naive scientist”. Following this idea, we as human beings are naive scientists in our daily life because we continuously try to explain other people’s behavior and try to construct reasonable pictures of the world around us. Our past experiences help us within this process of searching for meaning and reasons. There is an analogy between our daily life experiences and scientific experiences. Scientific theories represent (scientific) experiences in abstract forms. Our experiences or personal theories in daily life contain ideas about other persons, social groups or cultures. In the case of strangeness, our subjective “theories” are quite uncertain. In order to understand the “other” better and build a more differentiated picture of the other person, social group or culture we need to be open-minded as described in previous chapters. Back to the scientific view, entering a new research field would suggest using more explorative methods in the beginning which also require a kind of open-mindedness. This explorative character is needed to increase our understanding of a new phenomenon systematically and possibly develop a new or adapt an existing theory according to the increase of knowledge and understanding.

We gain our daily life experiences through reflections and interpretations of our daily impressions and perception within different situations and contexts. Comparisons between actual perceptions with past experiences lead to new experiences by accommodation of new knowledge and perceptions. The mechanisms I tried to point out here are similar to scientific mechanisms. In difference to scientific theories, our daily life experiences they are often more uncertain because they underlie more naive evaluation strategies. Our daily life evaluations are more intuitive, stronger affected by emotions and based on stereotypes and prejudice. Scientific evaluations should usually be more controlled according to certain scientific criteria which basically are concerned to reduce, control or estimate potential error influence. The aim is to reduce uncertainty. In daily life, we are usually not that aware of these uncertainties of our interpretations and the risks of possible error.

---

This short analogy showed that dealing with strangeness is a fundamental concept which even is related to the theory of science and research methods. However, the following considerations are more concerned to mediate the methodological idea underlying my procedure of developing a concept-based measuring instrument in the context of the more general concept of Intercultural Movement Education.

The following two sub-chapters (6.1 and 6.2) are focusing on the more general methodological background idea. I start with more general considerations of the underlying methodological idea in chapter 6.1 and apply these to my suggested strategy and statistical procedure for developing the STSQ (chapter 6.2). The next two chapters (6.3 and 6.4) represent the bridge between my theoretical and empirical part of my thesis. I placed the operational model (chapter 6.3) and the first structured version of my STSQ (chapter 6.4) in the methodological part. The more or less deductively derived first version of the STSQ is then starting point for my conducted empirical studies with the aim to develop and improve the STSQ successively (chapter 6.6 – 6.9) whereby chapter 6.5 was more a helping procedure to construct the first STSQ version.

### 6.1. Methodological background idea of developing a concept-based measuring instrument

The topic of this thesis points out that the STSQ was concept-based developed. This refers basically to the concept of IME. The conceptual character implicates that IME is in the first place a theoretical idea which cannot be observed directly. IME is a specific concept in the field of sport and is directed towards practice (application orientated approach).

To support the suggested concept-based development of a measuring instrument in the realm of IME, it seems helpful to point out the relationship between theoretical ideas or concepts and empirical observations. A few basic considerations around this relationship appeared to be natural when trying to develop a concept-based measuring instrument (cf. Hagtvet, 2006).

Scientific theories are rational constructions of our perceived “reality” (Kriz et al., 1996). They are assumed to illustrate reduced pictures of the world around us. Discovered structures and regularities make our environment understandable, and enable scientific predictions. Empirical theories are abstract and compressed systems of scientific statements, in the way that several empirical results may come out of minimal theoretical assumptions (Gadene, 1994b). Because of the unimaginable complexity we have to deal with, reduction is unavoidable. This

---

reduction is needed so the theory's practical function can give orientation in our complex world (Kriz et al., 1996).

The history of science has shown that theories and scientific knowledge is just temporary. A theory keeps its legitimacy until empirical results demonstrate the opposite. Following critical rationalism, a theory has to be given up if a contradictory observation can be found. But from a more pragmatic perspective, such a rigid demand is not necessarily required. Parts of the theory might continuously have some explanation force. The following example shall illustrate this: Einstein's general theory of relativity, which is in comparison to the classical gravitation theory of Newton a more abstract and general theory of gravity. It explains the phenomenon of gravitation more precisely and moreover corresponds in prediction of observable effects. Subsequently, Newton's theory can be deduced from Einstein's general gravitation theory, which turns this classical theory into a special case of the more general one. As a result, Einstein's theory is, compared to classical theories, more valuable because of its greater reliability and broader validity. In line with strict critical rationalism, Newton's theory must consequently be rejected, even though the consequences of classical physics are still useful in our daily lives. A practicable point of view would therefore suggest to keep a theory as complex as necessary, but as simple as possible. We do not even need having heard about "The Theory of Relativity" in order to measure our weight, but if we want to explain "weight" or the physical dimension "mass", only Einstein's theory delivers plausible explanations.

The above mentioned example of Einstein and Newton's theories should also illustrate that the research question or the research problem determines which theoretical perspective may bring light to the field one investigates. In particular, pedagogical research is principally application orientated because the main interest is to reflect critically and support learning and development of individuals in relation to the claims of society. In this sense, pedagogical research is not only interested in understanding or explaining different kinds of phenomena. Pedagogical efforts are also directed at change or intervention in order to support learning and development of human being in a responsible manner.

The latent character of a theory means in other terms that a theory implies more information than can be observed directly (Gadenne, 1994b). The typical circular reasoning "intelligence is exactly what the IQ-test measures" illustrates the problem to be pointed out here. If only confirmed empirical observations would fill a theory with content, then the theory would not be more than a collection of empirical observations and therefore only be of descriptive character, and finally would have no explanation force (Westermann, 2000).

Empirically based efforts are in general directed to provide theoretical assumptions with evidence. In this context, measuring is understood as a process of comparing empirical observations with theoretical assumptions. “Measuring” means the assignment of numbers towards observable events, expressions or variables. The relations between the numbers are then assumed to represent the empirical character/form of the observed phenomenon. This empirical representation of the theoretical ideas can be imagined as a kind of picturing process. An ideal (hypothetical) aim of this picturing process would be the creation of “isomorphic images” of the empirical observations. This consequently means that isomorphic images are unequivocal in both directions. But this idea is only achievable theoretically such as in mathematical theories. In practice, only “homomorphous” images are feasible. For instance, with the aim to arrange students according their mathematic performance by assigning grades from 1 to 6 towards each student, it is possible to link unequivocal one grade towards each student, but it is not necessarily possible to assign all grades to all of the students (cf. Kriz et al., 1996:90f.; Westermann, 2000:137ff.).

The crucial idea of measuring is threatened by errors. The process of measuring implies a kind of two-way transformation from theoretical terms to numbers first, and then, after statistical editing of the collected data, back from resulting numbers to interpretations regarding theoretical assumptions. These transformations are always influenced by errors such as incorrect assumptions about empirical proportions or misleading interpretations of the data. The more complex the data material the more differentiated the required statistical tools to edit the collected data. Consequently the risk can be larger for error influences as the arising difficulties of interpreting oblique rotated factor solutions or the multiple interactions of multivariate variance analyses (cf. Erdmann, 1988).

Nevertheless, an already mentioned problem regarding the empirical examination of theoretical constructs is the difficulty that the major assumptions of theories are not testable directly. Only through the derivation of additional assumptions or hypothesis (“Zusatzannahmen”) a theoretical construct can be operationalized and proved indirectly (Gadonne, 1994b). Applied to the topic of this thesis, additional hypotheses must be derived out of the fundamental assumption of the construct of strangeness such as how people may react when perceiving another person as strange, how people may feel in such situations, or what they might think when experiencing strangeness etc. Interpretations of the resulting observable feedback may give sufficient information about the empirical representation of the underlying construct. These procedures of operationalization only makes sense when a) the measured characteristic is assumed to show at least some stability over time which means that the observation exceed

---

a single case observation of a defined population, and b) it can be reasonably expected that the amount of observable indicators represent the theoretical construct in a sufficient way. Such operational procedures are required in order to achieve reasonable results when intending to measure the empirical representation of a theoretical construct (ibid. 1994b).

There are further problems linked to the problem of operationalization. If empirical testing of theoretical expectations lead to contradictory results, researchers often tend to assign the problem of such unexpected results towards errors in operationalization or population bias (Erdmann, 1988). The arising uncertainty during the search for possible error influences may attribute the errors towards external circumstances of the studies instead of considering the possibility that the theoretical assumptions might be responsible for inconsistent results (cf. Popper, 2002). It seems natural that it is difficult to give up theoretical assumptions which often represent the major incentive of a long research process and is consequently the baseline for the whole research. The often perceived pressure of publishing only statistically significant research results may support an attitude to look more for verifications instead of trying to treat theoretical assumptions critically or in the terms of Popper (2002) trying to falsify theoretical assumptions. It should additionally be pointed out that the knowledge of “wrong” theoretical assumptions may lead to important and meaningful knowledge as well. But instead of revising or giving up the “wrong” theoretical assumptions, researchers more often try to justify the theory with non-testable assumptions in order to “immunize” the theory against contradictory results. This strategy leads to the fact that the theory becomes more and more difficult to test empirically.

A practical point of view on scientific efforts suggests that a number of linked “simpler” constructed or demarcated research studies may bring light into a more complex area step by step. In this way, each sub-study can consequently be carried out under more controlled circumstances (Erdmann, 1988). Laboratory experiments in social sciences are usually expensive, difficult to arrange, and often underlie higher ethical restrictions. When dealing with a new and complex research field, it may be more fruitful and economic to clearly demarcate the research field into a number of smaller (easier to handle) and theoretically related studies (cf. Erdmann, 1996). An important demand is therefore to demarcate the research field distinctively, and keep the connections between each study. Careful theoretical groundwork is necessary because the theoretical concept creates the links between the studies. A logic consequence of such a design is that a theory-driven procedure requires lower methodological claims and standards (Erdmann, 1988).

My STSQ is in the first place developed deductively. This concept-based development should make predictions of theoretical assumptions possible. According to Erdmann (1988:279) the theory-based procedure does not call for large sample sizes and high psychometric standards. Data require lower scale qualities and does not oblige general standardization or norms of the measuring instrument, because of the lesser necessity for *ex post facto* interpretations. Mean values need not be interpreted *ex post facto* and be analyzed further with mathematically problematical procedures (e.g. questionable distributions of the data material, required data prerequisites for specific statistical methods, or unbalanced ratio of variables and/or persons). In extreme cases only nominal scale quality is required (Erdmann, 1979). Applied to the STSQ would this mean that the simplest way the STSQ is supposed to measure is to check either a criteria (item) is achieved or not.

Regarding the aim of treating a broader research topic with a number of linked small-sample studies, the consistency between the single results of theoretically unidirectional studies is more important than the level of significance or explained parts of variance (Erdmann, 1988:179).

Especially during theory development and operationalization the exhibited procedure seems helpful. Careful and sensitive proceeding, especially in developing processes, seems to prevent for jumping to unjustified conclusions or generalizations. The validity of a theoretical concept can moreover be narrowed down successively. A main advantage of this procedure is that problems of operationalization that can be met in progress which large single sample designs may not be able to catch up.

## 6.2. Design and statistical procedures for the development of the STSQ-items and STSQ-scores

Based on my previous argumentation, it was assumed that this theory-based procedure of item pool development does not call for one large-scale single sample analysis and sophisticated multivariate statistics but rather for a number of theoretically related smaller studies (Bortz & Lienert, 2003; Erdmann, 1988). The purpose was to develop the items of the STSQ in progress according to the increase of knowledge of each single study. Instead of reducing a large item pool by excluding unfitting items, it was intended to adapt the items in the progress based on the results of each data analyses. The relationship between each data collections was given by the adaptations of the STSQ from study to study. Provisional versions of the STSQ were tested with a sample of examinees and the STSQ items and/or summarizing item scores

were adapted according to the results. Then, new data was collected with the adapted version of the STSQ in order to test the result of previous adaptations.

Most of the samples were convenient samples collected in different contexts and designs such as field-studies and intervention including samples from different countries such as Norway, Poland, Germany, France and the Czech Republic. The samples were gathered from high school and university students because they might later become a major goal group for applications of the instrument. Some of the data collections were conducted in the context of another project. This EU-project was designed and conducted by a research group from the University of Freiburg in Germany. Even though my STSQ was still in development, I expected that an application of the STSQ within this theoretically related EU-project could generate additional knowledge about construct validity of the STSQ.

The key-element of the process of validation<sup>25</sup> is characterized by combining “scientific inquiry with rational argument” (Messick, 1995:742). Since the STSQ was in a developmental process the focus of my investigations was directed at exploring hints on construct validity in order to adapt single items successively. I preferred the term “validity hints” in order to point out the developmental character of investigations in distinction strict testing construct validity of the whole instrument. I intended to investigate how far respondents seemed to understand the items and compare that to the theoretically intended meaning of the items. The ways how different items of the STSQ were related with each other was supposed to reveal how the items seemed to be understood by a group of examinees.

Main statistical tools for those investigations were based on non-parametric correlations. Applications of non-parametric procedures require lower or weaker statistical prerequisites as parametric procedures. Parametric tests are only valid and meaningful under specific conditions such as normal distribution. These conditions must be examined in each case they are applied. These statistical tests to test required prerequisites are in most cases also based on parametric procedures and require the prerequisites to be tested (Bortz & Lienert, 2003:59). This makes statistical investigations of mathematical prerequisites a questionable procedure.

In the context of my thesis, the characteristics of the population regarding a STS are unknown. Non-parametric procedures can be considered as more conservative methods in the

---

<sup>25</sup> The conducted studies of developing the STSQ items are founded in classical test theory (Crocker & Algina, 1986; cf. Lienert & Raatz, 1998; Rost, 2004). The validation of a measuring instrument is understood as a process of collecting evidence to support the type of inferences to be drawn from a measurement score (Crocker & Algina, 1986). Consequently, validity testing is generally directed at investigating the property of inferences (Kleven, 2008; Lienert & Raatz, 1998; Rehm & Strack, 1994; cf. Shadish, Cook, & Campbell, 2002).

---

context of development a new measuring instrument. More conservative treatment of data should reduce the risk of inadmissible interpretations and randomly discovered significant results.

The decision between parametric or non-parametric procedures is usually related to the scale quality of the instrument. Parametric test such as t-test require interval scale level such as Likert scale type. Lower scale qualities suggest using non-parametric tests (Bortz, 2005). In particular, intended applications of the STSQ only needed to differentiate between different groups. For such an aim only rank order is needed. It is often assumed that the distances between each scale point of a Likert-scale are equal. In how far applications of a Likert-scale can follow such a claim is discussable and needs to be seen in relation to the underlying theoretical construct (cf. Rost, 2004:50). At least, it can be assumed that a Likert-scale is ordinal. Ignoring a potential discrete nature of a Likert scale can lead to inferential errors (cf. Clason & Dormody, 1994). However, non-parametric test are robust procedure that does not rely on the assumption that instrument provides precise distances on interval scales (cf. Van de Vliert & Kabanoff, 1990).

It can be stated that under certain statistical conditions non-parametric tests are less efficient than parametric tests (Bortz & Lienert, 2003). If the data set would fit required conditions perfectly (in theory), would an “equivalent” non-parametric test applied to the same data set require a larger sample size as the parametric procedure would require for detecting significant results. This comparison (regarding the efficiency of parametric vs. non-parametric tests) is quite theoretical, and probably irrelevant for practical applications. It seems irrational to assume that small sample sizes achieve perfect prerequisites for applying parametric procedures. But if required prerequisites are not given, applications of parametric test can become inappropriate and the aforementioned comparison can even become inverted; the parametric test would become less efficient than the non-parametric test (ibid., 2003).

Consequently, non-parametric procedures become also relevant for applications when dealing with smaller sample sizes and when the required prerequisites for parametric test are not given or questionable. It is quite certain that a discovered significant non-parametric test result with a small sample size is usually based on relative large effect size (cf. Bortz & Lienert, 2003). This is maybe the reason that (equivalent) effect size estimates are still lacking for non-parametric tests. Therefore, effect size is usually not reported within my results.

I tried to argue for the use of non-parametric procedures in the context of developing a new measuring instrument. However, non-parametric methods do not provide comparable statistics such as multivariate analyses. Consequently, I considered explorative factor analyses as addi-



tional procedure but being aware of the potential error risks when not meeting the required prerequisites. Results were therefore interpreted carefully because most of my analyses are based on lower sample sizes. In addition to the heterogeneity of the STSQ, results of factor analyses were only used to gain validity hints (cf. Bühner, 2004). Applying EFA should consequently not undermine my previous argumentation. It was more understood as supplement of a statistical procedure which appeared to be helpful and which is not provided in a comparable way by non-parametric methods.

The construct of STS was characterized as a multi-faceted model because the facets or categories were preliminary intended to structure the item pool and is not understood as factor model (cf. Figure 18). Even though explorative factor analysis was considered as a helpful statistical method in order to gain some first empirical feedback about the empirical structure and relationships of single items and item scores. Explorative factor analysis (EFA) helped to develop score keys of the STSQ. The explored structures in the relationships between variables were expected to help summarizing items and find representative labels. Consequently, EFA was applied as a structure detection method to see how the intended structure was reflected in the results of an explorative factor analysis. EFA was in this sense be used for discovering hints on construct validity as well (Krampen, 1981; cf. Lienert & Raatz, 1996)<sup>26</sup>. When data material did not provide optimal circumstances for an application of EFA<sup>27</sup>, it was important to carefully interpret resulting factor solutions in particular when applying EFA with smaller sample sizes and heterogeneous samples. The strict theoretical-based procedure was supposed to be one aspect to meet the mentioned threats and help considering the resulting factor structures. Kleven (2008) expresses the underlying methodological idea as follows:

“In addition to the rational assessment, empirical data may play an important role in construct validation, as first shown by Cronbach and Meehl (1955) for psychological tests. In quantitative research various correlational and factor analysis methods may help us to evaluate whether the operationalized constructs «behave» as we for theoretical reasons would expect the constructs to behave. (Kleven, 2008:225)

---

<sup>26</sup> Krampen (1981) uses the term “factor validity”

<sup>27</sup> According to Bühner (2004) factor analysis (principal axis factoring [PAF] method) does not require (multivariate) normal distribution and interval scale level in principle, but when both aspects are given this would produce optimal circumstances for conducting an explorative factor analysis. The more linear the relationship between items is, the more stable factors can be expected. Different aspects can threaten this linearity between items such as dependency on type of sample/population, sample size and number of items (Bortz & Döring, 2006; Bühner, 2006).

Further aspects needed to be considered regarding the application of EFA within the context of this thesis. Since my research on the STSQ was at its beginning, reliability of the items was unknown as well. According to Bühner (2004), it is therefore not advisable to use the criterion of “Eigenvalue”  $> 1$  as extraction method when applying an EFA. As a result, the graphical Scree test method<sup>28</sup> seemed to be more appropriate at the beginning. Independent from the selected methods, it was unreasonable to extract a number of factors which is not interpretable in a plausible way (Bortz, 2005). In order to make interpretations of “factors” easier, a *Varimax rotation method*<sup>29</sup> was selected.

For validation purposes, it was important that resulting factor solutions with different samples were interpretable of the underlying construct. In this sense, the plausibility and coherence of resulting factor solutions with different samples would be a supportive indicator for the validity.

Because of the heterogeneous structure of the STSQ and the lower scale qualities it was not expected to achieve high internal consistency of the STSQ scales. According to Lienert and Raatz (1998) a reliability coefficient between  $r=0,5-0,7$  is sufficient in order to differentiate between groups. However, intervention designs are one potential application of the STSQ in the future. The stability or re-test reliability was consequently a more relevant aspect to estimate for applications of the STSQ in the future than to Cronbach’s alpha. Therefore, the stability over time of different STSQ scores was investigated as well. The resulting knowledge was used to adapt the STSQ further.

#### *General conditions during data collection:*

Short and precise instructions made the STSQ self-explaining. Data collection could therefore be conducted relatively independent with respect to potential investigators previous knowledge. However, being in a process of developing a measuring instrument, my personal presence during data collection gained helpful information about the questionnaires applicability. Direct comments from respondents gave useful information in order to improve the measuring instrument’s understandability and practicability. Personal discussions with res-

---

<sup>28</sup> With the purpose of finding a number of extractions by using a Scree plot, one follows the run of the decreasing “Eigenvalue” curve within a Scree plot from the left to the right side, and searches for a significant decrease of the “Eigenvalue” (obvious bend). According to Bortz (1994 quoted in Bühner, 2004:162) the number of factors is determined by the number of “Eigenvalues” before the “Eigenvalue” curve follows a sharp bend.

<sup>29</sup> Varimax Rotation maximizes the squared factor loadings within a factor instead of maximizing the factor loadings of each item (Bortz, 2005; Bühner, 2004). This rotation method points out higher loadings and reduces at the same time loadings on other factors which consequently should make interpretations easier.

---

pondents after answering the STSQ showed that the questionnaire was usually perceived as interesting. The different types of approaches of the STSQ helped to keep the motivation for filling in the whole STSQ.

The smaller designed data collections made it easier to collect data under more controlled circumstances. Usually all studies were carried out in school or university in a quiet atmosphere. During all investigations data was collected anonymously and treated as confidential. Before respondents started filling in the STSQ they were informed about the aim of the measurement but only as much as they needed to fill in the STSQ correctly. When the STSQ was finished, a debriefing procedure was conducted which should give respondents an opportunity to ask further questions or to comment on the questionnaire. They were also informed more comprehensive about the aim of the research project they participated in.

In order to avoid a possible risk of “experimenter’s bias”, a script was worked out and used in all studies with the aim to control the given instructions and information. All tasks of the STSQ were explained in the instructions of the questionnaire so there was usually no further information needed. The respondents experienced my personal presence as positive, and seemed to reduce the drop-out-rate. It might further have supported the willingness of the participants to fill out the STSQ seriously - in comparison to more impersonal ways of data collecting methods such as per mail or internet-based data collection procedures.

The following table (2) gives an overview on each data collection (sub-study) within my project, their particular research aim, design, and further characteristics in summary. Each sub-study will be reported in a chronological order<sup>30</sup> in order to point out the successive character or the developmental process. Resulting consequences for adaptations of the STSQ are summarized at the end of each reported sub-study. The adaptations of respective STSQ items were starting point for new analyses with a new data set.

---

<sup>30</sup> except the first of two data collection within the EU-project. I report the result together in one chapter because they were conducted within the same project.

Table 2: Empirical steps of developing the STSQ successively

Aim	Chapter	Design	Sample	Method/ Statistical design	STSQ version <sup>31</sup>
<b>1. Development of a first item pool/ initial version of the STSQ (operationalization)</b> <ul style="list-style-type: none"> <li>• Exploration</li> <li>• Inspirations for item formulations, construction of relevant situations, and collection of relevant picture items.</li> <li>• Linguistic review</li> </ul>	Pre-studies (Chapter 6.5)	a) Group discussions  b) Writing tasks  c) Discussion	a) High school students (N=43)  b) Norwegian high school students Norwegian high school students with immigration background (N=35)  c) Experts within the field of IME (N=3)	Qualitative analysis: <ul style="list-style-type: none"> <li>• Writing task analysis, protocol analysis of group discussion</li> </ul>	
<b>2. Collecting hints for item adaptations/corrections of item formulations, instructions, score keys</b> <p>Sub-aims:</p> <ul style="list-style-type: none"> <li>• Development of score keys / STSQ score construction</li> <li>• General practicability</li> <li>• Validity hints</li> </ul>	Pilot study (Chapter 6.6)	<u>Single measurement design:</u> (STSQ v.1)  Single measurement (STSQ v.1)	Sport students (Germany), N=86  International students (ISS, Oslo), N=24	<ul style="list-style-type: none"> <li>• Correlations</li> <li>• Explorative factor analysis</li> <li>• Internal consistency/ Cronbach's alpha</li> </ul>	v.1
<b>3. Evaluation of effects of previous item changes</b> <p>Sub-aims:</p> <ul style="list-style-type: none"> <li>• Stability over time/ test-re-test reliability</li> <li>• Collection new hints related to construct validity in order to</li> </ul>	Repeated measurement design (Chapter 6.7)	<u>Repeated measurement design (without intervention):</u>  1 <sup>st</sup> measurement (STSQ v.2) Ca. four weeks space between measurement one and two 2 <sup>nd</sup> measurement (STSQ v.4)	1. Norwegian sport students, N=69 2. Norwegian sport students, N=51 → Resulting sample size of	<ul style="list-style-type: none"> <li>• Correlations</li> </ul>	First measurement: v.2  Second measurement: v.4

<sup>31</sup> A new version indicates that an adapted/revised version of STSQ is applied.

develop the item pool further

*identical persons for test-re-test reliability: N=35*

<p><b>4. Evaluation of effects of previous item changes</b></p> <p>Sub-aims:</p> <ul style="list-style-type: none"> <li>• Age differences and intellectual limitations</li> <li>• Explore further validity hints</li> </ul>	<p>Application in a Norwegian school (Chapter 6.8)</p> <p><i>Single measurement design with larger sample size:</i></p> <p>Norwegian gymnasium (STSQ v.5)</p>	<p>N<sub>total</sub> = 224</p> <p>N<sub>1st class</sub> = 80</p> <p>N<sub>2nd class</sub> = 76</p> <p>N<sub>3rd class</sub> = 68</p>	<p>v.5</p> <ul style="list-style-type: none"> <li>• Correlations</li> <li>• Mann-Whitney U-test (comparison of constructed extreme groups)</li> <li>• Explorative factor analysis</li> </ul>
<p><b>5. Evaluation of effects of previous item revisions</b></p> <p>Sub-aims:</p> <ul style="list-style-type: none"> <li>• Application of the STSQ within an intervention design</li> <li>◦ Hints on construct validity</li> <li>◦ Evaluation of measuring stability</li> </ul>	<p>Intervention design (Chapter 6.9)</p> <p><i>Repeated measurement – control group design:</i></p> <p>Pre, post, follow up measurement with one parallel control group measurement</p>	<p>Intervention sample (2006): N=33</p> <p>Control group sample: N=49</p> <p>Participating countries: GER, CZ, POL, FRA</p> <p>Intervention sample (2007): N=24</p> <p>Control group sample: N=11</p> <p>Participating countries: GER, CZ, POL, FRA</p>	<p>v.3</p> <ul style="list-style-type: none"> <li>• Correlations</li> <li>• Wilcoxon signed rank test (dependent group comparison, repeated measurements of the IG)</li> <li>• Mann-Whitney U-test (comparison of independent samples, IG vs. CG)</li> </ul>

### 6.3. Operational model of STS

Within this chapter I will introduce my developed operational model of STS. This model is based on previous theoretical considerations. The model was basically supposed to be a guideline for structuring the developed questionnaire from a theoretical point of view. The circles around these facets indicate the theoretical or latent character of the respective facets. The bottom of the model also shows the developed observable indicators (items of the first version of the STSQ). After the pilot study (chapter 6.6) each item could be assigned towards a latent variable as a starting point for further analyses. The indicated relationships between the different variables are based on theoretical considerations and assumptions. The (explorative) pilot study helped to summarize and label the observable indicators and assign them to theoretical facets of STS.

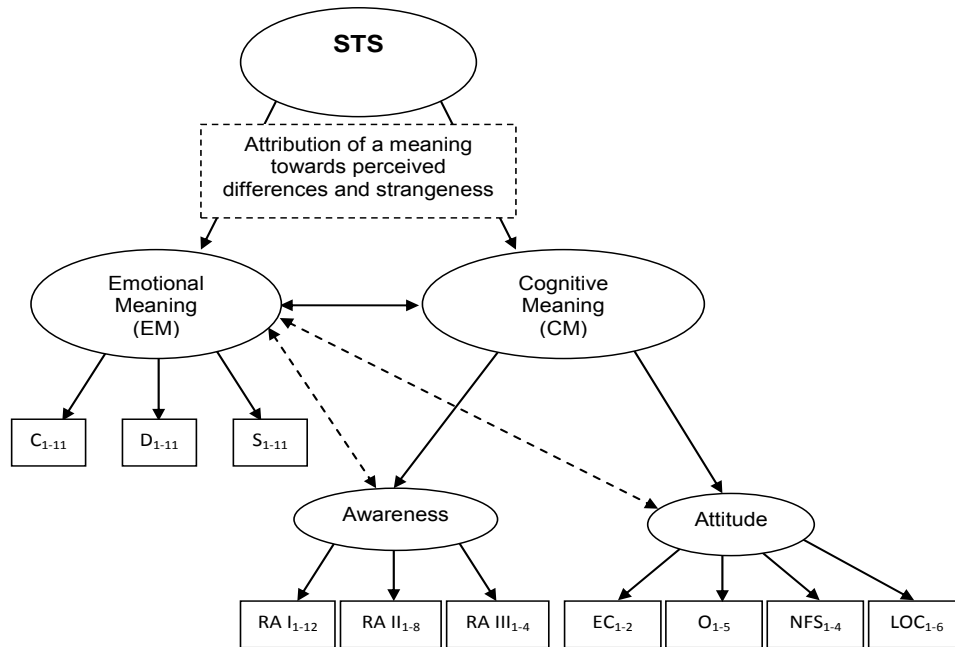
As shown in Figure 18 STS builds the head of the model. Strangeness itself was previously introduced as a social dimension, and all dimensions and facets are related to this dimension indirectly. The theoretical perspectives of subjective meaningfulness towards perceived difference and strangeness and ways people deal with difference and strangeness are theoretical guidelines for the model and the construction of observable indicators for measuring STS.

The subjective meaningfulness was introduced in the theoretical part of my thesis as a starting point for the concept of sensitivity. The subjective meaningfulness influences our behaviour and consequently our ways of dealing with difference and strangeness. Also according to my theoretical considerations, subjective meaningfulness is structured in two dimensions:

1. An emotional meaning (EM)
2. An cognitive meaning (CM)

Both dimensions are related to each other but are separated here for analytical purposes particularly in order to develop a structured item pool.

In addition, the cognitive dimension is separated in two further sub-facets, awareness and related attitudes because respective items are here more referring to cognitive aspects than to emotional dimensions even though emotional and cognitive interact within a strangeness related situation.



**Figure 18: Operational model of the STS (STS: Sensitivity towards Strangeness, EM: Emotional Meaning, CM: Cognitive/rational meaning, C: Closeness, D: Difference/Dissimilarity, S: Sympathy, RA I-III: Rational Attribution, EC: Ego centrism, O: Openness, NFS: Need for Security, LOC: Loss of control)**

### 6.3.1. Emotional meaning (EM)

Besides cognitive evaluations and judgments of “strangers” emotions influence the way people deal with strangeness. Already when people generate a first impression of a person they meet, they perceive the other person in a more or less pleasant way – dependent on expectations, associations, related experiences etc. (Fiske et al., 1991; Gudykunst, 1987).

When intending to measure an attributed emotional meaning in the sense of the underlying conceptualization, *closeness*, *difference* and *sympathy* are expected to represent relevant dimensions for an operationalization of an emotional evaluation of perceived differences (EM). At first, the dimension of “closeness” is used to illustrate a social relationship between people or groups. “Closeness” is supposed to indicate how central or relevant an object of perception is considered by a perceiver (cf. Bogardus, 1925; cf. Simmel, 1968).

The dimension of “difference” is understood in a similar way as the above mentioned dimension of closeness. People assign differences in order to demarcate themselves from other people (cf. chapter 3). A demonstration of being different can be interpreted as an attribution of meaning (subjective meaningfulness). This demarcation from another significant (meaningful) other is made because it is meaningful to the person. On the other hand, an obvious demonstration of similarities with another person can indicate that the person is a significant other but in the sense of social affiliation. Yet, neither indicating to be different nor to be similar shows that the object of perception is of lesser personal relevance.

An attribution of “sympathy” represents the third and last dimension with respect to emotional meaning linked to perceived difference and strangeness. This dimension is intended to show if the object of perception is perceived in a more pleasant (positive) or more unpleasant (negative) way.

The attributed personal meaning indicates a “willingness” to deal with the object of perception; the attribution of sympathy may moreover indicate the emotional direction of meaning. A more positive perceived person or situation is indicated by a higher attribution of sympathy.

A *moderate*, but principally positive estimation of perceived differences and strangeness indicates a sensitively attributed emotional meaning towards perceived difference (EM). An exaggerated assignment of sympathy towards perceived difference and strangeness in a positive or negative sense is more assumed to represent contradictory indicators with respect to a sensitive attribution of an EM. In order to deal constructively with strangeness, overwhelming emotions are more a bias which further may provoke irrational and unfair judgement of other persons or social groups. It is therefore assumed that a more rational interaction between na-



---

tive and stranger might be more helpful than being overwhelmed by positive or negative emotions such as related phenomena of exoticism or xenophobia illustrates. This does not mean that people should suppress emotions when dealing with strangeness. Being sensitive is here more understood as a way to be in better control of one's own feelings.

Cognitive and emotional processes are mutually related with each other. Emotional reactions can be evaluated and controlled by cognitive procedures. But causal attribution is influenced by our emotions, and can lead to distortions of our perceptions as presented previously. The following chapter focuses on the cognitive dimensions of the construct (STS).

### 6.3.2. Cognitive meaning (CM)

People try to understand other persons' behaviour – they attribute a cognitive meaning towards difference and strangeness – even though they lack information about the stranger. This may provoke misunderstandings and conflicts and support the before mentioned overwhelming feelings. A more rational way of dealing with perceived difference and strangeness is to be more sensitive to stereotypes and prejudice. Emotions are often difficult to handle and can reinforce the construction of stereotypes in particular feelings of uncertainty as argued in previous chapters. Rationality may help to be in better control of handling difficult emotions such as uncertainty. In addition, dealing with strangeness constructively is understood as a learning process. In order to perceive strangeness as a learning stimulus, it is necessary to meet the “stranger” open-minded. The implicit uncertainty can initiate critical reflections about one's own norms, standards and values. An openness and self-critical perception may lead to a more constructive dealing with strangeness and the person may possibly learn something new.

The learning idea finds support in cognitive theory where dealing with difference and strangeness can be understood as incentive for adaptations and re-arrangements of our existing cognitive structures (cf. chapter 4). Through active exchange and reflection people consequently become more and more familiar with the initially strange perceived person or situation. We may become aware of the fact that the other person is not that different as initially assumed. Situations may also show that perceived differences may not become understandable and just need to be accepted.

People explain an outcome of a situation differently. Such causal attributions often vary in rationality. An *awareness* of hypothetical argumentations and attribution strategies in the context of differences and strangeness may indicate a potential sensitivity towards strangeness as

well. The register of different rationality levels in causal attributions and reasoning strategies are supposed to differentiate between different levels of STS.

Dealing with feelings of uncertainty in a more rational way requires a stable or balanced developed identity. *Attitudes*<sup>32</sup> are usually related to cognitive, affective, and behavioural aspects. Furthermore, attitudes are important because they represent parts of our identity and influence our behaviour. Personal insecurity makes it more difficult to tolerate and further deal with the strangeness related feelings of uncertainty in a constructive manner. Insecure persons may more easily perceive strangeness as a threat to their own identity and may easier develop defensive attitudes and reaction strategies against “strangers”. Strangeness could then be avoided or devaluated in order to reduce its assumed menace. However, measuring relevant attitudes and insights related to STS was supposed to give information about people’s potential capability of tolerating and dealing constructively with related uncertainty.

---

<sup>32</sup> Attitude is understood as “the learned, relative stable tendency to respond to people, concepts, and events in an evaluative way” (Gerrig & Zimbardo, 2002:550).

#### 6.4. First structured item pool: Sensitivity towards Strangeness Questionnaire (STSQ)

In the following I will present my first structured version of the STSQ before presenting the empirical procedures which supported the development of this initial version. This initial version of the STSQ is actually a result of my theoretical considerations as presented before, the pre-studies which helped to find strangeness related pictures and situations, and the pilot study which helped explore how items of the STSQ can reasonably be summarized and labelled. The reason for presenting this initial version first is that it is probably easier for the reader to follow steps and procedures for item analyses and further development when already being familiar with the STSQ. In addition, it seems logic and it may support a better understanding of the STSQ when presenting it in line with the before presented operational model.

As shown in the conceptualization chapter, STS is understood as multi-faceted. The operational model shown in Figure 18 structures the developed item pool into different categories each related to the different dimensions of the STS. Considering the intended precision of measurement, it is important to point out that items of the same facet are not precisely to be understood as parallel items in the sense of Cronbach's alpha (cf. Cronbach & Shavelson, 2004). They are more to be understood as *complementary*. Items of the same facet are intended to measure a slightly different aspect but all related items are summarized by one superior facet. The resulting heterogeneously structured item pool is consequently supposed to measure a band width of the facets of the construct. This may reduce the precision of measuring one particular facet but the main intention is to measure the band-width of certain criteria.

The initial version of the STSQ is structured in three different types of measurement approaches:

- STSQ, part I:* A semi-projective device which focuses on an emotional valuation of perceived difference and strangeness. Pictures showing different persons or situations are used as stimuli for filling out three scales: closeness, dissimilarity and sympathy (cf. Figure 19).
- STSQ, part II:* A semi-projective device which focuses on the first cognitive dimension of the STS: the awareness of attribution and argumentation styles. Each of five characteristic situations include six different sub-items (suggested attributions and arguments) related to each situation separately (cf. Figure 24)

---

*STSQ, part III:* A statement item pool which is supposed to measure selected attitudes, insights, and understandings related towards differences, strangeness and relevant aspects in the realm of intercultural issues (cf. Figure 25)

In addition, a few demographic items are placed at the beginning and the end of the device. The purpose is to enable for later categorizations of the samples into sub-groups according to specific research goals when investigating the measuring characteristics of the STSQ.

#### 6.4.1. *STSQ, part I:* measuring emotional sensitivity toward strangeness

Attempting to measure people's meanings or attitudes towards intercultural issues may make associations by the respondents suspicious of a hidden agenda. They might think that it is their attitude towards racism that is really measured by the questionnaire. This provokes the already mentioned biases such as "political correct" or "social desirable" ways of answering the STSQ items. For this reason, indirect measuring methods were preferred to direct methods (Borkenau et al., 2005; Rost, 2004).

With the aim of measuring an emotional meaning (EM) attributed to perceived differences, a semi-projective measuring method seems to be suitable in order to meet this particular goal. The applied method is called semi-projective because it is using in part the idea of "projection". The respondents will get a visual stimulus (printed picture of a situation or person) which is supposed to provoke expressions of personal thoughts and feelings. Usually projective devices are open devices often applied in individual diagnostics where respondents are free to express their associations with respect to a given stimulus (e.g. Rorschach test). In contrast to projective devices, the here applied semi-projective method focuses on pre-selected emotions related to the concept of difference and strangeness. Respondents are supposed to indicate relevant perceptions on the three given scales. This reduction of measurable information seems reasonable when intending to develop a concept-driven measuring device that is intended to differentiate between groups. Measuring always implies a reduction of information compared to the content of a theory (Gadenne, 1994b). Therefore, it seems reasonable to focus on the most relevant and measurable facets of a theory instead of collecting as much information as possible (cf. chapter 6.1).

In general, the concept of "projection" characterizes a process where the respondent is supposed to project the „individual inside' to the „observable outside' (Fisseni, 1997). Projective devices are usually meant to measure unconscious aspects of one's personality (Rost,

2004). Such devices use a kind of defence mechanism<sup>33</sup> which leads to project one's opinions, norm and values towards other persons. The pictures of the selected item pool are intended to be an ambivalent stimulus when filling in the STSQ. Ambivalence of such projective items is needed in order to stimulate individual cognitive processes and enable to project personality characteristics or experiences into the shown situation or persons (ibid. 2004).

The principle of projection applied to the measurements of emotional STS (eSTS) means that pictures illustrating specifically selected situations and persons are going to be presented to the respondents (Table 3). Looking at the pictures is intended to provoke associations. Based on their individual associations related to the shown picture, respondents are supposed to indicate an emotional meaning attributed towards the given visual stimulus with respect to the three dimensions of closeness, dissimilarity (difference), and sympathy.

These three dimensions are operationalized as a five point rating scale where the respondents are instructed to cross out the respective number that is supposed to represent their personal attributed feeling (cf. Figure 19). Following the principles of semi-projective measuring methods, this device tries to take advantage of the principle of projection, but it is also intended to reduce some undesirable problems usually linked to (full) projective procedures. Data is often ambiguous and consequently difficult to interpret in too open devices (Fisseni, 1997). Relating the three scales to an open stimulus (picture) makes this part I of the STSQ semi-projective and is expected to increase reliability in comparison to complete projective measuring methods.

---


<sup>33</sup> The principle of projection stems from psychoanalysis and is referring to an externalization of (suppressed) thoughts and feelings. Such projections belong to defensive mechanism with the function to protect one's own identity. As a case in point, xenophobia can release such mechanism. However, in such a process personality traits of one person are projected to another person. Subsequently, one is convinced that this unconsciously projected or attributed personal characteristic really defines the respective person. Such a mechanism is used in projective measuring methods (Bibliographisches Institut & F.A.Brockhaus AG, 2008).

The item example illustrated in Figure 19 shows that each item consists of two main sections:

1. A picture, illustrating a particular person or situation where the respondents are instructed to write down their initial associations when looking at the picture
2. Furthermore, persons are asked to answer three different scales with the aim to indicate an emotional meaning towards the respective picture-stimuli.

**II. Have a closer look at each of the following pictures!**

**1 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!**



**1 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!**

*I feel ...*

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

*... towards the picture above.*

**Figure 19: Item example of the initial version of the STSQ, part I**

The written associations are mainly intended to insure that the respondents become more aware of what is shown on the picture by writing down their associations. Systematic analysis of this open question could be helpful for validation purposes of the assumed eSTS pattern score shown in Figure 20.

All three scales together are then assumed to measure dimensions of attributing an emotional meaning towards specific situations or persons shown on pictures. The indication of how *close* one feels related to the shown person or situation is assumed to represent a subjective meaningfulness or centrality related to their indicated associations. The perceived *dissimilarity*<sup>34</sup> is supposed to indicate in how far the respondent intends to distinguish him/herself from the shown person or situation. The attribution of *sympathy* is finally assumed to indicate how pleasant the respective situation or person is perceived.

<sup>34</sup> Instead of using the term “difference” in the questionnaire, the term “dissimilarity” is preferred in order to avoid terminologies from the theoretical construct.

It is intended that measurements of specific response patterns may give indications about eSTS. An indication on the dissimilarity scale such as “4”, “5” or “1”, “2” respectively is assumed to show that the person perceives him-/herself as more different or similar respectively in relation to the illustration. Higher scores on the dissimilarity scale show that a respondent intends to demarcate or identify him/herself with the illustration shown in the picture. This is interpreted as an indicator of a meaningful perception of a shown person/situation because the response scores 4, 5 (dissimilar) or 1, 2 (similar) will indicate a distinctive (meaningful) position (similar or dissimilar). Value “3” on the dissimilarity scale can be interpreted as an indication of lesser subjective meaningfulness. In such a case, the respondent does not feel similar or perceives him/herself as differently from a person/situation shown in the item. This can be interpreted as giving lesser subjective meaning.

Figure 20 shows the response pattern indicating an eSTS. The picture of the respective item is perceived as different from the individual’s own perception which is indicated by a value 4 or 5 on the dissimilarity scale. Those items which are indicated as different towards one’s own perception meet the ideal prerequisite for measuring an eSTS. The sympathy scale is the key-element for the attempt to measure an eSTS. A moderate positive attribution of sympathy (value 3-4 on the sympathy scale) is supposed to indicate a sensitive emotional evaluation of differences in the sense of the construct. An ideal response interpreted as an eSTS is illustrated in figure 20.

Dissimilarity:	1	2	3	4	5
Sympathy:	1	2	3	4	5

**Figure 20: eSTS pattern score =**  
**(Dissimilarity = 4 OR Dissimilarity = 5) & (Sympathy = 3 OR Sympathy = 4)**

The mean-value of all measured eSTS pattern scores of all shown picture items will lead to a summarizing eSTS-score of one respondent.

Reasonable alternative response patterns are possible. A low attribution of sympathy (score 1 or 2 on the sympathy scale) will indicate that a situation or person illustrated in the item is evaluated as relatively unpleasant. In combination with indicated dissimilarity (score 4 or 5) this pattern score is interpreted as an indicator for a kind of *xenophobic* tendency. That is be-

cause the perceived difference is emotionally valuated as lesser pleasant or unpleasant as illustrated in figure 21.

*Reasonable alternative response pattern:*

Dissimilarity:	1	2	3	<del>4</del>	<del>5</del>
Sympathy:	<del>1</del>	<del>2</del>	3	4	5

**Figure 21: “xenophobia” pattern score = (Dissimilarity = 4 OR Dissimilarity = 5) & (Sympathy = 1 OR Sympathy = 2)**

A second alternative and reasonable response pattern could be an “extreme” attribution of sympathy (indicated by a value 5) to an indicated perceived difference score (cf. Figure 22). Such a pattern is interpreted as an overvaluation of a perceived difference. The response pattern seems more to indicate feelings such as *pity* for a shown person on the picture item instead of measuring an eSTS. Because of its unconstructive character, *pity* is not assumed as helpful when dealing with differences constructively and is therefore no useful indicator when intending to measure eSTS.

Dissimilarity:	1	2	3	<del>4</del>	<del>5</del>
Sympathy:	1	2	3	4	5

**Figure 22: “Pity” pattern score = (Dissimilarity = 4 OR Dissimilarity = 5) & (Sympathy = 5)**

A third alternative and plausible response pattern can be expected when respondents indicate that they feel similar to a person/situation shown on a picture item (indicated by values 1-2 on the dissimilarity scale). In such cases respondents may associate personal experiences with the shown person or situation shown on a picture item. It would more likely indicate that the person or situation shown on the picture item seems familiar to the person. This is not a suitable prerequisite for measuring EM because the subsequent attribution of sympathy cannot be related to the perception of differences. An indication of perceiving the presented person or situation as similar to one’s own perception can be a signal for believing or wishing to be familiar with the shown persons or situations. The assignment of feeling more familiar than different (indicated by a 1 or 2 on the dissimilarity scale) in combination with an overwhelm-



ing attribution of sympathy (indicated by a 5 on the sympathy scale) are supposed to be an indicator for exotic attribution tendencies (cf. Figure 23).

Dissimilarity:	*	*	3	4	5
Sympathy:	1	2	3	*	*
















**Figure 23: “Exoticism” pattern score =**  
**(Dissimilarity = 1 OR Dissimilarity = 2) & (Sympathy = 4 OR Sympathy = 5)**

It has to be mentioned that the presented alternative response patterns are not of central interest here because it is intended to measure an eSTS. But demarcating eSTS pattern towards interpretable alternative pattern helps to differentiate the intended construct and to gain helpful validity hints when evaluating the STSQ.

Figure 19 illustrates that a picture is used as stimulus for answering the item. Table 3 gives an overview of selected picture items. The later reported pre-studies demonstrate how the pictures have been selected and validated (cf. chapter 6.5). Pictures marked with “\*” are not part of the initial item pool. Those pictures were added during the empirical investigations and will be reported in the respective chapter.

The following eleven pictures were selected for the initial item pool (I.1 – I.11):

**Table 3: Image used in part I of the initial version of the STSQ<sup>35</sup>**

				
I. 1	I. 2	I. 3	I. 4	I. 5
				
I. 6	I. 7	I. 8	I. 9	I. 10
				
I. 11				
				
*I. 12	*I. 13	*I. 14	*I. 15	*I. 16
* Further item images I.12 - I.16 added in research progress, and consequently are not part of the initial item pool				

The pictures above were used in order to stimulate associations regarding selected topics or categories. As introduced in the context of subjective meaningfulness (cf. chapter 3.3 and Haußer, 2007), a person's meaningfulness can be profiled by various subjective meaning themes. The pictures were pre-selected according to strangeness related themes which were collected with help of the pre-studies (cf. chapter 6.5):

<sup>35</sup> Picture I.3 and I.11 originate from the online version of the IAT (Nosek, Banaji, & Greenwald, 2005). All the other images used in the pre-studies and subsequent versions of the STSQ are free sources searched with Google picture search.

- I. 1 The first picture is originally a painting showing a woman who originates from Jamaica, but lives in Germany and is actually posing for this painting. She has tied a scarf around her head, and has a darker skin-tone (in relation to the majority around her). The main intended stimulus direction is “skin-tone”. But the ambivalent character of this picture may also represent categories such as “culture”. A characteristic such as “skin-tone” needs not necessarily be perceived as different because it is socially constructed. The category “skin tone” or an attributed “cultural background” is still a crucial starting point within intercultural conflicts. Consequently, the first picture is assumed to be relevant stimulus in measuring eSTS.
- I. 2 The next picture shows a young boy with “Down’s syndrome”. Physical or mental disability is often perceived as different. For people who are not used to Down’s syndrome, this picture may provoke a perception of strangeness.
- I. 3 The man on the next picture represents persons with overweight. In a society where fitness, health and a slim look is “declared” an ideal of beauty, such a person can quite often be perceived as different.
- I. 4 The next picture is a commercial from the 50<sup>ties</sup>. It shows an “old fashioned housewife” presenting kitchen tools. This item shows a difference between a conservative role patterns in comparison to a more emancipated perceptions of a women in the modern (western) societies. As discussed earlier, the perception of difference does not only refer to ethnic difference. This item is supposed to represent a stimulus category for social status difference.
- I. 5 Behaviour such as sexual orientations is often perceived as different and strange if diverge from the dominating (morally) accepted norms and values. A picture of two men kissing each other represents a minority of sexual orientation and is therefore a stimulus for perception of difference.
- I. 6 Religion is also a topic related to norms- and value-systems. Actual events demonstrate a high conflict potential between different religious groups. This is a controversial and complex category represented by an item showing a woman’s face covered with an opaque veil (*burqa*). This *burqa* covers her entire face except for a small region around her eyes.
- I. 7 The perception of difference in knowledge, competence or education can also be perceived as strange. People with no university or academic experience may feel out of

- place or strange within a group of university teachers. The respective item is illustrating a situation in a university auditorium. A teacher is writing notes on the black board in front of his students.
- I. 8 Similar to item 3 (overweight), this item is focussing on body image as well, but this time in an opposite direction. The person shown in this picture is a female fitness trainer demonstrating an exercise.
- I. 9 The next item shows a “white” man and a “black” woman holding hands and smiling at each other. This item is expected to be more ambivalent because of showing persons with differing skin tones.
- I. 10 Item 10 shows another, often unaccepted minority group. The picture shows the backs of three prostitutes waiting at the roadside. Prostitution is perceived as different from other professions, and moreover judged as an immoral profession. This category is closely linked to moral valuation and judgment. But this item might also be a more ambivalent item because not everyone may perceive the shown women as prostitutes.
- I. 11 Age difference can be responsible for experience of strangeness. Young people may have difficulties in understanding convictions or behaviour of older generations (and vice versa) because both parties have different historical background. This category may help additionally to estimate an eSTS with respect to relevant categories.

Using pictures as stimuli is problematical because the stimulus direction can be quite subjective. But it is needed in projective devices as indicated before. If a shown picture appears as strange or familiar to an examinee depends on one’s own culture. Without a doubt, culture is complex, but culture also shows more or less consistent pattern and structures. If a culture could not be demarcated to some extent, one crucial function of culture would get lost i.e. the function of giving orientation. For instance, language and communication is part of a culture. By all its complexity, if there were no consistency and common sense in language as part of a culture, communication would not be possible.

Another aspect needs to be considered regarding the selection of the pictures for part I. Using far foreign cultural groups as picture items (e.g. from a remote cultural group in the Amazon forest), which probably would be perceived as strange by most of the western population, would not be helpful to evaluate sensitivity towards strangeness. Simmel’s essay about the stranger points out why the stranger which is far away (e.g. remote cultural group) would probably be perceived as strange. But dealing with strangeness in a meaningful way, requires

a (social) closeness to strangers; closeness in terms of a meaningful relationship. Simmel tried to point out this idea with hypothetical aliens on a planet in a different galaxy. The stranger must be close and consequently more relevant to us. This idea suggests selecting pictures out of the respective population instead of selecting pictures of far distant people such as people from a remote cultural group in the Amazon forest.

As demonstrated, the use of pictures as stimulus is a quite delicate matter. Therefore, it is advisable to test and reconsider the selection of pictures (within a pilot study), when applying the item pool in different cultural context. If the items of part I meet later explained criteria, they probably will show some validity and consistency as well. Applying measuring instruments across cultures are a challenging task and require comprehensive investigations. Such aims are beyond the aims of my thesis but cross cultural validity investigations are interesting and relevant goals for the future.

#### 6.4.2. *STSQ, part II*: awareness of people's argumentation strategies and attributions styles with respect to relevant situations

The next part of the STSQ is directed to measure the first cognitive facet of the construct (cf. Figure 18). The intention is to measure people's awareness of potentially existing attributions and argumentations with respect to relevant situations. This part of the instrument is operationalized as a situational questionnaire. Five characteristic situations are functioning as a baseline for an evaluation of selected attributions and arguments about how people would argue or explain an outcome of a situation<sup>36</sup>. Respondents are instructed to imagine themselves in the described situation. Then, they are asked to estimate the probability of six suggested attributions or reasons linked to each situation on a scale from 1 to 4.

The aim is to register if people either believe that the suggested reason is probable or not. As illustrated in the following item example, the scale is characterized as a bipolar rating scale with two scale-intervals in both directions.

---

<sup>36</sup> The situations are constructed in a similar strategy as the picture items in the instrument part before. The later reported pre-studies helped to collect and construct the five different situations of the initial version of the STSQ.

III. Put yourself into the following situations!				
1) <i>Imagine people are entering a bus. All passengers pass by the conductor. One person was asked to show his/her ticket.</i>				
<b>What reasons could the bus driver have to stop this person? Please evaluate the probability of each suggested reason!</b>				
1. The person looks different	1 most improbable	2 improbable	3 probable	4 most probable
2. The person's behaviour seems to be suspicious.	1 most improbable	2 improbable	3 probable	4 most probable
3. The bus driver is irritated.	1 most improbable	2 improbable	3 probable	4 most probable
4. Generally, strangers are controlled.	1 most improbable	2 improbable	3 probable	4 most probable
5. The bus driver does not like foreigners.	1 most improbable	2 improbable	3 probable	4 most probable
6. It was a routine control.	1 most improbable	2 improbable	3 probable	4 most probable
Further reasons	_____			

**Figure 24: Item example of part II of the initial version of the STSQ**

In order to avoid “neutral” indication, the scale is constructed without a medium- or neutral position. The reason for this scale type is to provoke that the examinees make a decision when answering the items (cf. Figure 24). Six attributions or reasons are listed under each of the five situations. It is attempted to keep a particular structure within and between each situation. The guiding criterion for the item pool structure is that the items are supposed to vary in their level of rationality (RA) i.e. from level RA I - III. Each situation contains at least two different levels dependent on the situation. There are for instance arguments that are supposed to be of a stereotyping or stigmatizing character. Such items are consequently classified as more irrational because they are usually based on a lack of information and prejudice. On the other hand, people may be aware of different hypothetically existing attributions. Consequently, they may indicate a broader spectrum of suggested items which means that they estimate more items as probable than improbable.

The situational item pool uses the principle of “projection” in a similar way as described in part I of the STSQ. Potential respondents will be instructed to imagine a described hypotheti-

cal situation, and evaluate the probability of these suggested reasons/arguments existence in reality. In addition, people may to some extent “project” their personal opinion into their individual way of answering the items. It can be expected that the item responses may reflect the respondent’s personal reasoning and attribution style indirectly as well as the accumulated awareness over hypothetically existing attributions and arguments.

The semi-projective measuring procedure is chosen in order to reduce error influences. The applied process of projection is directed towards the ability to imagine a situation and then trying to imagine how the situation might be explained from the perspective of another person. Not asking respondents directly about how they would argue is supposed to reduce measuring errors such as political correctness or socially desirability or expected rating of the items. By transferring suggested item answers towards other hypothetical persons described in the situations, it will be possible to measure personal opinions and at the same time reduce relevant error influences which are specifically related to intercultural topics.

The items are summarized in categories of related items. As a result, there are at the most three different classes of attributions and reasons within the situations:

**RA 1:** More irrational such as stereotyping, stigmatizations, irrational or unreasonable attribution of responsibility

**RA 2:** More rational reasons/attribution level 1

**RA 3:** More rational reasons/attribution level 2

RA 1 represents the lowest, and RA 3 the highest rationality level related to the five specific situations.

Before analysing the collected data, item raw scores needed to be summarized (coded). The unbalanced number of items within each RA-level suggests combining raw scores into mean-scores. The three resulting mean-scores RA1, RA2, and RA3 are interpretable as indicators for the spectrum of attributions and argumentations on the mentioned three different levels. The differentiation between the three different types of RA-level moreover enables to distinguish between different qualities of aware attributions.

*Situation 1: Imagine people are entering a bus. All passengers pass by the conductor. One person was asked to show his/her ticket.*

In situation 1 the respondents are asked to imagine a situation where people are entering a bus. The crucial element with respect to difference and strangeness is that only one person is asked to show his/her ticket; all the other passengers can just pass by the conductor. The respondents now evaluate the probability of each suggested reason explaining why the conductor stopped just this particular person:

**Table 4: Items linked to situation 1, part II of the STSQ**

Item code	Items of situation 1	Category
1.1	The person looks different	RA 1
1.2	Generally, strangers are controlled	
1.3	The bus driver does not like foreigners	
1.4	The bus driver is irritated	
1.5	It was a routine control	RA 2
1.6	The person's behaviour seems to be suspicious	

Table 4 shows that situation 1 contains items of two different levels: RA 1 and RA 2. The first four items in Table 4 represent stigmatizing arguments. The conductor stops the person because he/she is perceived as different (cf. item 1.1), and he suspects that strangers usually do not buy tickets (cf. item 1.2). Another reason might be that he perceives the person as a foreigner, and he does not like foreigners (cf. item 1.3). The other two items of this situation may lead to more rational conclusions. The association of a routine control may represent a more value-neutral item. The conductor may also have some experiences how people behave when not having bought a ticket. In this case, the passenger can be perceived as behaving suspiciously because he/she appears to be afraid of getting caught.

*Situation 2: Imagine a basketball game. Team A is playing with a new team member who is expected to lead the team to the top. The match is over, and they have lost the game.*

The situation 2 refers to a basketball game. The focus is put upon one new player which is expected to lead the team to the top. Despite the coaches' expectations the team with the new player have lost the game. The items shown in Table 5 suggest different reasons which vary in rationality.



**Table 5: Items linked to situation 2, part II of the STSQ**

Item code	Items of situation 2	Category
2.1	The new player did not play as well as expected	RA 1
2.2	The new player did not adapt to the team	
2.3	The new player disturbed the team spirit	
2.4	The team opposed to the new player	
2.5	The new player was not integrated into the team	RA 2
2.6	The team was in poor condition	

The first four items of situation 2 represent the lowest level in rationality (RA 1) because they represent a kind of irrational explanation why the team probably has lost the game. The reason for not winning the game is only attributed towards the new player in all of these RA 1 items. As attribution theory shows, people tend to make an attribution error when dealing with strangers. Hence, the perception of undesirable behaviour of strangers is attributed stronger internally than externally. People tend to attribute failure towards other persons or groups and attribute success more to one self or the own group (cf. chapter 4). The result of making the new player (stranger) solely responsible for the failure represents this kind of attribution error and therefore represents a more irrational attribution style (RA1).

That the new player was not integrated into the team is understood as more rational attribution (Item 2.5, RA 2) because it might be lesser condemning than the first four items; it may further suggest that the new player is not yet integrated into the team. Not being integrated into the team relieves the new player from his/her exclusive responsibility. Basketball is a complex team sport and logically demonstrates that it is quite unreasonable to blame one player for losing the game. Therefore, the last two items (2.5 and 2.6) are assigned towards RA 2.

*Situation 3: Yildiz is a student at a high school for boys and girls. Outside of her home, she wears wide clothes and a head scarf. Last week Yildiz's family applied to exempt her from the co-educational Physical Education (P.E.) class because her family's Islamic faith does not permit girls to participate in sport together with boys.*

Situation three is inspired by a research study of Bender-Szymanski (2004; 2006). This study is related to the theoretical concept of "argumentation integrity" as described in chapter 2.3. Bender-Szymanski developed an experimental role-play with the aim of learning how to deal with religious and ideological conflicts in a more reasonable and democratic way. Situation three is based on a crucial situation described in the Bender-Szymanski's project, which

is originally based on a real conflict from a Germany court a few years ago. The design and the basic conclusions of Bender-Szymanski's research appear to fit the intended structure of the situational item pool. A basic argumentation pattern of Bender-Szymanski's project is transferred to situation 3 and adapted as item basis for situation 3 of the STSQ, part II. The described three categories of argumentation-integrity are relevant classifications with respect to the intended item structure. The situation for this item is that an Islamic family is trying to exempt their daughter Yildiz from P.E. class because they maintain that their Islamic faith does not permit girls participate in sports together with boys.

The following table shows the item structure according to their assigned rationality level RA 1-3:

**Table 6: Items linked to situation 3, part II of the STSQ**

Item code	Items of situation 3	Category
3.1	Co-educational teaching should have priority over individual beliefs	RA 1
3.2	Pupils should obey existing rules	
3.3	P.E. class could be organized gender-separated	RA 2
3.4	Students have the possibility to select exercises	
3.5	Different religious faiths should be respected	RA 3
3.6	P.E. can be arranged in a mutually acceptable way	

The first two items shown in Table 6 represent more rigid arguments which would reject the request of Yildiz's family. The next level represents willingness to compromise as regards separate boys and girls P.E. classes. It may also give Yildiz the right to choose exercises which will be acceptable to her family. The final two items represent the most open-minded arguments. They imply respect towards difference, and at the same time the willingness to work out balanced solutions by taking into consideration: the school's interests and the needs and wishes of Yildiz and her family.

*Situation 4: Imagine a company has to decide between two final applicants for a leadership position. The final candidates are a man and a woman.*

The next situation of the item pool refers to a company that has to decide between two final applicants for a leadership position. The two candidates are a man and a woman. The purpose is to measure an awareness of different argumentations with respect to gender. Gender is further a relevant concept within intercultural issues (cf. Gieß-Stüber, 2000a). The generalization that men are more efficient than women or that women are more integrative leaders are irrational assumptions based on stereotypes (cf. RA 1). The company's decision should be

determined by professional qualification or personal leadership. These items represent more rational arguments than the items mentioned before (cf. RA 2 vs. RA 1). The most differentiated arguments suggested are the last two items of Table 7 (RA 3).

**Table 7: Items linked to situation 4, part II of the STSQ**

Item code	Items of situation 4	Category
4.1	Men are supposed to be more efficient	RA 1
4.2	Women are supposed to be more integrative leaders	
4.3	Personal leadership determines the decision	RA2
4.4	Professional qualification determines the decision	
4.5	In case of equal qualification, the gender-balance determines the decision	RA3
4.6	The heterogeneity of the company structure will influence the decision	

The gender quote becomes relevant when both candidates show equal qualification. The items before do not even consider the gender balance. That the heterogeneity of the company structure will influence the decision (item 4.6) is assigned to the same RA level as the item before. But it indicates awareness of more complex considerations needed when considering the heterogeneity of the company's structure.

Even though gender differences can provoke feelings of strangeness, measuring gender differences in STS is not a main goal of this thesis. The conceptualization and operationalization of STS is the major focus of interest. There are no reasonable theoretical assumptions about how different genders differ with respect to sensitivity towards strangeness. Gender is assumed to be one relevant facet among others such as ethnic differences which can lead to experiences of strangeness. But a gender related research questions could be an interesting goal of subsequent research.

*Situation 5: Imagine a student begins to study at the university. Neither people, nor structures are familiar to him/her.*

The last situation of this part of the STSQ represents a typical situation where strangeness can be experienced. The respondents are instructed to imagine the first day at a new university. Because the instrument later mainly will be applied with college/university students, most of this population might have experienced such a situation before. The basic idea of this item is to register information of how far one can imagine what it is like to be new or strange. This situation is therefore relatable to the concept of empathy.

**Table 8: Items linked to situation 5, part II of the STSQ**

Item code	Items of situation 5	Category
5.1	This situation does not provoke specific irritations	<i>Ci</i>
5.2	It must be exciting for him/her to discover the new environment	<b>Pleasant perception</b>
5.3	He/she is confident to overcome the uncomfortable feelings of the first day	
5.4	The unknown expectation of the new surrounding makes him/her feel uncomfortable	<b>Unpleasant perception</b>
5.5	Getting to know people in such an impersonal situation must be difficult	
5.6	He/she feels insecure and disorientated	

Item 5.1 (Table 8) function as a kind of control item to get hold of possible measuring biases such as superficial answering of the questionnaire. An indication that this situation might not lead to any kind of irritations or confusion seems to be unreasonable and might indicate a possible measurement error.

The other items are meant to indicate some kind of empathy by imagining how it would feel to be in such a new situation or feel strange. While items 5.2 and 5.3 are assigned to a more pleasant perception, item 5.4, 5.5 and 5.6 are supposed to indicate unpleasant perceptions regarding the characteristic situation described in situation 5. Even though one might feel uncomfortable at the beginning, item 5.3 shows that these unpleasant feelings will disappear gradually. The item is referring to an understanding that strangeness can be deconstructed and changed more and more into familiarity. The items of unpleasant perceptions are an indicator of strangeness as difficult to handle.

The items indicate a potential awareness about how people might feel in strange situations. The items are in this sense expected to function as further indicators of measuring awareness. Situation 5 is also constructed for validation purposes of the STSQ. Correlations with related items might help to estimate the validity of respective items or item scores of the STSQ.

#### 6.4.3. *STSQ, part III: relevant attitudes, and understandings*

The third and final part (III) of the STSQ is an item pool of 17 statements measuring relevant attitudes and understandings related to dealing with difference and strangeness. The items indicate more or less complex ideas. The intention is to measure peoples' fundamental understanding of strangeness and related intercultural issues. A dynamic and pluralistic understanding of culture is needed in order to esteem other cultural values (cf. Auernheimer, 2003; cf. Chen et al., 2002; Fritz, Mollenberg, & Chen, 2002; Fritz-Gerald, 2002). Such a

differentiated understanding of “culture” is for instance a determinant indicator for being sensitive towards strangeness.

The following table shows four categories<sup>37</sup> which structure the initial item pool of part III of the STSQ:

**Table 9: Initial item pool, STSQ, part III (EC: Egocentrism, O: Openness, NFS: Need for security, LOC: Loss of control)**

Item code	Item pool, part III of the STSQ	Category
EC 1:	Other people are not as open-minded as I am.	EC
EC 2:	My personal way of life should be a model for other people.	
O 1:	It seems that different people have different values.	O
O 2:	The open result is responsible for the excitement of a competition.	
O 3:	To interact with different kinds of people is enjoyable.	
O 4*:	Sitting in a group of strangers provokes uncomfortable feelings.	
O 5:	There is no freedom without rules.	
NFS 1:	It is important to forward cultural values to the next generations.	NFS
NFS 2:	It is necessary to plan ahead in order to avoid surprises.	
NFS 3:	Being obliged to make decisions provokes uncertainty.	
NFS 4:	For activities, I prefer precise instructions to open suggestions.	
LOC 1*:	The number of friends I make depends only on me and my behavior.	LOC
LOC 2:	Most of the events in my life are determined by other people.	
LOC 3:	Fate determines whether I have more or fewer friends.	
LOC 4*:	In spite of individual differences people do not differ substantially.	
LOC 5*:	In spite of its apparently continuous change, culture contains some stable values.	
LOC 6:	Unforeseen events upset me.	

\*inverted item formulation with respect to the regarding category

Like part II of the STSQ, the scale is a bi-polar rating scale and varies between “strongly disagree - disagree – agree – strongly agree” respectively:

9. Unforeseen events upset me.	1	2	3	4
	strongly disagree	disagree	agree	strongly agree

**Figure 25: Item example of the third part of the STSQ (initial version)**

The two first items (cf. Table 9) measure relevant aspects of the previously discussed “egocentrism” (EC). These items are adapted examples from the “Intercultural Sensitivity Inventory” by Hammer, Bennett, and Wiseman (2003). People who point out their own personality as more valuable than others are expected to show a lesser degree of sensitivity when dealing with strangeness. They seem to feel overbearing with respect to other persons or groups of assumed distinguishing culture. STS requires a critical perspective of one’s own

<sup>37</sup> Also here, the categories were developed within the pilot study.

(dominating) cultural norms and standards while an egocentrism is characterized by an over-valuation of one's own culture. The result is an unreflective one-sided view of one's own culture. These two items are negatively formulated indicators with respect to measure a STS.

An open-minded attitude is needed when dealing with difference and strangeness in a constructive manner. As previously discussed, openness is dominated by data-driven perception patterns. Consequently, one has to be careful with unjustified conclusions or judgements of strangers because our interpretations of strangeness are based on a lack of information.

An open-minded attitude (openness) towards difference and strangeness cannot be observed directly because it is a latent variable of the construct. The ways one hypothetically deals with difference and strangeness may give some information about a person's openness towards strangeness. Agreeing with the statement that other people have different values is supposed to indicate a basic understanding that there are other values than one's own (cf. O 1, Table 9). The indication that one agrees with the statement that the open or uncertain outcome of a competition is the determined factor for its excitement, is supposed to measure a positive or open attitude towards strangeness (cf. O2, Table 9). This requires a competence in tolerance, and attributes in principle a more positive meaning towards uncertainty. The phrase in of item O2 "[...] open result [...]" creates a relation to the construct of uncertainty. When agreeing that the uncertain outcome is responsible for the excitement of a competition, the often negatively associated uncertainty seems to gain a more positive meaning.

The next two openness items indicate an enjoyment in dealing with difference and strangeness more directly (cf. O 3 and O4, Table 9). Item O 4 is formulated negatively and therefore needs to be recoded before data analyses. Items O3 and O4 are more directly related to the construct of the STS. In particular item O 4 is referring to the concept of uncertainty when meeting a stranger. Theoretically, it is assumed that uncertainty in this context may become difficult to handle when the intensity of related uncertain feelings is perceived as overwhelming. The situation can be experienced as difficult to stand. To deal with strangeness means that it is difficult to predict an outcome of the personal effort we may put into a situation. This unpredictability may show that we at least have to tolerate the implicit uncertainty; that we need to take a risk in order to "deconstruct" strangeness. We may thereby learn something about the "other person" – which means that we learn something about ourselves as well as previously discussed.

The term "freedom" is the supposed key-word of the next item (O5). Freedom is in principle associated with the individual right to act, speak or think as one wants. This general idea of freedom represents a further relevant indicator for openness or an open-minded attitude.

The specification of the item that “there is no freedom without rules” moreover implies an additional understanding. The statement (O 5) is supposed to represent the meaning of freedom in line with a basic understanding of democracy. This idea implies an individual responsibility to act according to rules which in principle ensure all people of a society including minority groups an opportunity and a right to freedom such as a free choice of religion.

As mentioned before, dealing constructively with strangeness requires tolerating implicit uncertainty or insecurity (cf. Buhr & Dugas, 2002). People who show a particular need for security (NFS) are expected to show intolerance towards uncertainty. They may at least perceive uncertainty as unpleasant, and probably try to avoid uncertain situations. The next four items (NFS 1-4, Table 9) indicate relevant aspects of this category. To forward cultural values to the next generation represents a need for security in trying to keep existing values. The idea behind this item (NFS 1) is to indicate a kind of unwillingness or resistance against possible cultural change (and consequently the implicit uncertainty as well) by stressing to keep “the good, old values”. But being sensitive towards strangeness and consequently perceive strangeness as an incentive for probable development, requires taking a risk and being open for considering possible changes.

The necessity to plan ahead in order to avoid surprises is a more direct and concrete item indicating a need for security (cf. NFS 2). The agreement with the statement that “being obliged to make decisions provokes uncertainty” (NFS 3) is additionally indicating a higher need for security. One is probably afraid of having responsibility and possible consequences. The final item of this category measures a similar idea. If one indicates to prefer precise instructions to open suggestions, this shows a need for security. Although, open suggestions would allow more freedom it also requires tolerating the implicit uncertainty in dealing constructively with a suggested task.

The final category of the third part of the STSQ is related to uncertainty conceptions. Now the connotation is closer directed to the German version of Rotter’s<sup>38</sup> “locus of control” concept (Krampen, 1981). Locus of control can be a situational variable indicating a feeling of being able to control the outcome of a situation. If people perceive themselves more to “the mercy of waves”, the arising feelings of uncertainty can be perceived as more difficult to stand. It may also easier to be interpreted as threatening towards the own personality. This may lead to an activation of defensive mechanisms in order to protect one’s self-image.

---

<sup>38</sup> Rotter, 1966

---

The first item of the LOC-category originates from the IPC-questionnaire (Krampen, 1981). It is one indicator in an internal locus of control scale. It is referring to the person's subjective perception of being in control of one's own events in live (cf. LOC 1, Table 9). The next item originates from the same source but represents a different dimension of the construct: "powerful others external control orientation" (Levenson 1972, quoted in Krampen, 1981:8). It is determined by a subjective feeling of powerlessness, and can be characterized as a feeling of a social dependency of other more powerful persons (cf. LOC 2). The last item originating from the IPC-questionnaire is referring to another external dimension of the "locus of control" construct. This item belongs to the "chance control orientation" (Levenson 1972, quoted in Krampen, 1981:8) and is determined by a kind of fatalism which means a general expectation ("Erwartungshaltung"). The world around us is perceived as unstructured and chaotic so one's life depends on fate, bad luck, and coincidence (cf. LOC 3). The latter aspect is represented by two further items (LOC 4, 5). These items are closer linked to the topic of strangeness because they are referring to the perceptions of other persons (LOC 4) and cultures (LOC 5) (cf. Fritz-Gerald, 2002).

The last item may appear ambiguously. On the one hand, it could be assigned towards the NFS category because its meaning is relatable to the items NFS 2, 3, and 4. But the additional aspect of unforeseen events upsetting one personally is closer related to the subjective feeling of controllability over events in life. The fact that one feels "upset" with respect to unforeseen events places this item theoretically into the category of LOC.



## 6.5. Pre-studies: helping procedures for item pool collection

Besides deductive item construction, the pre-studies were a helping procedure for item pool collection. These field studies were mainly directed at collecting items and help construct the first structured version of the STSQ questionnaire. The operational model was a guideline for the intended structure of the STSQ. These theoretical considerations in combination with the gained experiences of the pre-studies resulted in the first version of the STSQ (cf. chapter 6.4 and appendix 9.8.1).

### 6.5.1. Design

The pre-studies were basically concerned with collecting a picture item pool and situations which were intended to be used as item stem or item stimulus for the semi-projective parts of the STSQ.

The pre-studies were conducted in a sequence as indicated in table 10. I planned and conducted open interviews, guided group discussions, and writing tasks with high school students to get a first impression of the empirical representation of the phenomenon “strangeness”.

**Table 10: Pre-study<sup>39</sup>**

Study	Group	N	Age	Methodological procedure
1.1	School A: Norwegian high school	N <sub>1</sub> =12 N <sub>2</sub> =5 N <sub>3</sub> =2	17-18	Group discussions
1.2	School A: Norwegian high school	N <sub>4</sub> =5	16-17	Writing tasks
1.3	School B: Norwegian College school (Students with immigration background)	N <sub>5</sub> =19		Writing tasks
2.1	School A: Norwegian high school	N <sub>6</sub> =10 (m:9, f:1)	17-18	Questionnaire A, group discussions
2.2	School A: Norwegian high school	N <sub>7</sub> =20 N <sub>8</sub> =5	16-17	Questionnaire B, group discussions
3	Expert group within IME	N <sub>9</sub> =3		Constructive group discussions

Two questionnaires were supposed to validate the collected pictures and situations for the semi-projective parts of the STSQ (cf. appendix 9.9 and 9.10).

The resulting structured raw version of item pool was critically discussed together with three experts<sup>40</sup> working in the field of “IME”. Besides linguistic suggestions of single item

<sup>39</sup> The samples were mainly high school students which were considered as relevant group for applications of the STSQ in the future.

formulations, the discussions circled around the question of how far my so far developed item pool represents relevant facets of “intercultural movement education”.

#### 6.5.2. Method

##### *Pre-study 1: Exploration of empirical representation of strangeness*

At first, more open group discussions (pre-study 1.1) were carried out with the intention of getting hold of the students’ perceptions of strangeness, their personal awareness about social structures and mechanism leading to experiences of strangeness. These discussions helped to get useful ideas from the pupils’ level of reflection with respect to their personal awareness of the phenomenon of strangeness. The duration of discussions in each class lasted between 45-60 minutes.

The group discussions were guided by pre-selected questions and statements. First, students were asked to explain what it means that “only abroad people are perceived as strange” (“Fremd ist der Fremde nur in der Fremde”). Secondly, they should discuss the meaning of strangeness, report about situations they personally perceived as strange, and finally try to express how they felt in such a situation. All discussions were recorded in the form of field notations, and edited directly after each session with respect to gather strangeness related situations, feelings, and attempts of explaining personal experiences with strangeness.

Based on the results of the group discussion, I developed writing tasks. These tasks were carried out with a new class from the same school at the same high school level. In the first task students were asked to answer the following picture item (cf. figure 26).

---

<sup>40</sup> Professor Dr. Ralf Erdmann, initiator of IME; Dr. Elke Grimmering, researcher dealing with development of didactical concepts in the context of intercultural competence of teachers (cf. Grimmering, 2009); Carmen Cabrera-Rivas, also working in the field of IME and with practical approaches within the field of physical education.

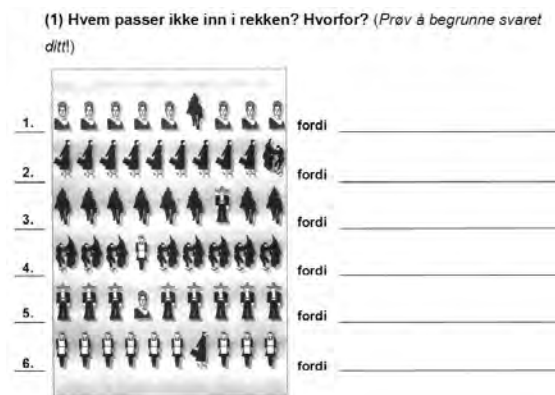


Figure 26: Task 1 of the writing tasks in the pre-studies

Resulting explanations were expected to give some feedback about pupil's knowledge of the social construction of strangeness. Similar to the open group discussion, they were additionally asked to write about typical situations they personally perceived as strange, and their reflections upon their feelings within these situations. Finally, they were asked to write down why they perceived described situations as strange (cf. appendix 9.1).

Both writing tasks and group discussions were personally guided by the author. In addition, it was important to keep the group discussions open. The arising dynamics in such open discussions were helpful in order to stimulate pupil's associations with respect to the given topic. Nevertheless, because of the powerful dynamics during the discussions it was difficult to control how much information was given to the students by the discussion leader, and how this affected the discussion. Therefore, the writing tasks (pre-study 1.2, 1.3) with different but comparable groups were conducted with the aim of gaining more controllable (written) information about the empirical representation of strangeness among the respective students.

*Pre-study 2: Cross-examination of the picture and situation items*

Further theoretical considerations, supported by the mentioned initial pre-studies (1.1-1.3), helped to construct 8 situations (cf. appendix 9.2). In addition, an internet based picture search with "Google" according to particular criteria led to the result of selecting 16 pictures as supposed stimulus for the picture-item pool of the STSQ. The following table shows the assumed relevant themes of the picture-item search on the Internet:

**Table 11: Key-terms for “Google” picture search**

Relevant themes for picture search
“Skin tone”
“Disability”
“Overweight”
“Housewife”
“Homosexuality”
“Fundamentalism, religion”
“Knowledge, science”
“Body image”
“Friendship, different skin tones”
“Prostitution”
“Age”

These themes were considered potential arenas for perceiving difference and strangeness.

The pre-selected pictures and constructed situations needed to be investigated systematically. For this purpose two slightly different questionnaires were developed. The questionnaires differ only in the direction of assigning characteristic words towards pictures and situations (questionnaire A) and assigning pictures and situations towards given words respectively (questionnaire B) (cf. appendix 9.2). As shown in table 10 these questionnaires were applied in different groups of the same high school. The first groups were asked to answer questionnaire A and the other groups were asked to fill out questionnaire B.

When analyzing the collected data, pictures and situations that differed dominantly from theoretical expectations were eliminated from the item pool or replaced with alternative pictures or situations respectively.

*Pre-study 3: Construction of a first structured version of the STSQ*

11 pictures and five situations represented the starting point for more concrete item formulations of the STSQ. In order to complete the construction of the STSQ as baseline for the discussions within a group of experts, I needed to prepare the following steps first:

- Formulation of item-answers (attributions, arguments) in relation to described situations and with respect to the pre-assumed structure of the STSQ, part II as indicated previously
- Development of a structured statement item pool (part III of the STSQ) in relation to the respective facets shown in my operational model of STS
- Reasonable scale selection for each part of the STSQ
- Construction of needed demographic items
- The formulation of a short and precise questionnaire introduction for all three parts of the STSQ
- The formulation of short and clearly formulated item instructions

- Reasonable arrangement of item order and questionnaire parts of the STSQ

Finally, the resulting STSQ was presented to the expert group in order to expose my questionnaire to a critical review from significant representatives of the underlying concept of IME.

### 6.5.3. Results and discussion

The pre-studies mainly supported the construction of relevant situations and reasoning/ argumentation strategies. In addition, these studies helped selecting pictures which were supposed to be related towards perception of difference and strangeness. In addition, the results of the pre-studies seemed to support basic assumptions of the construct as well. Described situations of the students were mainly related to experiences of being abroad, and how the respective new environments were perceived as strange. In addition, pupils reported from experiences they personally felt to be perceived as strange. Analyses of collected information pointed out general structures and mechanisms of strangeness related perceptions.

#### Main findings of pre-study 1:

The following characteristic terms and phrases were used by the pupils when they described and explained situations they personally perceived as strange:

*“I played basketball in the USA for the first time and I noticed that there were a lot of “black players” which I first perceived as threatening”*

*“Unusual training methods in football such as dance training”*

*“The only unknown party guest”*

*“Difficult to understand that a friend is not allowed by her parents to stay over for a night”*

*“Unexpected behaviour, comments or attitude of friends or family”*

*“When I entered a bus, I was the only person asked by the conductor to show my ticket”*

*“Different skills, competencies”*

*“Being like a boy in a ladies store”*

*“Looking different compared to the other guests at a party”*

*“Being unnoticed at a party”*

*“After a longer stay abroad, I notice more differences that I did before my journey”*

*“I was uncertain how to behave”*

*“I felt that I was not in control of the situation”*

*“Awkward”*

*“Unusual behaviour of other persons”*

*“New places”*

*“Difficult to orientate”*

*“The situation was difficult to handle”*

*“Difficult to interpret the behaviour of other persons”*

*“The only one with a different skin-tone which scared me at the beginning”*

*“[...] as a foreigner one is continuously made uncertain about who one actually is [...] it is difficult to have immigration background [...] natives seem to realize strangeness only if they will be confronted with strangers”*

*“At the beginning I felt threatened by so many black basketball players”*

*“I did not know the other people”*

*“People I do not know look at me”*

*“People do not realize that I am new”*

*“Everything was unfamiliar”*

In addition, pupils mentioned frequently characteristic strangeness feelings such as

*“confusing”, “angry”, “funny”, “surprised”, “insecure”, “afraid”, “unexpected”, “uncertain how to behave”, “difficult to control”, “unpleasant”*

However, analyses of discussions and writing tasks showed that most of the students had difficulties to explain their arising feelings. Yet, some students (N=5) were capable of explaining their personal experiences with difference and strangeness quite precisely and differentiated. Later discussions with their teachers showed that these pupils used to be reliable and interested students. In addition, these 5 students had discussed the phenomenon of strangeness in their German class before. Actually, one main chapter in their German teaching book is headlined with the topic “strangeness” (Andersen & Bali, 2000:125-258).

---

Main findings of pre-study 2:

Analyses of the assigning tasks and discussions of the picture items resulted in selecting thirteen pictures for part I of the STSQ. Picture B, K and O were excluded because the analyses showed that these items were difficult to classify with the given terms. In addition, 8 situations were analyzed. The analysis showed that all eight situations appeared to be classifiable with the given terms in both assigning directions (cf. appendix 9.2). Further considerations and revisions led to the 5 situations as described in chapter 6.4.

Main findings of pre-study 3:

Based on the results of pre-study 1 and 2, I constructed a provisional structured version of the STSQ consisting the three different parts of the STSQ and almost similar to the previously described version in chapter 6.4.

The experts within the field of IME suggested refinements of some items in order to point out its relation to the underlying concept of IME. Basically the suggestions were directed at refining the items linguistically. English is not my mother tongue and I developed the STSQ in English. The linguistic proof-readings were helpful in order to meet my intended meaning of the items in English.

To sum up, eleven pictures and five situations were selected for the item pool. Three different scales are supposed to measure related dimensions of the construct in part I of the STSQ. In addition, six reasons and/or attributions are refined and related to each situation. From a theoretical point of view, the differentiation between three rationality levels seemed to be optimal in situation 3 (cf. table 6). I tried to apply this structure as far as possible to all other situations in order to provide a certain degree of consistency. But it was not possible for all situations to follow this structure as differentiated as subsequent investigations demonstrated.

Part III of the STSQ counts finally 17 items structured according the in chapter 6.4.3 mentioned categories.

The so constructed initial version of the STSQ was the starting point for further item pool development. The next step was to collect analyze data with the STSQ as will be shown in the next chapter.

## 6.6. Pilot study: initial item analyses and exploration of reliability and validity hints

The first structured version of the STSQ was used for the following data collection (cf. appendix 9.8.1). Data was collected with two different convenient samples. The first sample was selected at a German university and mainly consisted of students in teacher education with sport as the main subject (N=86). Sport students were considered a relevant goal group in the future for applications STS and STSQ such as an on-the-job-training of physical education teachers.

A group of international students participating at the International Summer School in Oslo (ISS) was also available for data collection (N=24)<sup>41</sup>. Even though this international class was structured heterogeneously with respect to cultural background of the participants, this group was considered an interesting group for developing the STSQ items further because the participants joined teaching classes such as multicultural relations, and integration of minorities. These topics are related to STS. An application of the STSQ was expected to reveal some interesting validity hints by comparing both groups with each other. The resulting knowledge was supposed to help developing the items and summarizing item scores of the STSQ.

### 6.6.1. Aim

The main aim was to examine the applicability of the first item pool in general and get a first empirical feedback on how the constructed items were understood by a sample of respondents.

In addition, it was intended to develop item coding procedures. Even though it was not the major aim, the data analyses were supposed to generate first hints on reliability (internal consistency) and validity as well. But these hints were basically used to collect statistical information about the single items and adapt them according to the intended meaning. It was not intended to develop psychometric norms and standards at this early stage of developing the STSQ. But some information about reliability was considered supplying information for item revisions.

---

<sup>41</sup> Participating students came from various African countries, east and west European countries, and Asian countries.



### 6.6.2. Pre-considerations of the research design and methods

The investigations of this study were mainly designed as an explorative study. The results were interpreted on the baseline of the underlying theoretical background presented in previous chapters. Discovered confirming or contradictory results were followed up with the aim of obtaining information about how the items needed to be revised with respect to measuring the intended meaning. It was planned to use simple statistical procedures. Descriptive statistics of the items, correlations between single items and explorative factor analysis were applied to generate the required information in the first place.

As indicated when I introduced the STSQ, it was not expected to achieve high Cronbach's alpha values of related items within the same facet because the items of the same facets are not parallel items in the sense as required for using Cronbach's alpha (cf. Cronbach & Shavelson, 2004). But knowing this, estimating internal consistency by Cronbach's alpha can indicate how far the items are related to each other. The dominance of the indirect measuring procedures within the STSQ was also expected to reduce consistency estimations with Cronbach's alpha (cf. Rost, 2004).

**Table 12: Demographic data (pilot study)**

Sample		German students	ISS students
Total		86 (22)	24
Age	Mean	≈ 24	≈ 33
	Minimum	20	23
	Maximum	38	70
Sex	Male	40 (47%)	10 (42%)
	Female	45 (52%)	14 (58%)
	Missing values	1 (1%)	–
Abroad experience <sup>42</sup>	Yes	37 (43%)	14 (58%)
	No	49 (57%)	6 (25%)
	Missing values	–	4 (17%)

<sup>42</sup> Students who have been abroad at least more than two months

Because of technical problems at the German University, 64 of 86 questionnaires were duplicated in more or less bad quality. This affected the results of the first part of the STSQ because the presented pictures were in some cases difficult or only partial recognizable for the respondents. Analyses of part II and III of the STSQ could be conducted with the whole sample.

### 6.6.3. Expectations

The STSQ is a new and complex questionnaire. To deal with this complexity in a systematic way can be challenging and even overwhelming. Especially, explorative analyses can easily become confusing and difficult to handle. I therefore formulated my expectations about the intended relationships between certain variables of the STSQ before presenting the procedures and results in detail. This as a starting point should make the dominantly explorative character of my analyses more structured. The following expectations were extracted from my theoretical considerations (cf. part I of my thesis) and the intended meaning of the STSQ. I will start with part one of the STSQ:

Three different scales were used in part I of the initial version of the STSQ:

1. Closeness scale
2. Dissimilarity scale
3. Sympathy scale

A higher score on the closeness scale indicates a higher attributed meaning than lower scores. The dissimilarity scale measures further the direction of the attributed meaning. Consequently, an indicated attribution of meaning is specified by the dissimilarity scale.

The main function of the dissimilarity scale is to measure how different a shown situation or person is perceived. The intention of combining both scales (dissimilarity and sympathy) into one summarizing score functions as indicator for an emotional STS (cf. Figure 20). The more often the pictures are perceived as dissimilar, the more suitable is the item for evaluating an emotional valuation (attribution of sympathy) of perceived difference.

Theoretically, it seemed interesting to investigate a further assumed relation between the closeness and dissimilarity scale. Lower scores on the closeness scale indicate a lower attributed meaning (e.g. score value 1 or 2 on the closeness scale). Score 3 is understood as a medium or neutral position with respect to an indicated dissimilarity. The score 3 on the dissimilarity scale is supposed to indicate that a respondent either perceives a picture item as dissimilar or as similar to one's own perception which would indicate that the item appears to be irrelevant, meaningless or in Simmel's terminology 'distant'. If such a tendency can be supported

empirically, this matter of fact would support the validity of the dissimilarity scale in the way presented in the operationalization chapter. Such a result would also suggest that the dissimilarity scale (perceived difference) may represent the closeness scale in a sufficient way. The number of scales would therefore be reduced from three to two scales by excluding the closeness scale from the item pool.

The following expectations were examined accordingly:

1.1: The “closeness scale” was expected to be understood inverted by the respondents in relation to the “dissimilarity scale”. Consequently, these two scales should usually correlate negatively with each other.

1.2: The “closeness scale” was expected to correlate positively with the “sympathy scale”.

1.3: With the purpose of measuring an emotional STS, suitable picture-items needed to be perceived as different. Consequently, the results of the “dissimilarity-scale” needed to show an asymmetry from normal distribution: negative “skewness index”.

It can be argued that close persons may be perceived as more sympathetic than distant ones (1.2). If 1.2 can be supported by empirical results, it is reasonable to examine if the dissimilarity scale correlates negatively with the closeness scale. This would support additionally the decision to exclude the closeness scale from the item pool because measurements with the dissimilarity scale seem to include the intended meaning of the closeness scale.

Part II of the STSQ was examined according to the following expectations:

2.1: Items of the same facet/category within a situation should correlate according theoretical assumptions. But it was not expected to achieve high correlation values as one could expect when analyzing parallel items of within a factor model.

2.2: Item scores of the same facet should correlate between situations in the theoretically assumed way.

The item pool structure of part III was explored by factor analysis. The results were expected to summarize the single items and generate representative item scores. In this way, EFA was applied for discovering the empirical structure of the item pool and label these scores in accordance with my theoretical considerations.

The strict comparison of discovered empirical structure with theoretically assumed structure was supposed to gain first validity hints (construct validity). Comparisons (correlations) between items of different parts of the STSQ were also expected to gain some hints on validity.

#### 6.6.4. Procedure

The two independent samples (Germany Sport students and international students in Norway) were analyzed separately.

The first step was to examine if the item instructions and scale labels were understood by the respondents in the intended manner. For this purpose, the respondents' personal feedback during the data collection, such as questions or expressed difficulties in understanding particular item instructions, was supposed to help pointing out these general error influences. All relevant comments of the respondents during data collection were documented. Resulting information were useful with respect to enhance the clarity and appearance of the items (cf. Downing & Haladyna, 1997).

Internal constancies of the scales in part I of the STSQ were evaluated with Item-Total correlations and/or Cronbach's alpha. Non-parametric correlations were used to compare item scores within and between the different parts of the STSQ. The generation of the mentioned eSTS pattern implied a reduction of the scale quality. Therefore, non-parametric procedures were preferred for pattern score analyses. EFA was applied for investigating the empirical structure of the different parts of the STSQ and compared the results with the theoretically suggested structure.

However, all my studies within this thesis were based on relative low sample sizes. Lower sample sizes imply that it is more difficult to achieve optimal prerequisite of the data material such as optimal sample size, normal distribution and linearity for applying parametric statistical procedures. In accordance with Bortz and Lienert (2003) I therefore preferred non-parametric procedures<sup>43</sup> for most of my analyses (cf. chapter 6.2).

---

<sup>43</sup> A comparison of the results were calculated with both parametric and non-parametric correlations showed that both methods indicated similar tendencies, but the values based on non-parametric calculations were usually lower than the result of parametric procedures.

Explorative factor analysis over all items of part II of the STSQ did not lead to interpretable results. The item structure between all situations might have been too complex and the scale qualities too low. In addition, the samples sizes were quite low as well which made interpretations of EFA even more uncertain. It seemed therefore more reasonable to analyze the situations separately. Spearman correlations were used to investigate the relationships of items separately within each situation.

Explorative factor analysis and correlations of items were applied for exploring the item structure of part III. If resulting factor analysis led to an interpretable structure, internal consistency of the resulting scale items of each facet was estimated.

#### 6.6.5. Results

When filling out the STSQ, respondents frequently asked how the term closeness is to be understood. They had similar difficulties with completing the dissimilarity scale. But after explaining the meaning of these two scales the group seemed to be capable to complete all three scales.

#### *Results part I of the STSQ:*

The following Table 13 shows that the two different samples achieved relative similar Cronbach's alpha values on all scales.

**Table 13: Results of scale analysis (internal consistency), part I, STSQ v. 1, pilot study**

Sample	N <sub>Total cases</sub>	N <sub>valid cases</sub>	Scale	N of items	Cronbach's alpha
Sport students from Germany	22	20	Closeness	11	,79
International students in Oslo	24	15	Closeness	11	,77
Sport students from Germany	22	19	Dissimilarity	11	,64
International students in Oslo	24	13	Dissimilarity	11	,79
Sport students from Germany	22	21	Sympathy	11	,76
International students in Oslo	24	14	Sympathy	11	,67

Considering that part I is a semi-projective device, the consistency result appeared to be sufficient in order to differentiate between groups (according to Lienert & Raatz, 1998).

The next results in Table 14 show correlations between the different scales of part I of the STSQ. The reduced German sample (N=21) was used for the calculations. The correlation matrix shows a moderate and significant positive correlation between the closeness and sympathy scale. The table shows a moderate and significant negative correlation between dissimilarity and closeness, and between dissimilarity and sympathy.

**Table 14: Correlations (Spearman's rho) between closeness, dissimilarity and sympathy scale, N=21, German sample, good picture qualities**

	1	2	3
1. Closeness	--		
2. Dissimilar	-.55*	--	
3. Sympathy	.59**	-.49*	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

The negative correlation between dissimilarity and closeness is as expected. The positive correlation between sympathy and closeness seems also reasonable and supports the idea that more close persons are attributed more sympathy than less close persons.

#### *Results part II of the STSQ:*

The following five tables summarize the results of the item analyses for each situation separately (cf. chapter 6.4.2 for item coding used in the following tables):

Situation 1:**Table 15: Spearman correlation results of situation 1 part II of the STSQ (N=86, German sample)**

Item	1.1	1.2	1.3	1.4	1.5	1.6
1.1	--					
1.2	,21	--				
1.3	<b>,35**</b>	<b>,32**</b>	--			
1.4	,1	,19	,03	--		
1.5	<b>-,24*</b>	-,05	-,16	,06	--	
1.6	,20	-,06	,11	,18	-,06	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*.. Correlation is significant at the 0.01 level (2-tailed).

All correlations are quite low. The RA1 items 1.1, and 1.2 correlate significantly with the RA1 item 1.3. Item 1.5 was assigned to the RA2 category and correlate negatively with item 1.1 of the RA1 category. Since these two items were assigned to different categories (RA1 and RA2) the negative relationship seem to indicate a distinction.

Situation 2:**Table 16: Spearman correlation results of situation 2 part II of the STSQ (N=86, German sample)**

Item	2.1	2.2	2.3	2.4	2.5	2.6
2.1	--					
2.2	<b>,46**</b>	--				
2.3	<b>,43**</b>	<b>,39**</b>	--			
2.4	<b>,27*</b>	,07	<b>,32**</b>	--		
2.5	,20	<b>,41**</b>	<b>,34**</b>	<b>,33**</b>	--	
2.6	,03	-,08	-,06	-,14	-,08	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*.. Correlation is significant at the 0.01 level (2-tailed).

The results shown summarized in Table 16 support parts of the intended item structure. The strongest relationship within this situation is indicated by moderate and significant correlations between the items 2.1., 2.3, 2.4, and 2.5, and Item 2.5 was assigned towards RA2 level but correlates with items of RA1 level. Based on this result, the items related to this situation did not differentiate between the two theoretically intended rationality levels RA1 and RA2 (cf. Table 5).

Situation 3:**Table 17: Spearman correlation results of situation 3 part II of the STSQ (N=86, German sample)**

Item	3.1	3.2	3.3	3.4	3.5	3.6
3.1	--					
3.2	<b>,53**</b>	--				
3.3	-,21	,14	--			
3.4	,14	-,14		--		
3.5	<b>-,40**</b>	<b>-,24*</b>	,03	-,03	--	
3.6	,01	-,14	,12	,06	,11	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\* . Correlation is significant at the 0.01 level (2-tailed).

The negative correlation between 3.1 and 3.5, and 3.2 and 3.5 shown in the Table 17 indicate a reasonable relationship between items of RA1 and RA3 level. The significant positive relation between item 3.1 and 3.2 indicate that these two items can be differentiated from the other items in accordance to the theoretical expectations.

Situation 4:**Table 18: Spearman correlation results of situation 4 part II of the STSQ (N=86, German sample)**

Item	4.1	4.2	4.3	4.4	4.5	4.6
4.1	--					
4.2	<b>,26*</b>	--				
4.3	,04	,05	--			
4.4	-,20	-,17	,19	--		
4.5	0,10	<b>,31**</b>	,16	,02	--	
4.6	,11	,05	,00	,02	,07	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\* . Correlation is significant at the 0.01 level (2-tailed).

In addition to the results shown in Table 18 showed the analysis of the international sample (N=24) significant positive correlation ( $\rho=0,52$ , at a significance level of  $p<0,05$ ) between item 4.2 (RA1 item) and 4.6 (RA3 item). The correlations shown in Table 18 demonstrate that the item did not show a differentiated structure as theoretically intended. The indicated correlation between the two RA 1 items 4.1 and 4.2 is consistent with the intended relationship but also here the correlation is significant but quite low.



Situation 5:**Table 19: Spearman correlation results of situation 5 part II of the STSQ (N=86, German sample)**

Item	5.1	5.2	5.3	5.4	5.5	5.6
5.1	--					
5.2	<b>-,23*</b>	--				
5.3	<b>-,36**</b>	<b>,39**</b>	--			
5.4	,04	,00	,04	--		
5.5	,16	-,01	-,18	<b>,44**</b>	--	
5.6	<b>,23*</b>	<b>,27*</b>	-,06	<b>,26*</b>	,17	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\* . Correlation is significant at the 0.01 level (2-tailed).

In addition to the results shown in Table 19, the results of the international sample (N=24) indicated significant and positive correlation between item 5.2 and 5.3 ( $\rho=0,59$  at a sig. level  $p<0,01$ ) and a negative correlation between item 5.1 and 5.3 ( $\rho=0,56$  at a sig. level  $p<0,01$ ).

The next table shows correlations between summarized item scores (RA1, RA 2, RA3). The respective item scores were summarized within each situation according to the theoretically assumed structure and compared with each other between the situations by correlations. Considerations of the reasonability of the following correlation results were supposed to gain validity hints of the interpretability of the RA scores.

The results shown in Table 20 are structured according to a) validity supporting results b) unclear results and c) contradictory results.

**Table 20: Results of correlations (Spearman's rho) of RA1, RA2 and RA3 scores between each situation, part II of the unedited/initial version of the STSQ, N=86 (German sample)**

Results
<u>Supportive results for the validity:</u>
<ul style="list-style-type: none"> <li>• The RA 1 score of situation 1 correlates positive with RA 1 of situation 2 (.49**), with situation 3 (.22*), Unpleasant perception score (5.4+5.5+5.6) of situation 5 (32**)</li> <li>• RA1 score of situation 2 correlates low but significant with RA1 of situation 4 (.29**)</li> <li>• The RA 2 score of situation 1 correlates positive with RA2 of situation 4 (.22*)</li> <li>• The RA 3 score of situation 3 correlates negative with RA 1 score of situation 3 (-.23*).</li> </ul>
<u>Unclear results:</u>
<i>RA 1 scores correlate with RA 2 scores:</i>
<ul style="list-style-type: none"> <li>• RA1 sit. 1 correlates positively with RA2 sit. 2 (.31**) and RA2 sit.4 (.38**)</li> <li>• RA1 sit. 2 correlates positively with RA2 sit. 2 (.29**), and RA2 sit. 3 (.3**)</li> <li>• RA 2 sit. 1 correlates positively with RA 3 sit. 3 (.36**)</li> </ul>
<u>Contradictory results:</u>
<i>RA 1 scores correlate positively with RA 3 scores:</i>
<ul style="list-style-type: none"> <li>• RA 1 situation 1 correlates positively with RA 3 sit. 1 (.38**)</li> <li>• RA 1 situation 2 correlates positively with RA 3 sit. 2 (.34**)</li> <li>• RA 1 situation 3 correlates positively with RA 3 sit. 4 (.26*)</li> </ul>
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

The RA 1 score shows most empirical support by indicated reasonable correlations between the situations. The results above indicate, however, that it was difficult to differentiate between the intended three different rationality levels. Only within situation three the scores seem to differentiate reasonable between RA1 and RA3 (indicated by a negative correlation).

Probably because of the aforementioned problem applying EFA with small sample sized and too heterogeneous items, results of EFA were difficult to interpret and did not lead to new knowledge about the item structure in comparison to the aforementioned results. These results of EFA are shown in the Table 21.

**Table 21: Factor matrix of the RA sum scores between the situations, N=86 (German sample)**

Item	Factor		Communality
	1	2	
RA1sit1	.72	-.17	.54
RA1sit2	.80	-.00	.65
RA1sit3	.25	-.37	.20
RA1sit4	.39	-.25	.21
RA1sit5	.27	-.1	.08
RA2sit1	.07	.31	.10
RA2sit2	.40	.08	.17
RA2sit3	.24	.06	.06
RA2sit4	.18	.03	.03
RA3sit3	.08	.9	.81
RA3sit4	.46	-.14	.23
RA3sit5	.3	.13	.11
Eigenvalues (initial)	2,0 (2,7)	1,2 (1,6)	
% of variance (initial)	16,7 (22,2)	9,9 (13,6)	

Extraction Method: Principal Axis Factoring.  
Varimax rotation

The factor results indicate that the RA1 scores of situation 1 and 2 loads clearly on this “factor”. On the other hand, RA3 of situation 3 shows the highest and only clear factor loading on “factor” 2. All other factor loading are relative low and consequently difficult to assign.

However, RA3 of situation 3 loads highest on factor two. This clear factor loading (0,9) helps to consider in how far the other items are related (or not) to RA3 (situation 3) in a reasonable way. The RA1 scores load low and/or negative on factor two. Assuming that RA3 of situation 3 indicate more rational arguments in comparison to RA1 scores, the factor loadings on factor two seem (to some extend) to distinguish between the RA1 and RA3.

RA1 of situation 4, RA2 of situation 2 and RA3 of situation 4 show their highest loading on “factor” 1 but these loadings are relative low (around 0,4). Since RA1 of the situation 1 and 2 dominate loadings on “factor” 1 the factor loading of RA1 of situation 4 can reasonably related to RA1sit1 and RA1sit2. But RA2sit2 and RA3sit4 load also with a similar value on the same factor. This makes an interpretation difficult. Similar to the previous correlations, EFA showed that the RA scores did not differentiate between all three levels.

*Results part III of the STSQ:*

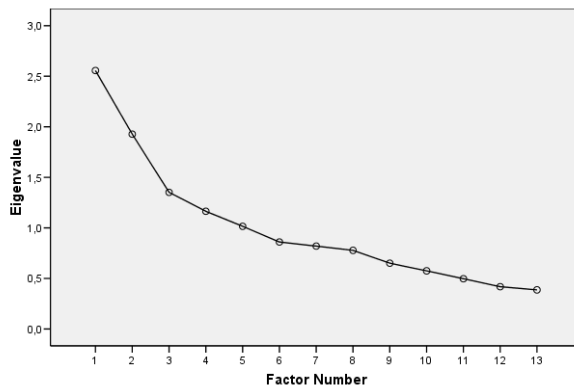
The structure of the final part III of the STSQ is analyzed with correlation and explorative factor analyses. Calculated correlations (Spearman's rho) between the first two items EC 1 and EC 2 with the German sample ( $N_{\text{Total}}=86$ ) showed that these two items either correlate significantly with each other or with other items of the item pool of part III.

The results of EFA were also difficult to interpret but some indicators were discovered which seemed to support the theoretically assumed item structure of part III but here also the explained variance was quite low.

The following parameters for factor analysis are set for the calculation:

- Extraction method: principal axis factoring,
- Rotation method: Varimax with Kaiser Normalization
- $N_{\text{Total}}=86$
- Number of items: 13 (items EC1, EC2, are excluded from analysis because these items showed low or no inter correlations; LOC 1 is excluded because of low communalities and low factor loadings)
- Extractions: graphical Scree plot analysis

The following Scree plot is interpreted as 2-factor solution for this item pool.



**Figure 27: Scree plot for the item pool analysis of part III, STSQ v.1**

The following Table 22 shows that both factors explain ca. 23% of the variance after rotation. This result was expected because of the intended heterogeneity of item pool.

**Table 22: Resulting factor loadings for the item pool analysis of part III, STSQ v.1**

Item	Factor		Communality
	1	2	
O 1 "It seems that different people have different values"	<b>,58</b>	,23	,38
O 2 "The open result is responsible for the excitement of a competition"	,30	,30	,17
O 3 "To interact with different kinds of people is enjoyable"	<b>,47</b>	,04	,23
O 4 "Sitting in a group of strangers provokes uncomfortable feelings"	<b>-,40</b>	,01	,16
O 5 "There is no freedom without rules"	,12	<b>,50</b>	,26
NFS 1 "It is important to forward cultural values to the next generations"	,22	<b>,50</b>	,30
NFS 2 "It is necessary to plan ahead in order to avoid surprises"	,16	<b>,55</b>	,33
NFS 3 "Being obliged to make decisions provokes uncertainty"	-,17	<b>,32</b>	,13
NFS 4 "For activities, I prefer precise instructions to open suggestions"	-,28	,35	,20
LOC 2 "Most of the events in my life are determined by other people"	<b>-,43</b>	,05	,18
LOC 3 "Fate determines whether I have more or fewer friends"	<b>-,47</b>	-,01	,22
LOC 5 "In spite of its apparently continuous change, culture contains some stable values"	<b>,44</b>	,10	,20
LOC 6 "Unforeseen events upset me"	-,30	,33	,2
Eigenvalues (initial)	2,6 (1,9)	1,3 (1,9)	
% of variance (initial)	13 (19,7)	9,9 (14,8)	

Extraction Method: Principal Axis Factoring.  
Varimax rotation

A careful interpretation of the factor loadings indicate that some items load on the same factor in a plausible way. O1, O3 and O4 load on the same factor which was considered as reasonable because all three items are openness items. In addition, the LOC2, LOC 3 are translated items of a German locus of control scale indicating external locus of control (Krampen, 1981). In order to point out that these items were applied in a different context (strangeness), I labeled these items as "loss of control. These items load negative on the same factor together with the openness items. This relationship can be interpreted as a supportive validity hint (regarding construct validity) for the openness items in line with my theoretical considerations around strangeness and identity theory. A self-perception that the events in one's life depend more on external factors represent an instable factor of one's self-concept and is an indicator for uncertainty. The implicit uncertainty when meeting strangers can therefore be perceived as unpleasant and people may tend to avoid such a situation because they might be more intolerant towards meeting uncertain situations.

However, I examined the items of each facet further of its internal consistency estimated by Cronbach's alpha. Table 23 shows more detailed item-total correlations. The selection of items were based on the before shown factor structure. Even though item O2 shows similar and weak factor loadings on both factors, it did not affected Cronbach's alpha that much if it were excluded from the scale (cf. Table 23). In addition, the results show a relative low item-total correlation. Item LOC2 and LOC3 also show a quite low item-total correlation.

**Table 23: Item-Total Statistics of factor 1 ( $\alpha_{\text{Cronbach}} = 0,62$ ), N=86, German sample**

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
O 1 "It seems that different people have different values"	18,03	4,19	,48	,34	,54
O 2 "The open result is responsible for the excitement of a competition"	18,55	4,7	,26	,09	,61
O 3 "To interact with different kinds of people is enjoyable"	18,22	4,15	,38	,32	,57
O 4* "Sitting in a group of strangers provokes uncomfortable feelings"	19,00	4,21	,34	,19	,58
LOC 2**"Most of the events in my life are determined by other people"	18,67	4,2	,3	,14	,6
LOC 3* "Fate determines whether I have more or fewer friends"	18,66	4,26	,28	,14	,6
LOC 5 "In spite of its apparently continuous change, culture contains some stable values"	18,66	4,6	,33	,18	,59

\* Data is recoded in relation to the openness scale meaning

Following Lienert and Raatz (1998) a reliability coefficient between  $r=0,5-0,7$  is sufficient in order to differentiate between groups. The internal consistency estimations also illustrate another aspect. The communalities in factor analyses are often interpreted as an estimate for reliability. But according to Bühner (2003) low communality does not necessarily imply low reliability. This is one of the reasons that I did not strictly exclude unfitting items according to cross loading or low communalities. On the other hand, it cannot be denied that the items with low communality and low item-total correlation do not represent optimal items for the item pool. Therefore, these items need to be revised and adapted so that they afterwards may fit better into the respective scale. This aspect illustrates a central idea of my intended procedure of developing my derived item pool further by collection empirical information showing which items need to be revised in a way that it fits better the intended meaning after the revisions were made. A typical standard procedure would have been to develop a larger item pool, excluded all unfitting items in order to achieve highest psychometric norms and standards. However, this is not the case here. The STSQ did not include enough items to follow such a

procedure. It was intended to carefully derive items out of theoretical considerations and support the constructed items with rational arguments and then develop them further by collecting hints from empirical data about if and how the items needed to be changed in order to meet the intended meaning.

The items of factor two showed the following scale characteristics (Table 24):

**Table 24: Item-Total statistics of factor 2 ( $\alpha_{\text{Cronbach}} = 0,57$ ), German sample,  $N=86_{\text{Total}}$**

	Scale Mean if Item De- leted	Scale Vari- ance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
O 5 "There is no freedom without rules"	12,80	3,52	,39	,19	,49
NFS 1 "It is important to for- ward cultural values to the next generations"	12,63	3,8	,30	,18	,53
NFS 2 "It is necessary to plan ahead in order to avoid sur- prises"	13,16	3,45	,39	,18	,49
NFS 3 "Being obliged to make decisions provokes uncer- tainty"	13,28	3,72	,26	,11	,54
NFS 4 "For activities, I prefer precise instructions to open suggestions"	13,34	3,7	,25	,07	,55
LOC 6 "Unforeseen events upset me"	13,52	3,82	,27	,10	,54

Consistency estimations with the international sample ( $N=24$ ) led to almost similar Cronbach's alphas in comparison to the German sample ( $N=86$ )<sup>44</sup>.

A comparison of the STSQ mean scores between the German and international sample showed no statistically significant differences<sup>45</sup> (cf. Table 25).

<sup>44</sup> Factor 1:  $\alpha_{\text{Cronbach}} = 0,64$  and factor 2:  $\alpha_{\text{Cronbach}} = 0,5$

<sup>45</sup> Mann-Whitney test

**Table 25: STSQ mean scores separated by the two samples: German sample (GER) N=86 and international sample (Int.) N=24**

STSQ score	German sample		International sample	
	Mean	SD	Mean	SD
eSTS pattern	0,2	0,17	0,13	0,15
RA1	2,57	0,41	2,41	0,35
O	3,13	0,36	3,09	0,42
LOC	2,17	0,38	2,03	0,46
O LOC	3,07	0,32	3,07	0,35
NFS	2,64	0,4	2,82	0,48

A comparison of the STSQ scores with each other indicated low but significant correlations. O correlated negative with the NFS score ( $\rho = -0,21$  with a sig. level of  $p < 0,05$ ). Even though the correlation is low, this relationship seems plausible in line with the theoretical concept. A need for security may reduce an openness in a similar way as loss of control was assumed to reduce a willingness to “take a risk” of meeting strangers. The RA1 score showed also a negative correlation with the NFS score ( $\rho = -0,23$  with a sig. level of  $p < 0,05$ ).

#### 6.6.6. Discussion and conclusions

Being aware of the heterogeneity of the items and the low sample size, EFA and Cronbach’s alpha were probably not optimal procedures for analyzing the empirical structure of the STSQ. Kleven (2008) points out that Cronbach’s alpha originates from a psychometric tradition which basically was interested in individual differences. Individual diagnostic with psychometric measuring procedures require high norms and standards as the term “psychometric” suggests. Regarding the aims in the context of my project I did not require high psychometric standards in the first place as pointed in my pre-considerations in chapter 6.1 and 6.2.

There are three or probably four different uncertainties to deal with, the construct’s validity, the STSQ’s validity and the appropriateness of statistical procedure in relation to the sample and the type of measuring instrument, and errors linked to the selected samples. Validity was basically understood according to the following quote:

“Validity is a property of inferences, and the relevance of various types of validity depends of what kinds of inferences are drawn, not on what kind of methods used to collect the data.” (Kleven, 2008:220; cf. also Messick, 1995).

This suggests discussing validity aspects with reference to the most relevant types of validity according to Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). My investigations refer basically to construct validation. I tried to collect empirical hints which were supposed to help considering how the indicators (items) needed to be adapted in order to represent the



---

intended meaning (of each single item as representative indicator for different facets of the construct). But another aspect was considered regarding the appropriateness of the applied statistical procedures. This discussion is more related to “statistical-validity” (cf. Kleven, 2008) or statistical conclusion validity which refers to considerations whether a tendency should be considered trivial or worthy of an interpretation. Some of these aspects were discussed with the presentation of the results.

Part I of the STSQ was designed as semi-projective method. The feedback from the respondents showed that it was difficult for them to understand the closeness scale. The correlations between the used three scales supported the assumption that the closeness dimension can be measured indirectly with the two other scales (cf. Table 14). Consequently, the closeness scale was excluded from the item pool of part I for the next data collection.

In addition, strangeness starts with the perception of a difference. The dissimilarity scale appeared to be more relevant in order to measure an emotional STS compared to the closeness dimension.

The results indicated that the sympathy scale was understood as theoretically indented and was consequently kept as an indicator for measuring an emotional dimension (sympathy) attributed towards the shown picture item.

Respondents indicated during data collection that the dissimilarity scale was frequently understood as similarity scale. Therefore the task instructions needed to be revised in order to make the intended meaning of the item task clearer. Instead of using the term “dissimilarity”, “familiarity” might be the more precise terminology because “familiarity” is understood as the opposite of “strange”. The term strangeness as scale label was avoided on purpose because it might provoke a bias such as political correct answers. Using the counterpart terminology of strangeness was more reasonable in order to avoid the underlying construct’s name directly as an indicator. Discussions with English native-speakers showed that the term “familiarity” is probably easier to evaluate than the term “strangeness”. Figure 28 illustrates all revisions based on the resulting knowledge of the results.

<p>1 a) What comes first to your mind when you see the following picture? Please describe your initial associations!</p> <div style="border: 1px solid black; width: 60px; height: 60px; margin: 10px auto; text-align: center; padding: 5px;">           Picture stimulus         </div> <div style="margin-left: 100px;"> <hr style="width: 100%;"/> <hr style="width: 100%;"/> <hr style="width: 100%;"/> </div> <p>1 b) Please evaluate your feelings towards the picture above on each of the following scales! Indicate your subjective relationship by crossing out the respective number: 5 = highest value; 1 lowest value!</p> <p>I feel ...</p> <table style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td style="width: 10%;">1.</td> <td style="width: 40%;">close</td> <td style="width: 5%;">1</td> <td style="width: 5%;">2</td> <td style="width: 5%;">3</td> <td style="width: 5%;">4</td> <td style="width: 5%;">5</td> </tr> <tr> <td>2.</td> <td>dissimilar</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>3.</td> <td>sympathetic</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </tbody> </table> <p>... towards the picture above.</p>	1.	close	1	2	3	4	5	2.	dissimilar	1	2	3	4	5	3.	sympathetic	1	2	3	4	5	<p>1 a) Please give a headline to the situation shown in the following picture!</p> <div style="border: 1px solid black; width: 60px; height: 60px; margin: 10px auto; text-align: center; padding: 5px;">           Picture Stimulus         </div> <div style="margin-left: 100px;"> <hr style="width: 100%;"/> <hr style="width: 100%;"/> <hr style="width: 100%;"/> </div> <p>1 b) Please indicate how familiar (1.) and how sympathetic (2.) you feel imagining the given situation in the picture above. Please cross out the respective number:</p> <p>1. Familiarity:      low <span style="border: 1px solid black; padding: 2px 5px;">1</span> <span style="border: 1px solid black; padding: 2px 5px;">2</span> <span style="border: 1px solid black; padding: 2px 5px;">3</span> <span style="border: 1px solid black; padding: 2px 5px;">4</span> <span style="border: 1px solid black; padding: 2px 5px;">5</span> high</p> <p>2. Sympathy:      low <span style="border: 1px solid black; padding: 2px 5px;">1</span> <span style="border: 1px solid black; padding: 2px 5px;">2</span> <span style="border: 1px solid black; padding: 2px 5px;">3</span> <span style="border: 1px solid black; padding: 2px 5px;">4</span> <span style="border: 1px solid black; padding: 2px 5px;">5</span> high</p>
1.	close	1	2	3	4	5																
2.	dissimilar	1	2	3	4	5																
3.	sympathetic	1	2	3	4	5																

**Figure 28:** Changes and adaptations of the items construction of part I of the STSQ. On the left side, an item of the STSQ version 1; on the right side, the item after revisions according to the results of this pilot study.

The first part of the item was initially formulated as an open question about what respondents associate with the shown picture. The intention of changing the instructions was to induce that the respondents have dealt more intensively with the shown picture. Analysis of this open item task showed that respondents wrote down a lot of different associations. In order to make the item task more efficient, the instructions were changed. The respondents were now asked to give a headline to each picture item instead of writing down their associations when imagining the illustrated situation or person on the picture-items. A headline requires more reflections compared to free associations. A headline may therefore indicate a more representative attributed meaning of the respondents. Associations could easily be more influenced by situational factors and could therefore provoke inconsistent responses of this item over time.

Because of the above mentioned technical difficulties during data collection in Germany, a larger part of this sample filled out the STSQ with “bad picture qualities”. I decided therefore to keep all pictures for the next data collection in order to control these single items and scale characteristics with an appropriate picture quality.

The first results of scale consistency measured by Cronbach’s alpha showed suitable results at this early stage of STSQ’s development especially when considering that projective measuring devices usually show lower reliability coefficients (cf. Rost, 2004). The above mentioned adaptations were expected to increase the reliability.

The second part of the STSQ was also designed as a semi-projective device but now with situations instead of pictures as projection stimulus. It was attempted to follow a pre-defined item structure within all five situations in a comparable way. The items of the same category

were related to the items of the same category nested within different situations. The relationships of items of the same category were expected to show a significant correlation.

The results of the factor analysis over all items did not support the intended complex structure of the item pool between the situations. Besides sample size, the heterogeneity of a sample made interpretations of factor loadings difficult.

Nevertheless, the correlation results of each situation pointed out that the RA1 score appeared to be clearest within each situation. The comparison of the different RA mean scores between the situations indicate a reasonable (negative) correlation between RA1 and RA 3 score (Table 20 and 21). But it was not able to distinguish between the intended three different levels of RA.

The last part of the STSQ was constructed as a statement item pool. The extracted two factors of the conducted explorative factor analysis were to some extent interpretable on the basis of my theoretical concept. The first factor appeared convergent with the concept of “openness” because the openness-items dominate this factor; O1 loaded highest with 0,58 on factor 1. The second factor included items of the theoretically intended facet NFS; item NFS1 (0,50) and NFS2 (.55) (cf. Table 22). As previously indicated, the relative low factor loadings and explained variance was considered to be a result of the heterogeneity of the item pool.

Some of the LOC items loaded negatively on factor 1 together with the O-items. These negative correlations of the same factor could be interpreted. Being open-minded in strangeness related situations requires a certain degree of self-confidence; otherwise implicit uncertainty may become difficult to tolerate and may have a negative effect on a person’s open-mindedness (cf. Buhr et al., 2002).

Contrary to what was expected, item O5 loaded on the second factor. This item seemed to be misunderstood by the respondents, because it correlated (Spearman correlation) more with NFS items (NFS 1, 2 and 3) than with O items. Item O5 correlated significantly with item of part II of the STSQ: item 1.1, 1.3, and 2.3. Those items were supposed to indicate a kind of stigmatizing attributions or arguments (RA1). The correlations of O5 with RA1 items seemed to indicate that the respondents understood item O5 as too determinant or restrictive. Item O5 needed to be reformulated to emphasize the meaning that people need rules in order to insure freedom (cf. Table 26).

LOC 6 loaded on both factors in a reasonable way (factor 1: -.30, and on factor 2: +.33). The slightly higher factor loading and the better reliability result suggested assigning the

LOC6 item to the second factor. Both facets, LOC and NFS are related to uncertainty and appeared to be related as Table 21 illustrates.

An internal consistency analysis (item total correlations) of the resulting factors showed satisfying first results taking into consideration that the items are heterogeneously structured. Reliability results showed that Cronbach's alpha had almost the same values for both scales when estimating the two samples separately. In both samples the items of the second factor achieved lower Cronbach's alpha compared to the scale items of the first factor. This result seems to be caused by the lower number of items within factor two.

The following table summarizes the changes and adaptations of items in part III of the STSQ according to the results of item analyses.

**Table 26: Revision of items according to the resulting knowledge of this pilot study (STSQ v.2, part III)**

<b>Item</b>	<b>STSQ v.1</b>	<b>Revised item</b>	<b>Comment/reason for revision</b>
EC 1	<i>Other people are not as open-minded as I am.</i>	O1.1: Open-mindedness helps understanding.	Changed to a further O-item in order to point out the O-facet
EC 2	<i>My personal way of life should be a model for other people.</i>		
O3	<i>To interact with different kinds of people is enjoyable.</i>	O3.1: To meet different kinds of people is enjoyable.	Linguistic revisions
O5	<i>There is no freedom without rules.</i>	O5.1: Freedom does not mean the absence of rules.	Seemed difficult to understand in the first version
NFS2	<i>It is necessary to plan ahead in order to avoid surprises.</i>	NFS2.1: To plan ahead helps to provide security.	Linguistic revision in order to point out the meaning
LOC4	<i>In spite of individual differences people do not differ substantially.</i>	LOC4.1: People differ substantially	This change should make the item content clearer (easier to understand).

The results of all parts were presented and discussed separately. At the end of each discussed part, suggested adaptations of instructions, single items, and/or score keys were summarized and supported with arguments. Item changes were followed up during the next steps of development.

Firstly, the scales of part I were reduced and changed into two scales: familiarity and sympathy. Considering the circumstances of such a type of measuring method (semi-projective), sampling procedures/ types of samples and stage of development, first reliability estimations showed satisfying results of part I. Cronbach's alpha varied between 0,6 and 0,8 (Lienert & Raatz, 1998).

Correlations indicated reasonable relationships between the different scales. The whole picture pool was kept for the next data collection in order to test the effects of the so far made adaptations of item instructions and scale labels.

Secondly, data analyses of part II pointed out that RA1 items appeared as the clearest facet. Items of the other RA-levels correlated with RA1 items so the item pool did not reflect the item structure empirically as differentiated as theoretically intended.

Finally, the analyses of part III indicated an inverted relationship between LOC and O items. This result seemed plausible. It was therefore considered to summarize the LOC and O items into one summarizing STSQ score. The NFS items could be differentiated from O and LOC items. Part III can be summarized with two scores, O and the inverted LOC items (O+LOC score), and items which were interpretable as NFS items (NFS score).

## 6.7. Repeated measurement design

The intention was to apply the STSQ in the context of intervention research in the future. An evaluation of the intervention designs the STSQ should be able to differentiate between different groups. Consequently it was important to investigate how far measuring results with the STSQ were repeatable with an identical sample. Either the test results or the ranking<sup>46</sup> of persons should be similar between two different measures (cf. Bortz & Lienert, 2003). The correlations of the STSQ scores between two measurements were used as an estimate of the stability over time (Crocker & Algina, 1986).

Besides this major aim, changes and revisions of items of the previous application of the STSQ were followed up. The resulting knowledge was expected leading to new adaptations in order to improve the STSQ further.

### 6.7.1. Pre-considerations of the research design and methods

With the purpose of estimating the STSQ's stability over time, two measurements were conducted within one sample of respondents.

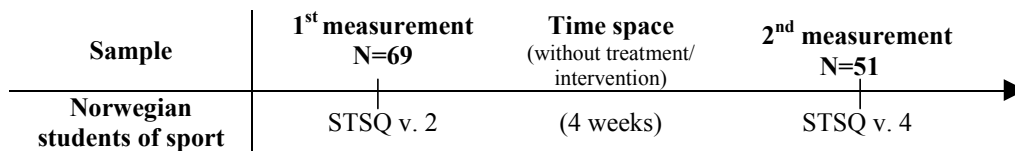


Figure 29: Repeated measurement design

As shown in Figure 29 there was a time space of four weeks between the measurements without intervention or treatment of the sample. Facets of the STS were related to relative stable personal characteristic such as knowledge and attitudes. It was important to choose an appropriate time space between first and second measurement when intending to estimate a test-re-test reliability. The longer time space between measurements, the more relevant error influences may become. The shorter the time space between the two measurements, the more relevant undesirable remembering effects may become (Lienert & Raatz, 1998). However, aiming at an evaluation of the stability over time, four weeks between each measurement was

<sup>46</sup> Similar rank means that the person with the highest scores at measure one, get the highest scores at measure two.

considered an appropriate time space in order to keep a reasonable relation between the two measurements.

Data was collected with a sample of Norwegian sport students at the Norwegian School of Sport Sciences. The number of participants in first measurement were N=69. The repeated measurement was carried out four week later with same group. In this second measurement N= 51 persons participated. N= 35 persons could be identified as identical persons participating in both first and second measurement. In order to identify and consequently compare identical persons of the first and second measurement, each questionnaire needed to be coded<sup>47</sup>.

**Table 27: Sample characteristics**

Sample		1. measurement	2. measurement
N total		69	51
N identical/match			35
Age	Mean	21	21
	Minimum	18	18
	Maximum	28	27
Sex	Male	48	31
	Female	21	19
	Missing values	0	1
Abroad experience <sup>48</sup>	Yes	19	12
	No	49	38
	Missing values	1	1

Between the two measurements of stability testing, some of the STSQ items were changed, and the second measurement of this data collection was carried out with the slightly adapted version of the STSQ. It was not expected that these smaller adaptations of the STSQ could affect the stability results significantly because the respective items were adapted to make the intended meaning clearer. Since an estimation of test re-test reliability was based on correlations between first and second measurement, the influence of smaller revision was expected to be controllable. In addition, smaller changes of the questionnaire items of the second measurement may have reduced the remembering effects. But they were supposed to measure the same aspects as before.

<sup>47</sup> In addition, the codes shall avoid that the participating persons need to reveal their names on the questionnaire, and can still keep their anonymity when participating in the study.

<sup>48</sup> Students who have been abroad at least more than 2 month

*The following items were used for stability estimations:*

At the first measurement of the stability study, data was collected with item I.1 – I.11 of part I of the STSQ. The second measurement was conducted with a reduced item pool consisting of item I.1, I.3, I.9. These three items of the second measurement appeared to be more representative for measuring a STS than the rest of the item pool as will be demonstrated in the result chapter (cf. chapter 6.7.4).

Table 28 shows the item pool of the second part of the STSQ which was used for data collections within both measurements.

**Table 28: Item pool of part II applied in both measurements**

<b>Item</b>	<b>Situation 1</b>
1.1	The person looks different
1.2	Generally, strangers are controlled
1.3	The bus driver does not like foreigners
1.4	The bus driver is irritated
1.5	It was a routine control
1.6	The person's behaviour seems to be suspicious
<b>Situation 2</b>	
2.1	The new player did not play as well as expected
2.2	The new player did not adapt to the team
2.3	The new player disturbed the team spirit
2.4	The team opposed to the new player
2.5	The new player was not integrated into the team
2.6	The team was in poor condition
<b>Situation 3</b>	
3.1	Co-educational teaching should have priority over individual beliefs
3.2	Pupils should obey existing rules
3.3	P.E. class could be organized gender-separated
3.4	Students have the possibility to select exercises
3.5	Different religious faiths should be respected
3.6	P.E. can be arranged in a mutually acceptable way
<b>Situation 4</b>	
4.1	Men are supposed to be more efficient
4.2	Women are supposed to be more integrative leaders
4.3	Personal leadership determines the decision
4.4	Professional qualification determines the decision
4.5	In case of equal qualification, the gender-balance determines the decision
4.6	The heterogeneity of the company structure will influence the decision



The mean scores for both measurement (1 and 2) were calculated as follows, and in accordance with previous results:

$$\overline{RA1}_{arith. mean} = \frac{1}{n} \sum_{i=1}^n RA1_i = \frac{1.1 + 1.2 + 1.3 + 2.1 + 2.2 + 2.3 + 2.4 + 3.1 + 3.2 + 4.1 + 4.2}{11}$$

$$\begin{aligned} \overline{RA2}_{arith. mean} &= \frac{1}{n} \sum_{i=1}^n RA2_i = \\ &= \frac{1.4 + 1.6 + 2.5 + 2.6 + 3.3 + 3.4 + 3.5 + 3.6 + 4.3 + 4.4 + 4.5 + 4.6}{12} \end{aligned}$$

Since the pilot study showed that it was not possible to distinguish as differentiated as intended, only two RA score were calculated. But it is the aim of following procedures to adapt the items more and more in order to enable for differentiation between these two RA levels.

Table 29 shows the item pool of the third part III of the STSQ also applied within both measurements:

**Table 29: Item pool of part III applied in both measurements**

Item Code	Item
O 1.1:	Open-mindedness helps understanding.
O 1:	It seems that different people have different values.
O 2:	The open result is responsible for the excitement of a competition.
O 3.1:	To meet different kinds of people is enjoyable.
O 4*:	Sitting in a group of strangers provokes uncomfortable feelings.
O 5:	Freedom does not mean the absence of rules.
NFS 1:	It is important to forward cultural values to the next generations.
NFS 2.1:	To plan ahead helps to provide security.
NFS 3:	Being obliged to make decisions provokes uncertainty.
NFS 4:	For activities, I prefer precise instructions to open suggestions.
LOC 1*:	The number of friends I make depends only on me and my behavior.
LOC 2:	Most of the events in my life are determined by other people.
LOC 3:	Fate determines whether I have more or fewer friends.
LOC 4.1*:	People differ substantially.
LOC 5*:	In spite of its apparently continuous change, culture contains some stable values.
LOC 6:	Unforeseen events upset me.

\*inverted item formulation with respect to the regarding category

### 6.7.2. Expectations

According to the aim of testing the stability over time, STSQ scores were expected to correlate with each other significantly between the two measurements. But the empirical experiences with the STSQ were still in the beginning. It was difficult to formulate more concrete expectations about which STSQ score might show the most stable relationship between the two measurements.

Part I of the STSQ were investigated further especially because of the mentioned technical problems (with the copy machine in Germany) during the previous data collection which made interpretations of these previous results difficult. The scales of part I were analyzed for validity as well. It was also examined if the responses of the familiarity and sympathy scale could be interpreted in accordance with the theoretically assumed eSTS response pattern. Four response pattern scores were calculated: *eSTS response pattern*, *xenophobia pattern*, *exoticism pattern*, and *pity pattern*. As mentioned previously, the alternative patterns were mainly supposed to demarcate eSTS from alternative response patterns. The alternative pattern scores are not specifically conceptualized, but these phenomena can be related to STS as shown in previous chapters.

### 6.7.3. Procedure

According to Rost (2004) the way of estimating stability over time lead to underestimations of stability because reliability and stability over time are estimated simultaneously. Structural equation modelling (SEM) would allow a separation between reliability and stability, and would lead to more exact estimates (cf. Steyer & Eid, 2001). Such statistical procedures are more appropriate to apply to test models which are theoretically or empirically well-founded (Bühner, 2004). Model testing with confirmative factor analysis (CFA) or SEM could for instance become a relevant goal in the future for estimating construct validity.

The first step of investigating stability over time was to compare STSQ scores of the first measurement with the second measurement. The different STSQ mean scores were used for calculating correlations between the two measurements.

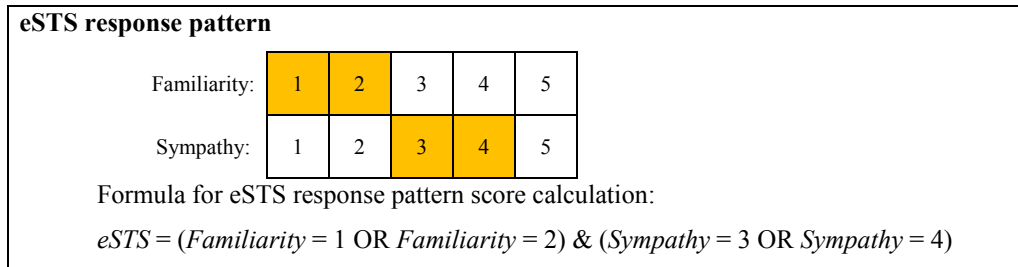
Descriptive statistics and correlations were applied in order to examine the STSQ with the purpose to gain hints on how to improve single items and scoring procedures further. Explorative factor analyses were used in order to explore the structure of data. If the results could support a differentiation between different scales items in the theoretically associated way, such a result was expected to serve as supportive hint for construct validity. If not, a closer

look on how the items load in on different factor in relation to other items was expected to reveal some hints on how the items needed to be adapted in order to point out the intended meaning.

At first, part I of the STSQ:

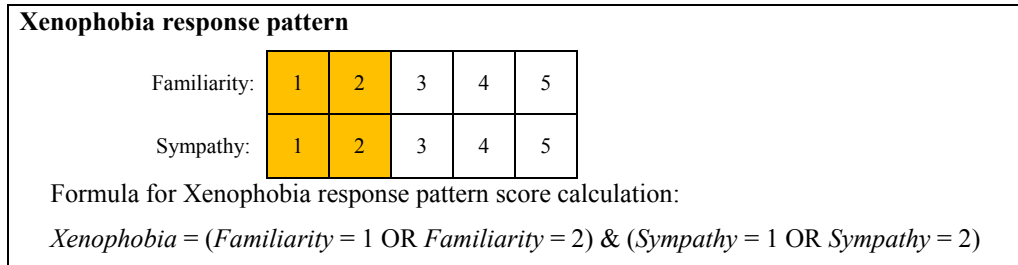
Each scale of part I was analyzed separately at first (familiarity and sympathy). The stability of the assumed response pattern scores was also examined.

The following response patterns were compared with each other empirically:

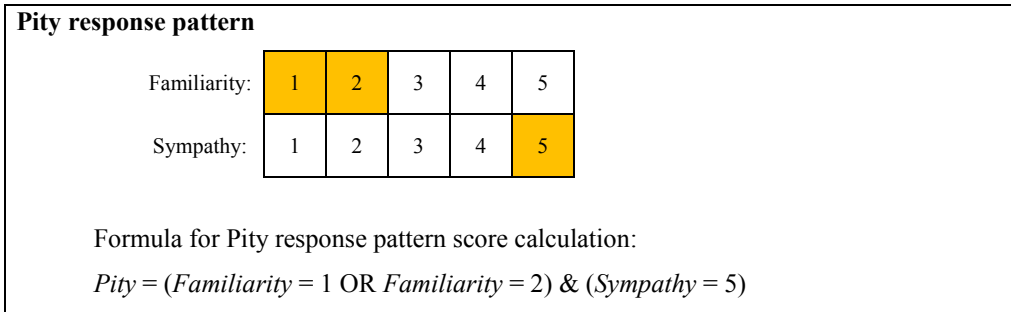


**Figure 30: Calculation of the STS response pattern score**

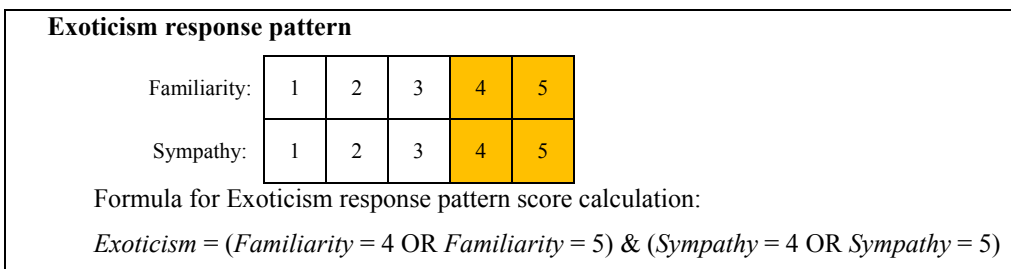
The next patterns illustrate the calculation of reasonable alternative response pattern.



**Figure 31: Calculation of the xenophobia response pattern score**



**Figure 32: Calculation of the pity response pattern score**



**Figure 33: Calculation of the exoticism response pattern score**

The single scale scores on both the familiarity and sympathy scale was transformed into response pattern scores as illustrated above. Each observed or measured pattern was assigned with a score 1. The mean score overall picture items represented the respective responses.

The correlation was estimated by non-parametric correlations for both pattern-score comparisons and scale comparisons as argued previously. The calculation of eSTS pattern scores also implied a reduction of scale quality. The raw scores on both scales (familiarity, sympathy) were combined and the resulting pattern score was transformed into dichotomous score as shown in figure 30.

Part II of the STSQ was basically examined with respect to stability of the now two RA scores. Correlations were calculated with non-parametric procedures as argued previously.

A sub-goal for this item analyses with a new sample was a calculation of the difficulty index for each item (Guttman, 1950). Dichotomous items were required for such an analysis. The scale quality of the items of part II was therefore changed into dichotomous level (probable – improbable) in order to investigate the item difficulties (cf. *ibid.*, 1950). The results are shown in the appendix (9.5). The items of each situation were then arranged in a new order according to increasing difficulty index (cf. Rost, 2004).

*Thirdly, part III of the STSQ:*

The third part of the STSQ was analyzed according to its stability over time. The items were summarized as theoretically intended, but the empirical results of previous studies led to some revisions of item membership as previously illustrated (cf. chapter 6.6.6).

Descriptive statistics and correlations were applied to follow up the previous item revisions.

#### 6.7.4. Results

*Test-retest analysis (stability over time)*

The following Table 30 shows an overview of mean score results of both measurements within this stability study.

**Table 30: Stability results of the first measurement and second measurement (Spearman's rho)**

Scale/ score	N	First Measurement (1)		Second measurement (2)		rho
		Mean (1)	SD (1)	Mean (2)	SD (2)	
Familiarity	34 (35)	2,65	,67	2,50	,8	,66**
Sympathy	33 (35)	2,91	,51	3,11	,72	,66**
eSTS pattern score	33/35	0,32	,3	0,28	,27	,35*
RA1	34 (34)	2,4	,34	1,49	,18	,62**
RA2	34/34	2,78	,29	1,62	,14	,25
O	32/33	3,21	,33	3,27	,30	,26
NFS	32/31	2,85	,32	2,89	,38	,45*
LOC	33/32	2,45	,37	2,35	,21	,3
O + LOC	30/31	2,97	,24	2,92	,20	,48**
Valid N (listwise)	23					

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

According to Table 30 the scales, familiarity and sympathy, show moderate and significant correlations between the first and second measurement.

The STS response pattern score show low but significant correlations at a 0.05 level between first and second measurement.

A stability analysis of part II shows a moderate and significant correlation between first and second measurement of the RA 1 score, but no significant correlation between the two measurements of the RA 2-score could be found (cf. Table 30).

The items initially assigned to O and LOC separately were combined to one summarizing score. This O+LOC<sup>49</sup> score seemed theoretically reasonable and gained empirical support within the previous sub-study. Table 30 indicates a moderate and significant relationship of the O+LOC scores between the two measurements ( $\rho=0,48$  with a significance level of 0,01). In a similar way, the NFS score correlated moderately and significantly with both measurements ( $\rho=0,45$  with a significance level of 0,05) and indicated some stability overtime.

In addition to estimations of stability over time, selected items and item scores were examined further to obtain validity hints, and control the effects of previous item adaptations.

*Part I of the STSQ:*

Comparisons between the STS response pattern and alternative patterns led to some significant results. The “xenophobia pattern” correlates negatively with the eSTS pattern at a significance level at 0.05 of a Spearman correlation in both measurements ( $\rho= -0,42$ ).

The “eSTS pattern” showed a negative significant correlation at the significance level of 0.01\*\* with the “exoticism pattern” in measurement 1 ( $\rho= -0,34$ ).

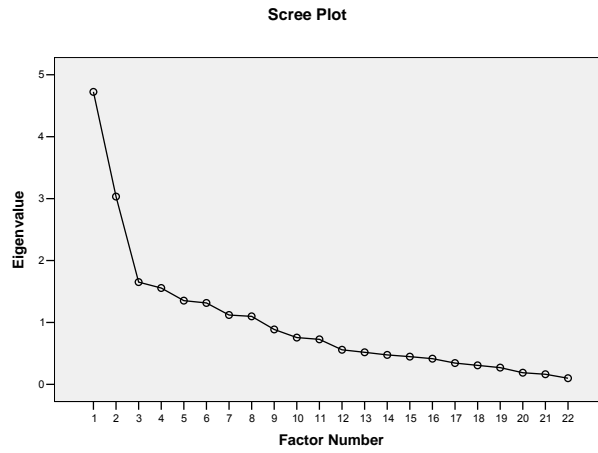
Both the familiarity and sympathy scales over 11 items showed moderate to high Cronbach’s alpha values as shown in the next table.

**Table 31: Scale analysis part I, STSQ v.2**

Scale	N <sub>Total cases</sub>	N <sub>valid cases</sub>	N of items	Cronbach’s alpha
Familiarity	69	66	11	,789
Sympathy	69	65	11	,752

The following Scree plot indicates a clear two factor solution which can be interpreted as familiarity (factor 1) and sympathy (factor 2).

<sup>49</sup> For this purpose the score of the LOC items needed to be inverted/ recoded so that they suit in a mentioned reasonable way to the O-items.



**Figure 34:** Scree plot part I of the STSQ (N=69)

The factor loadings show that most of the familiarity scale items load highest on factor 1. The sympathy scale items are more difficult to assign because they show lower factor loadings. The sympathy scale of item 2, 3 and 10 are very low. The familiarity scale item 6 show relative high factor loadings on both factors. The sympathy scale of the same item (6) shows a factor loading of 0,42 on factor 1 and 0,18 on factor 2.

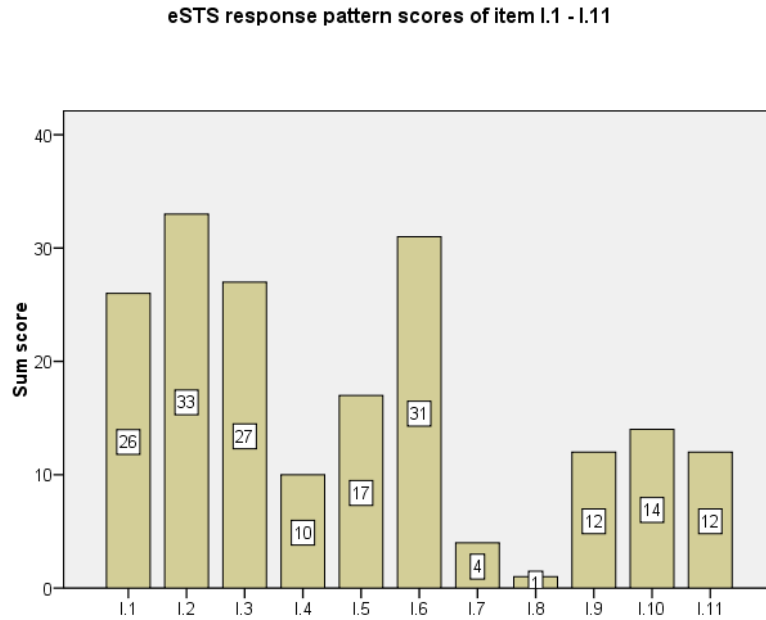
**Table 32: Results of factor analysis, part I of the STSQ**

Scale	Item	Factor		Communality
		1	2	
<b>Familiarity</b>	I.1	,43	,19	,19
	I.2	,62	,39	,39
	I.3	,65	,43	,43
	I.4	,38	,15	,15
	I.5	,66	,45	,45
	I.6	,85	,73	,73
	I.7	,39	,22	,22
	I.8	,37	,19	,19
	I.9	,49	,24	,24
	I.10	,26	,11	,11
	I.11	,53	,28	,28
<b>Sympathy</b>	I.1	-,09	,40	,40
	I.2	,13	,14	,14
	I.3	-,15	,07	,07
	I.4	,00	,35	,35
	I.5	,25	,44	,44
	I.6	,42	,18	,18
	I.7	,08	,35	,35
	I.8	-,05	,35	,35
	I.9	,14	,43	,43
	I.10	,24	,13	,13
	I.11	,28	,27	,27
Eigenvalues (initial)		3,62 (4,72)	2,87 (3,03)	
% of variance (initial)		16,47 (21,47)	13,052 (13,785)	

Extraction Method: Principal Axis Factoring.  
Varimax rotation



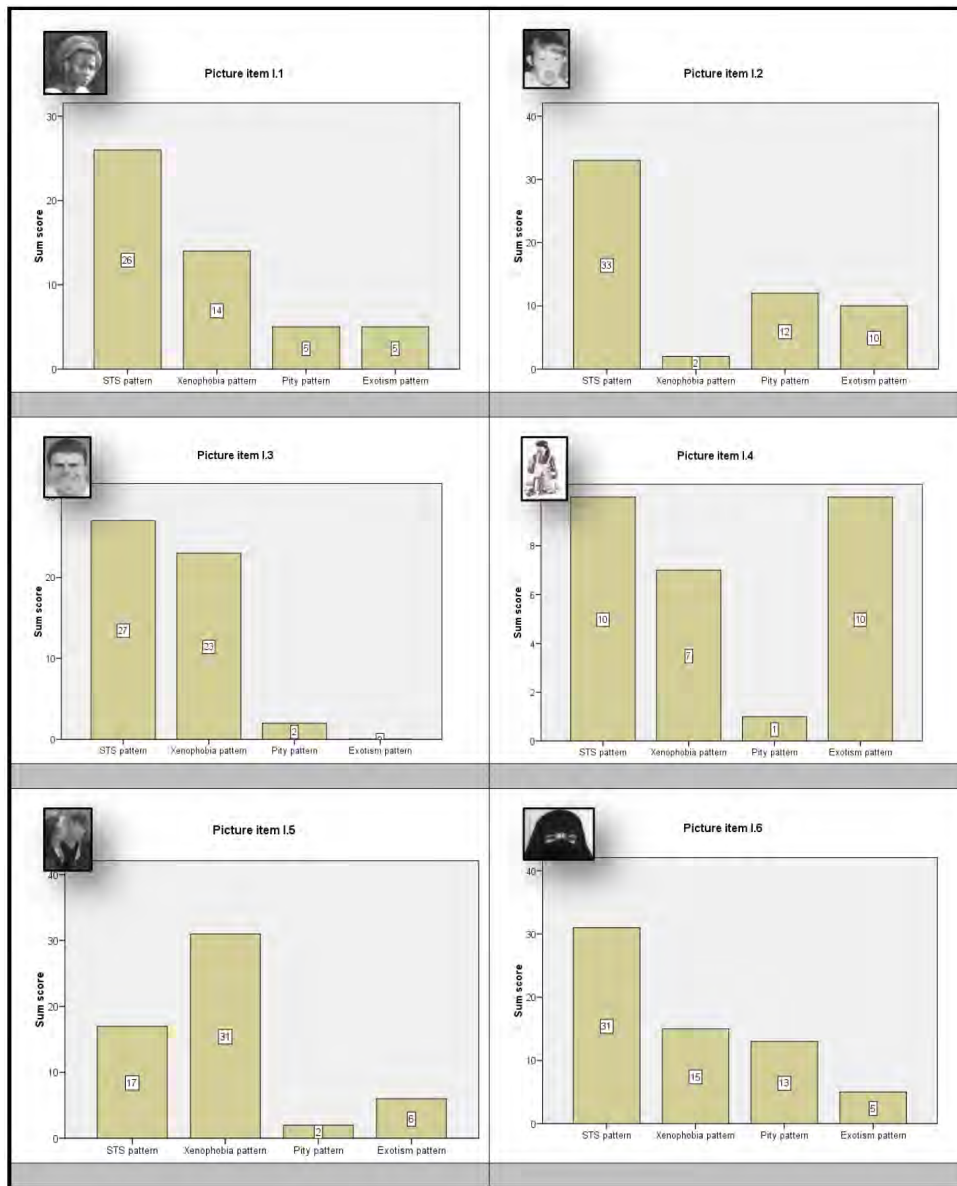
The following figure shows the frequencies of eSTS pattern scores of each item of part I.



**Figure 35: Frequencies of STS response pattern score of item I.1 – I.11**

The items I.1, I.2, I.3, and I.6 measured the highest number of STS response pattern scores (by a sample size of N=69).

The next figure shows the results of frequency analysis including the alternative response patterns scores with respect to each item. The intention of the frequency analyses was to find out which item seemed to be the most sensitive with respect to measure STS. Two criteria helped to decide on measuring sensibility: a) total frequency score, and b) dominance of STS frequencies compared to alternative response pattern scores. The items that counted the highest frequencies on the eSTS pattern score were more suitable to measure STS than lower frequencies. The eSTS pattern frequencies should be determinant compared to the alternative response pattern scores.



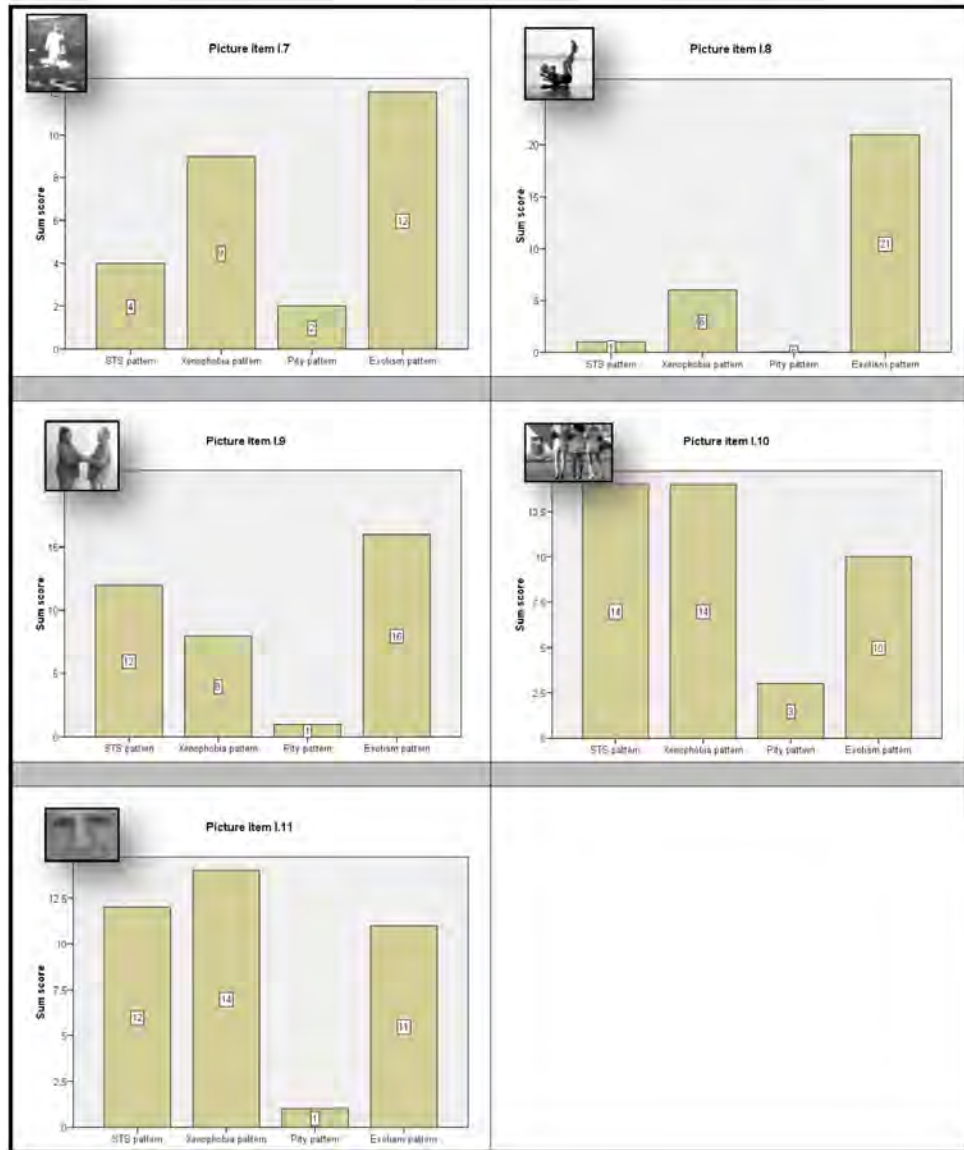


Figure 36: Frequencies of response pattern analysis in comparison to alternative response pattern

The comparison of STS and alternative pattern scores showed that item I.2 and item I.6 count a relative high number of pity pattern (12/13) in relation to item I.1, and I.3. Furthermore, item I.5, and I.3 showed a high number of xenophobia pattern (31/23). Item I.8 of the exoticism pattern dominates this item.

*Part III of the STSQ:*

According to the results of the previous study, an explorative factor analysis was calculated with the following parameters: 2 factors were extracted with principal axis factoring method. A Varimax rotation was conducted.

**Table 33: Results of factor analysis part III (sample of the first measurement of the stability study, N=69)**

Item	Factor		Communalities
	1	2	
O1.1 "Open-mindedness helps understanding"	<b>.66</b>	.04	.44
O2 "The open result is responsible for the excitement of a competition"	<b>.31</b>	-.03	.1
O3.1 "To meet different kinds of people is enjoyable"	<b>.45</b>	-.39	.36
O4 "Sitting in a group of strangers provokes uncomfortable feelings"	-.04	<b>.4</b>	.16
O5.1 "Freedom does not mean the absence of rules"	<b>.7</b>	-.14	.51
NFS1 "It is important to forward cultural values to the next generations"	<b>.73</b>	.16	.56
NFS2.1 "To plan ahead helps to provide security"	<b>.37</b>	.07	.14
NFS3 "Being obliged to make decisions provokes uncertainty"	<b>.02</b>	<b>.58</b>	.34
NFS4 "For activities, I prefer precise instructions to open suggestions"	-.20	.13	.06
LOC2 "Most of the events in my life are determined by other people"	<b>.06</b>	<b>.67</b>	.45
LOC3 "Fate determines whether I have more or fewer friends"	-.05	<b>.54</b>	.29
LOCr4 "People differ substantially"	<b>.21</b>	-.12	.06
LOC5 "In spite of its apparently continuous change, culture contains some stable values"	<b>.41</b>	-.11	.18
LOC6 "Unforeseen events upset me"	-.12	<b>.41</b>	.18
Eigenvalues (initial)	2.17 (2.91)	1.65 (2.2)	
% of variance (initial)	15.47 (20.76)	11.77 (15.64)	

Extraction Method: Principal Axis Factoring.  
Varimax rotation

The results of the factor analysis are shown in the table above. The extraction of two factors explained 27,2% variance which is not much. Because of the heterogeneity of the item pool it was not expected to achieve high percentage of explained variance. EFA was approximate approach for item analyses. The factor loadings highlighted in Table 33 can be seen in a reasonable relation to the other items dominating the same factor. Factor loadings which appear plausible with theoretical expectations can be interpreted as supportive hints for construct validity. The highlighted factor loadings appear plausible their relationship other items on the same and/or other factor. These statistical relationships can be interpreted as supportive

hints for construct validity (Crocker & Algina, 1986). Except item O-4 the O-items can be assigned to factor 1. Item O-4 can be assigned towards factor 2 together with some LOC and NFS items which appears plausible because LOC, NFS and item O-4 indicate uncertainty.

#### 6.7.5. Discussion and conclusions

##### *Part I of the STSQ:*

The results of the test-retest analyses showed that the two scales (familiarity and sympathy) correlated moderately and significantly ( $\rho=0,66$ ) with the first and second measurement. This result indicated a relatively stable measurement of the two scales over a time space of four weeks. With respect to the type of measurement (projective measuring device), this result is acceptable on the single scale level. The results of the pattern analysis indicated some measuring stability over time for the STS response pattern by a significant correlation between first and second measurement.

The negative correlations between the eSTS pattern score and the xenophobia-pattern and the exoticism-pattern can be interpreted as supportive validity hints regarding interpretations of the eSTS pattern. The result supported theoretical assumptions about the relationship between both concepts; xenophobia was assumed as a contrary facet in relation to STS. The exoticism-pattern was supposed to be negatively related to eSTS pattern.

Further theoretical reflections about the supposed meaning of the selected pictures lead to the conclusion of excluding some pictures/items from the item pool. Item I.2 “the boy with trisomi 21” (cf. item I.2 shown in Table 3) was excluded in the next data collections. A disease such as trisomi 21 might be problematical in order to measure a STS because this picture may bias the sympathy scale. The boy’s condition is genetically determined. It is unreasonable to make this boy responsible for his potentially perceived “strange” behavior. Descriptive data supported this theoretical consideration. The mean score was quite high on the sympathy scale of this item (mean = 4,1). This indicated that respondents usually scored above average on the five point sympathy scale. As a result, this item was not helpful to differentiate between different groups regarding a STS. The open-question-part of the item supported this impression as well. This item was often headed with terms such as “poor boy” which was more directed to measure “pity” instead of STS.

One might assume that the skin-tone item (item I.1 shown in Table 3) needed to be excluded from the item pool as skin-color is also genetically determined. On the contrary, distinguishing skin tone is often associated with a person’s character or specific behavior such as

criminal activity (Hall, 2005). This item may allow differentiation between relevant extreme groups and still be an appropriate indicator for a STS. Pattern analysis also showed that this item appeared as a sensitive indicator for measurement of STS (cf. Figure 35 and 36). Accordingly, people with overweight are often made self-responsible for their "weight problem"; so the image shown in item I.3 (Table 3) was also considered to be a useful stimulus in order to differentiate according to the construct.

Item I. 5 (showing a homosexual couple kissing) was often commented with terms such as "disgusting". Such a comment is consistent with the frequently observed xenophobic (or in this case homophobic) response pattern score (cf. Figure 35 and 36). Item I.8 (fitness woman) is not suitable because of the high mean score on the familiarity scale and was excluded from further data collections. The sample of persons for this data collection was a group of students from the Norwegian School of Sport Science. It therefore appeared reasonable that these sport students would perceive the picture as familiar. But the descriptive statistics in Table 69 (cf. appendix 9.1) showed in addition that the items I.1, I.2, I.3, I.5, and I.6 show a low familiarity mean scores. They therefore seemed to meet one prerequisite for measuring an STS.

As discussed before, item I.2 is a more suitable indicator when intending to measure a construct such as pity. The results of the pattern analysis, however, showed that this item counted the highest number of eSTS pattern scores even though the previously argumentation demonstrated that this image provoked feelings such as pity. Assuming that the argumentation above is reasonable, it appears to be difficult to differentiate between STS and pity. The reason for this may be that a STS response is defined with score 3 or 4 on the sympathy scale. Considering that the scale range is from 1 to 5, a score 4 can be interpreted as a more positive sympathy ascription, and score 5 as maximum sympathy assignment. Measuring errors such as response tendencies might be responsible for the difficulty of differentiating precisely between STS and pity.



The results of explorative factor analysis of part I showed that the two factors could be interpreted as familiarity and sympathy scale. But the results in Table 32 also revealed that some items were problematic. The items I.5 (homosexuality), I.6 (religious faith), I.10 (prostitution), were originally supposed to represent relevant topics of strangeness. The discussed results and further theoretical reflections led to the conclusion, that the latter mentioned items are too value loaded, and could be perceived as provocative by the respondents. This might then lead to reactions such as denying filling out the item seriously. The headlines on the open question analysis supported this impression as it did with the homosexuality item example.

With respect to the rest of the picture item pool a general problem arose when analyzing it. In order to measure an eSTS the respondents dominantly needed to perceive it as different. If not, the sympathy score could not theoretically be related to the construct and would therefore not be a representative indicator<sup>50</sup>. This problem was also related to the selected projective measuring procedure in general. In order to increase the reliability of this part of the instrument, it might help to add items according to the defined criteria. The items I.4, I.7, I.8, I.11 did not show the required distribution on the familiarity scale (indicated by higher mean values and low or negative skewness values in Table 69 (appendix 9.1)).

The intention was to keep item I.9 for the time being because the image shown in this item might be perceived ambiguously. It might therefore be valuable in differentiating between groups even though the results did not support this idea.

In addition to the mentioned changes, the following overview shows item examples of part I which illustrates the smaller linguistic revisions between the two measurements of the stability study. The main idea was to make the item instructions easier to understand:

**Table 34: Applied item pool between different data collections. Revisions are illustrated with an item example, STSQ, part I (EM)**

Stability study, first measurement	Study 2, second measurement
<p>Applied items: I. 1-11</p> <p>II. Please have a closer look at <u>each</u> of the following pictures!</p> <p>1 a) Please give a headline to the situation shown in the following picture!</p> <div data-bbox="373 1234 496 1357" style="text-align: center;">  </div> <p>1 b) Please indicate how familiar (1.), <u>and</u> how sympathetic (2.) you feel imagining the given situation in the picture above. Please cross out the respective number:</p> <p>1. Familiarity:      low <input type="text" value="1"/> <input type="text" value="2"/> <input type="text" value="3"/> <input type="text" value="4"/> <input type="text" value="5"/> high</p> <p>2. Sympathy:      low <input type="text" value="1"/> <input type="text" value="2"/> <input type="text" value="3"/> <input type="text" value="4"/> <input type="text" value="5"/> high</p>	<p>Applied items: I.1, I.3, I.9</p> <p>II. Please have a closer look at <u>each</u> of the following pictures!</p> <p>1 a) Please give a headline to the following picture!</p> <div data-bbox="884 1252 1007 1375" style="text-align: center;">  </div> <p>1 b) Please indicate how familiar (1.), <u>and</u> how sympathetic (2.) you feel, when you imagine the situation illustrated above. Please cross out the respective number:</p> <p>1. Familiarity:      low <input type="text" value="1"/> <input type="text" value="2"/> <input type="text" value="3"/> <input type="text" value="4"/> <input type="text" value="5"/> high</p> <p>2. Sympathy:      low <input type="text" value="1"/> <input type="text" value="2"/> <input type="text" value="3"/> <input type="text" value="4"/> <input type="text" value="5"/> high</p>

Because of the mentioned difficulties of the picture suitability, it was important to find other items, which were difficult to classify and could function as stimulus for a perception of difference. When searching for new or additional items it is important to more ambiguous

<sup>50</sup> - because STS was defined previously as dealing with differences and strangeness constructively.

items in order to achieve a more balanced distribution on the sympathy scale. In the meantime, the chosen items of the existing item pool (I.1, I.3, and I.9) were used.

*Part II of the STSQ:*

The stability estimations showed some stability over time of the RA1 score. The RA2 scores showed no stability over time. This result could be interpreted as consistent with results of the previous data collection (cf. pilot study). But part II of the STSQ was analyzed more specifically in one of the following data collections.

The items of the second part of the STSQ were revised as shown in Table 35:

**Table 35: The table shows revisions of items of part II of the STSQ according to the result of the parallel conducted intervention study. The revisions are only applied to the second measurement of this stability study.**

Item	STSQ part II	Revised items	Comment/reason for revision
<i>All items part II</i>		Change from 4 level interval scale to dichotomous scale; new item order according to their difficulty index (cf. appendix 9.5)	Determination of item difficulty in order to re-arrange the item order according to increasing difficulty index (cf. appendix 9.5)
5.2	<i>It must be exciting for him/her to discover the new environment</i>	5.2.1: It is difficult to discover the expectations of the new surroundings.	Reformulation in order to point out the uncertainty aspect
5.4	<i>The unknown expectation of the new surrounding makes him/her feel uncomfortable.</i>	5.4.1: The new surrounding makes him/her feel uncomfortable.	Reformulation in order to make this item easier to understand

*Part III of the STSQ:*

As shown in the results, the combined O+LOC score showed acceptable stability results (Table 30). The NFS facet showed some stability.

The EFA result showed the effects of previous item revisions. In accordance with the results of the pilot study, two factors were extracted. Except items NFS1 and NFS2, the first factor is dominated by O-items according to the factor loadings in Table 33. Factor two appeared to be dominated by LOC items and one NFS item. The revised item O5 now showed high factor loading on the same factor as the other O-items. The revised item O3 showed relatively high factor loadings on two factors, but both loadings were interpretable according to the construct because this item loaded positive on factor one and negative on factor two. This result was consistent with the construct's assumptions because the factor two showed dominant factor loading of LOC and NFS items which were contrary to the O items. Item O4



---

showed a high factor loading on factor two that will be interpreted on the basis of the construct. If a person feels uncomfortable within a group of strangers, this might be one indicator of a NFS according to the construct.

Table 33 shows some unreasonable results. Item NFS1 loaded unexpectedly on factor one which earlier was interpreted as an openness facet (0,73). In addition, according to previous results revised item NFS2.1 showed an unreasonable factor loading on factor one as well (0,37).

Because of low inter-item correlations both in previous studies and this study, item LOC1 was excluded from the item pool. Item LOC2 and LOC3 show the highest factor loadings in factor two together with item NFS 3 and O4. This relationship was interpreted reasonably in line with the construct STS. Previous results have already indicated an inverted relationship between the O- and LOC-indicators (cf. Table 22). In this sense, item LOC6 could reasonably be assigned to factor two, but this item showed a low but negative factor loading on factor one.

As the mentioned results showed, a few items needed to be revised because they were difficult to assign towards one specific category. It was important not to change too much at once. Too many changes could mix up the whole factor structure and it could be difficult to control the effects. To follow the mentioned methodological background idea, it was important to adjust the STSQ step by step and carefully control each change and its effect.

The following table shows subsequent item revisions based on the empirical result so far. The table showed the item abbreviation, the item before revisions are made, the revised item formulation and a short reason for the changes being made. The third column (revised item) also shows how the item is labeled after item revisions. The last number of the item label indicates the number of changes/revisions of the respective item to distinguish changed items from the initial item pool.

**Table 36: Revision of items according to the resulting knowledge of this stability study, STSQ v.2, part III**

Item	STSQ part III, study 2, first measurement	Revised item	Comment/reason for revision
O1	<i>It seems that different people have different values</i>	O1 is excluded from the item pool	Unclear item; showed unreasonable component loadings; loads positive on two contradictory assumed category
NFS 4	<i>For activities, I prefer precise instructions to open suggestions.</i>	NFS4.1: When engaging in an activity, I prefer receiving clear instructions to open suggestions.	Linguistic revision in order to increase the factor loading on the respective factor
LOC1	<i>The number of friends I make depends only on me and my behavior.</i>	LOC1: Excluded from the item pool	No significant loadings on the suggested components
LOC 5	<i>In spite of its apparently continuous change, culture contains some stable values</i>	LOC 5.1: Life brings continuously changes LOC 5.2: Culture has some stable values.	Original item was too complex. Therefore, it seems reasonable split-up this item into two items LOC5.1 and LOC5.2.
LOC 6	<i>Unforeseen events upset me</i>	LOC6.1: I like unforeseen events.	Linguistic change and inverted formulation is supposed to increase factor loading because the item might be easier to answer.

Considering the previous results together with the results of this study, it seemed that all initially three facets of part III are related with each other. The O-items usually correlated negatively with both LOC and NFS items. LOC and NFS items correlated positive with each other. This tendency might support the idea of summarizing the consistent items of part III into one summarizing score. Based on the results of explorative factor analysis within this sub-study, item O1, O2, O3, O5 and LOC 5 determined the openness dimension. In addition, O4\*, NFS3, LOC2, LOC3, and LOC6 determined the other facet which was interpreted as a kind of uncertainty facet because items of both categories (NFS and LOC) are related to “uncertainty”. If this tendency finds further empirical supported it might be reasonable to summarize the items in the mentioned way.

In summary, the correlations between 0,4 and 0,7 between the two measurements (within four weeks interval) indicated moderate stability results for nearly all STSQ scales and scores. The familiarity and sympathy scale, and the RA1 score appeared to be the most stable facets of the selected sample ( $\rho=0,62$ ,  $p<0,01$ ). The RA 2 score showed a very low correlation between the two measurements ( $\rho=0,25$ ). The eSTS pattern scores showed a low (around  $\rho=0,35$ ,  $p<0,05$ ) but significant correlation between the two measurements.

The O score showed a low correlation, but the results of the combined O+LOC score indicated a significant and moderate stability ( $\rho=0,48$ ,  $p<0,01$ ). The results of the NFS-score indicated a moderate and significant relationship between the two measurements ( $\rho=0,45$ ,  $p<0,05$ ).

Negative correlations between the alternative xenophobia ( $\rho=0,42$ ,  $p<0,05$  in both measurements) and exoticism ( $0,32$ ,  $p<0,01$ ) pattern types appeared to be reasonable in relation to the eSTS response pattern and supported therefore the supposed meaning as indicator for measuring emotional STS in the sense of the construct.

As argued previously, the picture items I.1, I.3, I.9 were most reasonable as stimulus for measuring a eSTS pattern.

## 6.8. Application of the STSQ in a Norwegian high school

The STSQ is designed for populations such as university students, teachers, or social workers. An application of the STSQ could be in the realm of on-the-job training programs for teachers. An interesting question would be if the teachers are able to transfer gained intercultural competence to pupils or students. The STSQ would then be applied to measure possible effects of the training program on the pupil's level.

It was investigated if the STSQ can be applied on high school level. So far, it seemed not reasonable to apply particular part II and III of the STSQ under the age of 16. It can not necessarily be expected that these students show an awareness and reflection ability about the contents of these parts of the STSQ.

Following my earlier argumentation, it seemed reasonable to investigate the age-difference of applying the STSQ in school classes. Comparisons between groups of students of three different high school levels may indicate intellectual differences by STSQ score differences. It could also provide further validity hints of STSQ items and item scores.

### 6.8.1. Considerations of the research design

A comparison of STSQ mean scores of high school levels was expected to give information about the differences according to the pupils' intellectual development by comparing these results to theoretically anticipated ones.

Besides comparing different high school levels, the sample was divided according to criteria (variables). The sample was divided into separated groups such as persons with and without immigration background, students with abroad experiences of more than 2 months versus pupils with no abroad experiences and first year students versus third year students. A comparison between the results of these sub- or extreme groups was supposed to gain further validity hints.

Anticipated relationships between different facets of the conceptualization were compared with correlations between measured STSQ scores and the results gain further validity hints of the measuring instrument. Similar to previous procedure, correlation studies between single items were expected to improve the validity of the STSQ measurements.

This study was carried out with a sample of students at a Norwegian high school (N=224). The school is a specialized sport gymnasium (Norges Toppidrettsgymnas). The pupils were used to speaking English. They were being trained by coaches from different countries who

often speak English. The students represented young Norwegian competitive athletes on a top-level within their respective age group and discipline.

Data was collected in all three levels of the high school. The age of the pupils varied between 15 and 18 years. First year students were usually between the age of 15 – 16 years; second year class students are from 17 till 18 years, and the last school year students are mainly 18 years old. As shown in Table 37, boys were dominating this high school sample (about 74% boys, and ca. 26% girls in total).

**Table 37: Demographic frequencies of the high school sample**

High school year		1. year	2. year	3. year
<b>Age</b>	15 years	11	-	-
	16 years	67	2	-
	17 years	2	69	7
	18 years	-	5	60
	19 years	-	-	1
<b>Sex</b>	Male	59	54	52
	Female	21	22	16
<b>Multicultural background<sup>51</sup></b>	Yes	10	20	15
	No	65	50	51
<b>Abroad experience<sup>52</sup></b>	Yes	10	6	12
	No	65	65	54

The sample represented almost all pupils of the high school level at this school.

The applied questionnaire version was based on the results of the previous sub-studies. Part I of the STSQ was conducted by using the familiarity scale and sympathy scale and the picture items I.1, I.3, I.9 (N= 173). The data collection was spread over a period of ca. 4 weeks. The item pool of part I was adapted so that N=51 students from 2<sup>nd</sup> and 3<sup>rd</sup> grade at the same high school completed the extended item pool including items I.1, I.3, I.9, I.12\*, I.13\*, I.14\*, I.15\*, I.16\* (cf. Table 3).

<sup>51</sup> At least one parent originates from a different country than Norway

<sup>52</sup> Students who have been abroad at least more than 2 month

The items of part II of the STSQ were dichotomously coded (probable/improbable), and the items of situation five of part II of the STSQ were also used for analyses.

#### 6.8.2. Expectations

It was expected that STSQ score differences between different (extreme) groups helped to gain further validity hints according to the respective aspects/criteria. In particular the following expectations were investigated:

##### Group comparison analyses:

**Table 38: Expectations of (extreme) group analysis**

Criterion	Indicator	Expectations
High school level	Mean score group differences	Students of last high school year were expected to achieve higher STS scores of part III and probably part II of the STSQ than first year students because of an assumed higher intellectual level of last year students.
Abroad experiences ( $\geq 2$ months)	Mean score group differences	Students having long term experiences ( $\geq 2$ months) from abroad were expected to be more STS than students with no experiences from abroad.
Immigration background	Mean score group differences	Based on similar assumptions as before, students with immigration background were expected to show more differentiated experiences regarding strangeness. These students may achieve higher STS scores than pupils without such a background.

##### Search for further validity hints:

The openness score (O) is expected to correlate negatively with both the “need for security” score (NFS) and “loss of control score” (LOC). Re-considering the intended meaning of the LOC and NFS facet, these two categories were of similar character so the LOC and NFS scores should correlate positively with each other.

If data analyses would show reasonable correlations between STSQ scores of the different parts this result could be interpreted as supportive evidence of the STSQ’s validity. Reasonable correlations between the different parts would then be similar to an external validation because differing measuring procedures are applied. Even though each part should measure different facets of STS, it is reasonable that these facets are related to each other as indicated in the operational model. It was explored if and how far the different parts of the STSQ correlated with each other.

Nevertheless, it is theoretically uncertain how far the emotional and cognitive STS scores correspond. The emotional indicators might be influenced by situational factors. The cognitive items such as attitudes, insights and understandings might be stable indicators. It could

---

therefore be difficult to find empirical support for the relationship between emotional and cognitive STS scores.

#### 6.8.3. Procedure

Non-parametric statistics were preferred for reasons mentioned previously and the validation requires comparisons of variables with differing scale levels. Spearman-rho was used as correlation indicator in order to test relationships between different variables and STSQ scores as mentioned in the chapter above.

The Mann-Whitney U-test was applied in order to test mean score differences between two independent samples such as extreme group comparisons.

The items of situation five (cf. Table 8 and Table 35 showing previous revisions) were used as an additional indicator for group differences. Items 5.2, 5.3, 5.4.1, 5.5 and 5.6 were summarized as a mean score and labeled as uncertainty score. This score represented a hypothetical perception of uncertainty when imagining situation five. The results were compared to LOC and NFS scores and functioned as an indicator for validation purposes.

#### 6.8.4. Results

The first table (39) shows the most relevant results of STSQ mean scores comparisons between first and last year high school students. The mean differences were tested with Mann-Whitney test of statistical significance. Table 39 also shows three results which were marginal significant. The LOC score showed a statistically significant mean difference but from a theoretical point of view this difference seemed to be plausible as will be discussed later. The extended O+LOC score showed a significant difference indicating a similar tendency as the LOC difference. The eSTS pattern and the NFS mean differences are also not statistically significant but they could be indicators which helped to consider the validity of the respective STSQ scores.

**Table 39: Significant and marginal significant mean score results of the comparisons between (a) first and last high school year, (b) abroad experiences (yes or no), and (c) immigration background (yes or no)**

Score	Group	N	Mean	SD	p
RA 2	1. year students	72	1,63	0,17	0,02*
	3. year students	64	1,72	0,16	
LOC	1. year students	69	2,21	0,35	0,06
	3. year students	62	2,08	0,27	
O+LOC	1. year students	68	2.87	0,04	0,02*
	3. year students	61	2.98	0,03	
eSTS pattern score	Abroad experience, YES	6	0,04	0,06	0,07
	Abroad experience, NO	39	0,16	0,16	
LOC	Abroad experience, YES	27	2,0	0,25	0,02*
	Abroad experience, NO	161	2,2	0,32	
O+LOC	Abroad experience, YES	26	3.1	0,05	0,03*
	Abroad experience, NO	159	2.9	0,02	
NFS	Immigration BG, YES	6	3,04	0,19	0,07
	Immigration BG, NO	159	2,9	0,36	

\*. Significant difference in means at the level  $< 0.05$  (Mann-Whitney test).

In addition to the extreme group comparisons, all statistically significant correlations are listed in Table 40. The highest and most significant positive relationship is indicated between eSTS pattern and RA2 score. The lower indicated relationship between RA1 and RA2 (0,34) shows that RA1 and RA2 do not differentiate as theoretically expected within the third year sub-sample.



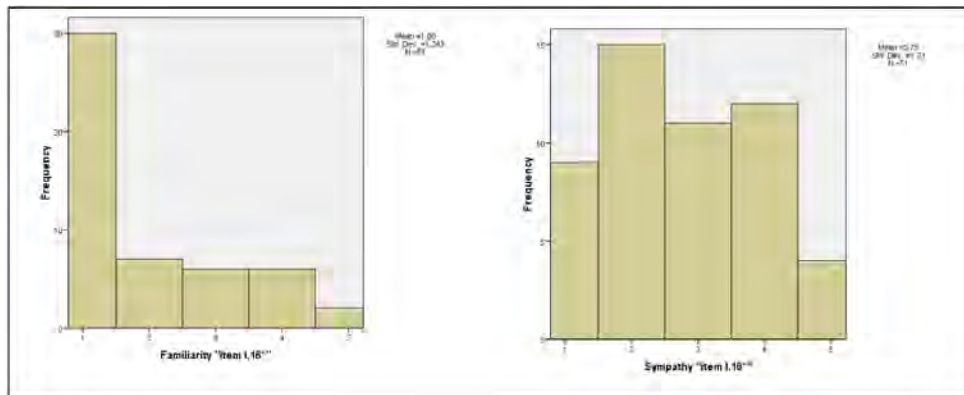
**Table 40: Significant STSQ score inter correlations separated according to high school levels**

High school level	N	STSQ score	Spearman's rho
3	63	RA 1	,34**
		RA2	
3	63	O	-,28*
		RA2	
2	20	RA2	,56**
		STS	
2	66	O	,27*
		RA1	
1	69	RA1	,3*
		RA2	
1	72	O	,24*
		RA2	
1	67	NFS	,25*
		Uncertainty	

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

The next figure shows a result of part I of the STSQ. The histograms shown in Figure 37 illustrate the score distribution of the familiarity and sympathy scale of item 1.16 (Table 3). The score distribution of the familiarity scale is clearly displaced to the right. The sympathy scale is broader distributed along all scores.



**Figure 37: Histogram familiarity and sympathy scale if item 1.16\* (STSQ, part I)**

#### 6.8.5. Discussion and conclusions

Students generally develop intellectually through the three years of high school. It was expected that students of the third high school year were more rational than students of the first year.

The significant RA2 mean score difference shown in Table 39 indicated that the pupils of the third high school year show higher RA2 score results than first year students. The RA2 items then indicate a higher rationality level compared to RA1. RA2 also showed a positive correlation with the eSTS pattern score results which could be interpreted as an indicator for the intention that RA2 items measure awareness of more rational argumentations in strangeness related situations (cf. Table 40).

The measured significant group difference of LOC score between students with abroad experiences ( $LOC_{MEAN}=2,0$ ) compared to no abroad experiences ( $LOC_{MEAN}= 2,2$ ) shown in Table 39 appeared reasonable as well. Abroad experience can have a positive influence on self-concept whereby LOC is assumed to be negatively related to self-concept (cf. Giess-Stüber, 2008; cf. Grimminger, 2009). It could be argued that pupils who take the “risk” of traveling abroad more than two months may feel more secure in new surroundings than pupils with no abroad experiences. In addition, traveling a longer period abroad can be challenging (e.g. regarding language, contacts, orientation etc.). Managing daily tasks in a new environment may have a positive influence on one’s own identity.

A correlation analysis between the different STSQ scores separated by high school years showed that the NFS score correlated low but significantly with the uncertainty score of situation 5 (part II):  $\rho = 0,25^{53}$ .

The LOC and O score correlated low ( $\rho = -0,2$ ) in all sub-samples with each other but not statistically significant with non-parametric correlations<sup>54</sup>. This result was consistent with previous results and supported the idea to summarize the O and LOC scale to one O+LOC score.

Further theoretical considerations and interpretations of empirical results led to additional changes of items. The following changes should make the items easier to understand, and the intended meaning of items was pointed out. The following table shows all revisions according to previous results and reconsiderations:

---

<sup>53</sup> Significant correlation ( $p < 0,05$ ),  $N=67$



<sup>54</sup> Pearson correlation indicated a significant correlation:  $r=-0,21$  with  $p < 0,01$

*STSQ, part I:*

Single item analysis showed that the following pictures are suitable as item stimulus in order to provoke a perception of difference: I.1, I.3, I.12, I.13, I.16 (cf. Table 3).

The items of part I were changed as shown in Table 41:

**Table 41: Item revisions STSQ, part I**

Present item structure	New, revised item structure
<p>9 a) Please give a headline to the following picture!</p>  <p>9 b) Please indicate how familiar (1.), <u>and</u> how sympathetic (2.) you feel, when you imagine the situation illustrated above. Please cross out the respective number:</p> <p>1. Familiarity:      low 1 2 3 4 5 high</p> <p>2. Sympathy:      low 1 2 3 4 5 high</p>	<p>5 a) Please give a headline to the following picture!</p>  <p>5 b) Please cross out the respective number!</p> <p>1. How <u>different</u> do you perceive yourself compared to the person illustrated in the picture above? low 1 2 3 4 5 high</p> <p>2. How <u>sympathetic</u> do you feel, when you imagine the person illustrated above? low 1 2 3 4 5 high</p>

The picture shown in the item example above seems to meet the intended stimulus criteria in an optimal way compared to other selected item images. The shown person on this item was difficult to classify for most of the respondents which therefore usually was indicated by an attribution of differences (higher scores on the difference scale). The results shown in (Figure 37) supported this impression. This item provoked a perception of difference which was needed in order to measure STS whereby the sympathy scale showed a more balanced distribution.

The scale instructions were changed in order to make the items easier to understand with respect to what was intended to measure i.e. perception of difference and an attribution of sympathy.

*STSQ, part II:*

The present STSQ version of this part II used dichotomous scale level as mentioned previously. To increase the variance, I decided to change back the scale level of these items to the bi-polar interval scale level. For application purposes, a dichotomous scale might be easier

to analyze, but with the aim to test reliability and validity aspects, interval scale level bear more information than dichotomous items.

The following table shows further smaller revisions of item formulations. Table 42 shows both the old and revised item formulation. The table indicates in addition a short reason for the revisions being made.

**Table 42: Further item adaptations, part II of the STSQ**

Item label	Present item	Item label	Revised item	Comment/reason for revision
1.6	The person's behavior seems suspicious.	1.6.1	The passenger appears suspiciously.	Point out the more rational meaning/ RA2 membership with the aim to make RA2 more distinctive compared to RA1.
1.1.1	The person looks different.	1.1.2	The passenger looks different.	"Passenger" appears to be a more precise description of a person entering a bus.
1.2	Generally, strangers are controlled.	1.2.1	Strangers are often asked to show their tickets.	Linguistic revision which may increase the correlation to the same scale items
2.5	The player was not integrated into the new team.	2.5.1	The new player was not yet integrated into the team.	The supplied term "yet" is supposed to point the meaning of this item with respect to increase the correlation with other RA2 items with the aim to make RA2 more distinctive compared to RA1.
2.6	The team was in a poor condition.	2.6.1	The team has to work on a new game strategy.	The initial item showed no reasonable correlations. The new item is supposed to correlate with the RA2 items because this statement is supposed to be more rational than the initial version.
3.6	P.E. can be arranged in a mutually acceptable way.	3.6.1	Physical Education can be arranged in a mutually acceptable way.	P.E. was not always associated with Physical Education.
3.4	Students have the possibility to select exercises.	3.4.1	Students are given the opportunity to choose between exercises in Physical Education class.	The reformulation of this item is supposed to make it easier to understand.
3.3	P.E. class could be organized gender-separated.	3.3.1	Physical Education class could be organized gender-separated.	P.E. was not always associated with Physical Education.
4.5	In case of equal qualification, the gender-balance determines the decision.	4.5.1	In case of equal qualification, the balance of gender within the company determines the decision.	Linguistic revisions
4.6	The heterogeneity of the company influences the decision.	4.6.1	The heterogeneity within the company influences the decision.	Linguistic revisions

*STSQ, part III:*

The next table shows further item adaptations of the third part of the STSQ.

**Table 43: Further item adaptations, part III of the STSQ**

Item label	Present item	Item label	Revised item	Comment/reason for revision
LOC5.1	Life brings continuously changes.	LOC5.1.1	Life brings changes continuously.	Linguistic revisions
O2	The open result is responsible for the excitement of a competition.	O 2.1	The uncertain outcome is responsible for the excitement of a competition.	Linguistic revision
NFS3	Being obliged to make decisions provokes uncertainty.	NFS3.1	Having to make decisions makes me feel uncertain.	Linguistic revisions
		NFS5	Having responsibility makes me feel uncertain.	Additional NFS item which is expected to correlate positively with the NFS3.1 item. In particular, having to make decisions is related to responsibility in general. A significant correlation would support this assumption, and complement to the other NFS items i.e. "responsibility" which makes one feel uncertain.

Analyses of this data collection in a Norwegian high school, helped to adapt STSQ items and item instructions further as illustrated above. The comparisons of STSQ scores between the different high school levels showed that the measurement of part II and III varied reasonably with high school level. The uncertainty related scores of the STSQ (LOC, NFS) showed a higher uncertainty in the first high school year than in the last. Higher grades scored significantly higher on the more rational items of part II of the STSQ (measured by the RA2 score).

Correlations in all three sub-samples of the different high school levels consistently indicated negative correlation between O and LOC score and may support the idea to combine the O and LOC score to one O+LOC score.

One crucial result was shown in Figure 37. The score distribution of the familiarity and sympathy scale represented an optimal item characteristic as indicator for part I of the STSQ (emotional STS). The picture shown of this item was the most representative stimulus for a perception of strangeness.

In addition, the RA scales appeared to be most consistent between the three high school years even though the total Cronbach's alpha reached only moderate values.

### 6.9. Application of the STSQ within the Comenius project 2.1 “The development of intercultural competence through sports in an expanding European Union”<sup>55</sup>

Besides the intended evaluations of previous adaptations of the STSQ, this data collection was conducted within a theoretically related EU-project with the title “The development of intercultural competence through sports in an expanding European Union” (Gieß-Stüber & Blecking, 2008). The specific design was supposed to gain additional knowledge about the STSQ to help improving the STSQ items further.

The general goal of the EU-project was a survey of existing “movement cultures” within the participating countries. The research group<sup>56</sup> intended to examine the field from a historical, sociological, and educational perspective. They emphasized especially the role of minorities. The project was expected to generate knowledge required for a development of teaching and learning methods in the realm of teacher training.

This EU-project included three events or training weeks once a year with an international group of teacher students from the participating countries Germany, France, the Czech Republic and Poland. Each year, another country hosted the training weeks in their country. The training weeks were conducted in order to test in practice the mediation of developed concepts of sport and integration within an international group of students. The training weeks were in addition aiming at a promotion of an intercultural competence of future P.E. teachers. The conducted training weeks cannot directly be understood as intervention. The explorative character dominated the investigations. The project leaders were interested in discovering how the students reacted towards a new and unusual approach in the context of sport and physical education.

The following ideas were background and guidelines for the training weeks of the EU-project (Gieß-Stüber, 2008):

- Authentic experience as subjective reference for reflections and insight
- Close relationship between theory and practice
- Irritation of so far unquestioned, familiar perceptions of the own person or cultural habits

---

<sup>55</sup> EU project number: 119019-CP-1-2004-1-DE-COMENIUS-C21

<sup>56</sup> Professor Dr. Petra Gieß-Stüber is leader of this EU financed project and one of the initiators of “Intercultural Movement Education”

- Experiencing a variety of new, unusual (strange) and different sport disciplines, activities, movements etc.

Living and learning together for one week within an international group was meant to create the needed learning arena for aforementioned guidelines. Specific arrangements in the context of sport activities (indoor and outdoor), play, and movement were supposed to create emotions and experiences. The experiences were used as initiator for discussions and considerations of basically intercultural learning in and through sport.

The activities and theoretical topics of the training weeks were related and structured according to the following topics (Gieß-Stüber, 2010a:28):

- A) Re-thinking teacher training in an expanding EU
- B) Potential modules for teacher education in the future:
  - i. Sport and migration in historical perspective
  - ii. Sport and integration
  - iii. Sport and regional movement cultures
  - iv. Sport and strangeness
  - v. Sport, ethnicity and gender
- C) Perspectives in the promotion of intercultural competence among student teachers.

An example shall give an idea about how the students were introduced into the first training week (according to Gieß-Stüber, 2008)<sup>57</sup>:

*Topic 1: Getting to know each other - realizing familiarity and strangeness*

All participants were divided into nationally heterogeneous groups. Equipped with Cameras, they were given the task to explore the city, and solve different tasks on their journey through the city together as a group. Solving the tasks required cooperation and communication. The groups were asked to take pictures of things they perceived as strange.

The intention of this game (“*Stadtrallye*”) was directed towards two perspectives: the guests should get familiar with new and unfamiliar environments, whereby the natives should become aware of how the foreigners perceived the native’s environment.

<sup>57</sup> More detailed information of the training weeks can be found on the projects webpage: <http://portal.uni-freiburg.de/sportpaedagogik>



---

Gieß-Stüber (2008) summarized her experiences of the trainings weeks as productive and challenging learning opportunities in an international context. The formal teaching classes of the training weeks were characterized by theoretical considerations around the aforementioned topics and a variety of intensive experiences. The participants were in some cases irritated when for instance being confronted with unusual (as too strict perceived) teaching methods (cf. Gieß-Stüber, 2008). But also the informal contact and interactions between the participants were characterized by intense experiences of uncertainty, frustrations, but also with deeper insights into the phenomenon of strangeness.

#### 6.9.1. Considerations of the given research design

I was invited to collect data with my STSQ within this project because my project refers to the same underlying framework concept of *IME*. An application of the STSQ within these training weeks appeared to be a benefit for both projects – even though the given design was not optimal and my STSQ was still under development. An application within this theoretically related and kind of intervention design was expected to gain some hints of construct validity for my project. Resulting knowledge was expected to develop the STSQ further. The EU-project would receive a feedback on possible effects of their training weeks regarding a STS. The German research team was responsible for the design and data collections with the STSQ and other evaluation instruments. I analyzed the STSQ data according to the aforementioned aims of improving the STSQ items.

I received data from the two first training weeks of the Comenius project. In 2005 the training week was conducted for the first time and in 2006 for the second time with a new group of participants. Each of the training weeks went over the duration of six days once a year. Since the EU-project was conducted parallel to the development of the STSQ, different versions of the STSQ were applied. In 2005, an early version<sup>58</sup> was used, and in 2006 the last version of the STSQ was used for data collection. Even though the goals and design of the EU-project did not follow an optimal intervention design and were not specifically directed to improve STS, the common theoretical baseline of *IME* and related topics within the training weeks made an application of the STSQ attractive for validation purposes.

From a methodological point of view, the differentiation between groups can be used as procedure for construct validation (Crocker and Algina 1986; Lienert and Raatz 1996). For

---

<sup>58</sup> The version of the STSQ used within the repeated measurement design was applied during the first year (cf. chapter 6.7)

instance, contrasting mean scores from different groups would show if they differed in a hypothesized direction and could generate useful information of construct validity. Finding expected differences would consequently support the construct and/adequacy of the instrument vice versa. The aim was to demonstrate that subjects, who have received a specific treatment to alter their standing on the construct, differed from subjects who had received no treatment (Crocker & Algina, 1986). If expected differences were not found, the main possible explanations were failure of the theory underlying the construct, inadequacy of the measuring instrument, or failure of treatment (ibid., 1986).

The following Figure 38 illustrates the given measurement design of the STSQ.

<b>Measure</b> <b>Group</b>	<b>m<sub>1</sub></b>	<b>Intervention</b>	<b>m<sub>2</sub></b>		<b>m<sub>3</sub></b> <b>(follow up)</b>
Intervention Group (IG)	Pre measurement	<b>Six days training program</b>	Post measurement	3 month without specific intervention	Follow up measurement
Control Group (CG)	--	--	CG measurement	--	--

**Figure 38: Measurement design of the STSQ data collection within the EU Comenius project 2.1**

The Comenius project only investigated relative change of the participants regarding aspects of an intercultural competence. Only one control group was therefore needed (Gieß-Stüber, 2010b). The control group was selected according to parallel criteria with respect to the “experimental” group. The criteria for control group selection were participating country, age, sex, semester/study years, and the combination of study subjects. Control group data was collected as far as possible right after each training week. The intervention group was followed by three measurements: The first day of the training (m1), the final day of intervention (m2), and a follow up measurement after three month. Besides measurements with the STSQ, the two additional questionnaires were applied by the project group in Germany within the intervention group; one (Q1) parallel to m2 and the other one (Q2) parallel to the follow up measurement.

The following table shows an overview on the number of participants differentiated by citizenship, and measurement:

**Table 44: Sample structure and size of the intervention design in Germany 2005 and in The Czech Republic 2006**

Country	Intervention group (IG)	Control group (CG)	N <sub>2005</sub>	N <sub>2006</sub>
GER	Pre		11	6
	Post		11	6
	Follow up		--	5
		CG	22	--
FRA	Pre		7	5
	Post		6	5
	Follow up		6	4
		CG	7	5
CZ	Pre		5	6
	Post		4	6
	Follow up		5	6
		CG	10	6
POL	Pre		8	7
	Post		7	7
	Follow up		--	6
		CG	10	--
LUX	Pre		2	--
	Post		2	--
	Follow up		2	--
		CG	--	--
TOTAL, IG			33	24
TOTAL, CG			49	11

### 6.9.2. Measuring instruments

There was one year between each event of the EU-project. The STSQ was under development. Therefore, different versions of the STSQ were applied for each year.

In 2005 an early version of the STSQ was applied. The STSQ included at this stage of development the following items:

#### *STSQ, part I:*

The following picture items were used for each measurement (cf. also Table 3):

**Table 45: Applied item pool of the data material from the EU-project in 2005 (STSQ, part I)**

Measurement	Items
Pre, post, control group	I.1, I.3, I.9
Follow up	I.1, I.3, I.12*, I.13*, I.14*

\* According to the results of following studies, new pictures are added to the item pool (cf. Table 3)

Both the familiarity and the sympathy scale were applied. The eSTS pattern score and the alternative pattern scores were calculated as shown previously.

*STSQ, part II:*

Table 46 shows the item pool applied within this data collection. The STSQ scores of this part are calculated as following:

$$RA1_{mean} = \frac{1.1 + 1.2 + 1.3 + 2.1 + 2.2 + 2.3 + 2.4 + 3.1 + 3.2 + 4.1 + 4.2}{11}$$

$$RA2_{mean} = \frac{1.4 + 1.6 + 2.6 + 2.5 + 3.3 + 3.4 + 3.5 + 3.6 + 4.3 + 4.4 + 4.5 + 4.6}{12}$$

**Table 46: Item pool of part II; the item order is changed according to difficulty index (cf. appendix 9.5)**

Item pool applied for the pre, post, and CG measurement		Item pool applied for the follow up measurement <sup>59</sup>	
<b>Situation 1</b>		<b>Situation 1</b>	
1.1	The person looks different	1.6.8	The passenger appears suspicious
1.2	Generally, strangers are controlled	1.1.1	The passenger looks different
1.3	The bus driver does not like foreigners	1.5	It was a routine control
1.4	The bus driver is irritated	1.2.1	Strangers are often asked to show their tickets
1.5	It was a routine control	1.3.1	The bus driver does not like foreigners
1.6	The person's behaviour seems to be suspicious	1.4	The bus driver is irritated
<b>Situation 2</b>		<b>Situation 2</b>	
2.1	The new player did not play as well as expected	2.5.1	The player was not yet integrated into the new team
2.2	The new player did not adapt to the team	2.2	The new player did not adapt to the team
2.3	The new player disturbed the team spirit	2.1	The new player did not play as well as expected
2.4	The team opposed to the new player	2.4	The team opposed to the new player
2.5	The new player was not integrated into the team	2.3	The new player disturbed the team spirit
2.6	The team was in poor condition	2.6.1	The team has to work on a new game strategy
<b>Situation 3</b>		<b>Situation 3</b>	
3.1	Co-educational teaching should have priority over individual beliefs	3.5	Different religious faiths should be respected
3.2	Pupils should obey existing rules	3.6.1	Physical Education can be arranged in a mutually acceptable way
3.3	P.E. class could be organized gender-separated	3.2	Students should obey existing rules
3.4	Students have the possibility to select exercises	3.4.1	Students are given the opportunity to choose between exercises in Physical Education class.
3.5	Different religious faiths should be respected	3.1	Co-educational teaching should have priority over individual beliefs
3.6	P.E. can be arranged in a mutually acceptable way	3.3	Physical Education class could be organized gender-separated
<b>Situation 4</b>		<b>Situation 4</b>	
4.1	Men are supposed to be more efficient	4.4	Professional qualification determines the decision
4.2	Women are supposed to be more integrative leaders	4.3	Personal leadership determines the decision
4.3	Personal leadership determines the decision	4.5.1	In case of equal qualification, the balance of gender within the company determines the decision
4.4	Professional qualification determines the decision	4.2	Women are supposed to be more integrative leaders
4.5	In case of equal qualification, the gender-balance determines the decision	4.1	Men are supposed to be more efficient
4.6	The heterogeneity of the company structure will influence the decision	4.6.1	The heterogeneity within the company influences the decision.
<b>Situation 5</b>		<b>Situation 5</b>	
5.1	This situation does not provoke specific irritations	5.4.1	The new surrounding makes him/her feel uncomfortable
5.2.1	It is difficult to discover the expecta-	5.6	He/she feels insecure and disorientated

<sup>59</sup> Besides some revisions of item formulation, the item order has changed according to increasing difficulty index within each situation.

	tions of the new surroundings		
5.3	He/she is confident to overcome the uncomfortable feelings of the first day	5.2.1	It is difficult to discover the expectations of the new surrounding
5.4.1	The new surrounding makes him/her feel uncomfortable	5.3	He/she is confident to overcome the uncomfortable feelings of the first day
5.5	Getting to know people in such an impersonal situation must be difficult	5.1	This situation does not provoke specific irritations
5.6	He/she feels insecure and disorientated	5.5	Getting to know people in such an impersonal situation must be difficult

*STSQ, part III:*

The next table shows the used item pool of part III of the STSQ.

**Table 47: Applied item pool of part III**

Item Code	Item
O 1.1:	Open-mindedness helps understanding.
O 1:	It seems that different people have different values.
O 2:	The open result is responsible for the excitement of a competition.
O 3.1:	To meet different kinds of people is enjoyable.
O 4*:	Sitting in a group of strangers provokes uncomfortable feelings.
O 5.1:	Freedom does not mean the absence of rules.
NFS 1:	It is important to forward cultural values to the next generations.
NFS 2.1:	To plan ahead helps to provide security.
NFS 3:	Being obliged to make decisions provokes uncertainty.
NFS 4.1:	When engaging in an activity, I prefer receiving clear instructions to open suggestions.
LOC 1*:	The number of friends I make depends only on me and my behavior.
LOC 2:	Most of the events in my life are determined by other people.
LOC 3:	Fate determines whether I have more or fewer friends.
LOC 4.1*:	People differ substantially.
LOC 5.1*:	Life brings continuously changes.
LOC5.2*:	Culture has some stable values.
LOC 6.1*:	I like unforeseen events.

\*inverted item formulation with respect to the regarding category

The STSQ scores of part III were calculated based on previous results. The STSQ scores were summarized as shown in Table 48 and 49.

**Table 48: Item sub-scores (O, NFS, and LOC score)**

Openness (O) score	Need for security (NFS) score	Loss of control (LOC) score
O1.1	NFS1	LOC2
O2	NFS2.1	LOC3
O3.1	NFS3	LOC4.1
O4*	NFS4.1	LOC5.1*
O5.1		LOC5.2*
		LOC6.1*

\*inverted item formulation with respect to its regarding category

**Table 49: Item scores based on the results of the last study**

Openness (O) score	Need for security (NFS) score
O1.1	NFS1
O2	NFS2.1
O3.1	NFS3
O4*	NFS4.1
O5.1	
LOC2*	
LOC3*	
LOC4.1*	
LOC5.1	
LOC5.2	
LOC6.1	

\*inverted item formulation with respect to its regarding category

The following three items were selected from the additional questionnaire of the EU-project because they seemed to be relatable to the STSQ, and could gather some construct validity hints:

**Table 50: Items of an additional questionnaire<sup>60</sup>**

Item code	Item
Item iii 1	Did your attitude towards the issue of intercultural life change since the course in September?
Item iii.4	How do you estimate the probability of success of endeavours to find peaceful solutions to conflicts between social groups?
Item iii 8	Did your attention or behaviour linked to the topics of the course in September change after that event?"


The latest version of the STSQ (cf. appendix 9.8.2) was applied to collect data at the event in 2006 in the Czech Republic.

The STSQ scores were calculated as illustrated in the chapters before including all adaptations suggested according to previous results. Part I consisted finally of five items: I.1, I.3, I.12, I.13, and I.16 (cf. Table 3). Each item included two ranking scales: a) difference scale and b) sympathy scale (cf. Figure 39).

<sup>60</sup> The selected items originate from an unpublished evaluation questionnaire of this EU-project (119019-CP-1-2004-1-DE-COMENIUS-C21).

The sensitivity score was calculated as shown in the following example:

**5 a) Please give a headline to the following picture!**



\_\_\_\_\_

\_\_\_\_\_

**5 b) Please cross out the respective number!**

1. How different do you perceive yourself compared to the person illustrated in the picture above?

low 

1	2	3	4	5
---	---	---	---	---

 high

2. How sympathetic do you feel, when you imagine the person illustrated above?

low 

1	2	3	4	5
---	---	---	---	---

 high

**Figure 39: STS response pattern score**

In order to measure STS-response pattern, a respondent has to indicate score 4 or 5 on the first scale (difference scale) as shown in Figure 39. The respondent needs, additionally, to indicate score 3 or 4 on the second scale (sympathy). The respondent then gets one eSTS pattern score on the respective picture item. The mean score over all five items then represents the measured eSTS pattern result.

Part II of the STSQ was represented by RA1 and RA2 mean scores. The following tables illustrate which items were used for score calculations. The scales were now changed back to a bi-polar scale with two intervals in both directions as illustrated in the following item example:



**Part II: Put yourself into the following situations!**

1) *Imagine people are entering a bus. All passengers pass by the conductor. One passenger was asked to show his/her ticket.*

**What reasons could the bus driver have to stop this passenger?  
Please indicate the probability of each suggested reason!**

1. The passenger appears suspiciously.	1 most improbable	2 improbable	3 probable	4 most probable
2. The passenger looks different.	1 most improbable	2 improbable	3 probable	4 most probable
3. It was a routine control.	1 most improbable	2 improbable	3 probable	4 most probable
4. Strangers are often asked to show their tickets.	1 most improbable	2 improbable	3 probable	4 most probable
5. The bus driver does not like foreigners.	1 most improbable	2 improbable	3 probable	4 most probable
6. The bus driver is irritated.	1 most improbable	2 improbable	3 probable	4 most probable

Further reasons \_\_\_\_\_

**Figure 40: Item example part II**

The scores were calculated as the following formula illustrates and include the last revised items (cf. Table 42):

$$\overline{RA1}_{arith. mean} = \frac{1}{n} \sum_{i=1}^n RA1_i = \frac{1 \cdot 1.2 + 1.2 \cdot 1 + 1.3 + 1.4 + 2.1 + 2.2 + 2.3 + 2.4 + 3.1 + 3.2 + 4.1 + 4.2}{12}$$

**Table 51: RA1 score items (nested within and between situations)**

Item code	Item
1.1.2	The passenger looks different.
1.2.1	Strangers are often asked to show their tickets.
1.3	The bus driver does not like foreigners.
1.4	The bus driver is irritated.
2.2	The new player did not adapt to the team.
2.1	The new player did not play as well as expected.
2.4	The team opposed to the new player.
2.3	The new player disturbed the team spirit.
3.2	Students should obey existing rules.
3.1	Co-educational teaching should have priority over individual beliefs.
4.2	Women are supposed to be more integrative leaders.
4.1	Men are supposed to be more efficient.

The following formula illustrates the calculation of the RA2 scores and the subsequent Table 52 shows the items in detail:

$$\overline{RA2}_{arith. mean} = \frac{1}{n} \sum_{i=1}^n RA2_i = \frac{1.6.1 + 2.5.1 + 2.6.1 + 3.3.1 + 3.4.1 + 3.5 + 3.6.1 + 4.3 + 4.4 + 4.5.1 + 4.6.1}{11}$$

**Table 52: RA2 score items (nested within and between situations)**

Item code	Item
1.6.1	The passenger appears suspiciously.
2.5.1	The new player was not yet integrated into the new team.
2.6.1	The team has to work on a new game strategy.
3.5	Different religious faiths should be respected.
3.6.1	Physical Education can be arranged in a mutually acceptable way.
3.4.1	Students are given the opportunity to choose between exercises in Physical Education class.
3.3.1	Physical Education class could be organized gender-separated.
4.4	Professional qualification determines the decision.
4.3	Personal leadership determines the decision.
4.5.1	In case of equal qualification, the balance of gender within the company determines the decision.
4.6.1	The heterogeneity within the company influences the decision.

The items of part three of the STSQ were structured as the following tables illustrate:

**Table 53: O score items**

Item code	Item
O1.1	Open-mindedness helps understanding.
O2.1	The uncertain outcome is responsible for the excitement of a competition.
O3.1	To meet different kinds of people is enjoyable.
O4*	Sitting in a group of strangers provokes uncomfortable feelings
O5.1	Freedom does not mean the absence of rules.

**Table 54: LOC score items**

Item code	Item
LOC2	Most of the events in my life are determined by other people.
LOC3	Fate determines whether I have more or fewer friends.
LOC4.1	People differ substantially.
LOC5.1.1*	Life brings changes continuously.
LOC5.2*	Culture has some stable values.
LOC6.1*	I like unforeseen events.

**Table 55: NFS score items**

Item code	Item
NFS1	It is important to forward cultural values to the next generations.
NFS2.1	To plan ahead helps to provide security.
NFS3.1	Having to make decisions makes me feel uncertain.
NFS4.1	When engaging in an activity, I prefer receiving clear instructions to open suggestions.
NFS5	Having responsibility makes me feel uncertain.

Each facet was represented by the items shown in Table 53 - 55. Mean score were calculated for each score separately. The summarizing O+LOC mean score was calculated as well. When calculating the scores, all items marked with a "\*" needed to be re-coded in an inverted way consistent with the other items. The label of each facet determined the direction of measurement.

### 6.9.3. Expectations

It was expected to measure mean score differences between the different measurements of the intervention groups. A comparison of results of the intervention group with the results of the control group, should allow inferring carefully the relative mean score changes of the intervention project. If measured differences were consistent with anticipated results they could be supportive indicators of validity because the event activities were based on similar theoretical assumptions.

Related items of the applied additional questionnaire were expected to gain some useful knowledge about the STSQ's construct validity indicated by reasonable correlations with and STSQ scores.

The follow up measurement was conducted about 4-6 months after the training program was carried out. The idea was to measure possible long term learning effects. The third measurement could on the other hand give some further hints on the stability over time of the STSQ. According to McDonald (1999) it is best to obtain a set of retest correlations over a series of increasing time intervals, if we wish to study either the stability of the measurement or course that it follows through time. In this sense, the design shown in Figure 38 could help to approximately estimate the stability in addition to the re-test measurement results shown in chapter 6.7.

#### 6.9.4. Procedure

STSQ scores were calculated as mean scores because they were easier to compare directly when the number of items is different or unbalanced between measurements. Mean scores were compared between the different measurements overall and between sub-groups separated according to citizenship. This procedure was supposed to make possible differences clearer. Otherwise, the heterogeneity of the whole sample could cover possible effects of intervention so that mean differences would not become obvious. Statistics were therefore calculated for the whole intervention/ control group sample, and for each sub-group according to the respondents indicated country of origin.

Mean score differences were tested for statistical significance with non-parametric inference statistics. The Wilcoxon signed rank test was applied when comparing possible differences in mean scores of the pre, post, and follow up measurement of the intervention (dependent) sample (Bortz & Lienert, 2003). All measurements were finally compared pair wise with the results of the control group. In order to insure that possible differences in STSQ scores between measurements were sufficiently reliable, the repeated measurements needed to be examined by its reliability when applying the Wilcoxon signed-rank test (ibid., 2003). For this purpose, the stability results might function as additional reference with respect to test-retest reliability.

Mann-Whitney U-test was applied for comparisons of mean scores between intervention and control group samples (ibid., 2003).

Correlation between the same scores of different measurements (pre, post, and follow up measurement) was supposed to show how stable the measuring results appear empirically. It

---

was expected that the results show some instabilities over time. It needed to be considered how far the intervention activity could be responsible for possible inconsistencies.

The following table illustrates the statistical procedures used in order to estimate measured group differences, and relationships between different STSQ scores:

**Table 56: Statistical methods used for the last analyses**

<b>Goal</b>	<b>Statistical method/index</b>
Comparisons of STSQ scores	Spearman's rho
Comparison of dependent samples	Wilcoxon signed rank test
Comparison of independent samples	Mann-Whitney U-test

## 6.9.5. Results

I start by presenting the results of the 2005 sample and continue with the 2006 results. The following table summarizes the clearest results of group difference analysis. Most of the results in Table 57 are statistically significant and could possibly reveal some useful information about the STSQ's ability to differentiate between groups.

**Table 57: Significant and marginal significant group differences<sup>61</sup> of the 2005 sample**

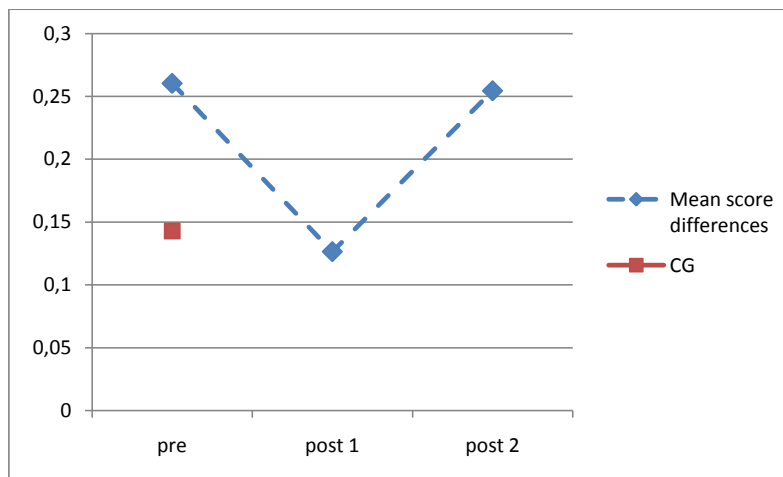
Score	Sample	Measurements	N	Mean	SD	p
eSTS pattern	ALL	Pre	32	0,26	0,34	0,02*
		Post	29	0,13	0,19	
eSTS pattern	ALL	Pre	32	0,26	0,34	0,03*
		CG	49	0,14	0,19	
eSTS pattern	ALL	Follow up	11	0,25	0,13	0,02*
		CG	49	0,14	0,19	
Xeno pattern	ALL	Follow up	11	0,25	0,2	0,01**
		CG	49	0,11	0,22	
Exo pattern	ALL	Follow up	11	0,07	0,13	0,04*
		CG	49	0,31	0,32	
RA1	CZ	Follow up	4	2,7	0,14	0,06
		CG	6	2,82	0,17	
RA 2	CZ	Pre	5	2,87	0,35	0,04*
		Follow up	5	2,98	0,33	
RA 2	ALL	Pre	33	2,8	0,3	0,03*
		Follow up	11	2,84	0,3	
LOC	ALL	Post	30	2,4	0,3	0,02*
		Follow up	11	2,06	0,32	
LOC	CZ	Pre	5	2,23	0,32	0,04*
		Post	5	2,6	0,28	
LOC	CZ	Post	5	2,6	0,28	0,01**
		CG	6	2,28	0,31	

\*. Mean difference is significant at the 0.05 level

\*\* . Mean difference is significant at the 0.01 level

<sup>61</sup> All groups and sub-group results are illustrated in the appendix 9.2

The following Figure 41 illustrates the results of eSTS pattern scores of the intervention group in comparison to the control group. The control group result is similar to the eSTS pattern score result of the post measurement. Both pre and follow up (follow up) measurement results are similar. The pre and post (1) measurement, and the pre and control group result differ significantly (cf. Table 57).



**Figure 41:** eSTS mean score differences of the intervention group in comparison to control group mean score (2005 sample)

Table 57 shows further that the LOC score at the follow up measurement of the IG is significantly lower than at the post measurement which does not differ significantly from the pre measurement result.

The sample of the third measurement within the IG (N=13) showed some correlations. The xenophobia response pattern score showed a significant and positive correlation with the RA1 mean score ( $\rho=0,85^{**}$ ), and with the LOC mean score ( $\rho=0,58^*$ ) as well. In addition, the NFS mean score showed a negative and significant correlation with the O mean score ( $\rho=-0,81^{**}$ ).

The NFS mean score correlated moderately with a question from the additional questionnaire: NFS showed a correlation of  $\rho=-0,53^*$  with item iii.4 “How do you estimate the probability of success of endeavours to find peaceful solutions to conflicts between social groups?”. The indicated negative correlation is reasonable because item iii4 is supposed to indicate a respondent’s optimistic attitude of finding peaceful solutions between social groups.

The following tables show the correlations of each STSQ score between each measurement (pre, post and follow up). Spearman's rho is used as an indicator for the stability over time between the three measurements of the intervention group. It seems that instabilities reflect the measured differences as shown in the tables above.

Table 58 shows the correlations of the eSTS pattern score. The follow up result correlates negative with pre and post (1) results.

**Table 58:** Correlation results (Spearman's rho) of the eSTS response pattern scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=11)

eSTS pattern score	m1	m2	m3
m 1	--		
m2	,30	--	
m3	-,35	-,39	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

The correlations in Table 59 indicate moderate positive relationship between first and second (0,51), and between second and third (0,46).

**Table 59:** Correlation results (Spearman's rho) of the RA1 mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=10)

RA1 scores	m1	m2	m3
m 1	--		
m2	,51**	--	
m3	,35	,46	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).



Table 60 shows that the RA2 scores of measurement one correlate positive and quite high (ca. 0,8) with RA2 scores of the second and third measurement.

**Table 60:** Correlation results (Spearman's rho) of the RA2 mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=33; m2: post measurement, N=30; m3: follow up measurement, N=11)

RA2 scores	m1	m2	m3
m1	--		
m2	,79**	--	
m3	,79**	,55	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

Table 61 indicates positive and moderate relationships between the three O-scores of all three measurements.

**Table 61:** Correlation results (Spearman's rho) of the O mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=27; m3: follow up measurement, N=10)

O scores	m1	m2	m3
m1	--		
m2	,51**	--	
m3	,52	,63	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

The next Table 62 indicates almost no relationship between the NFS scores of the three measurements.

**Table 62:** Correlation results (Spearman's rho) of the NFS mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=11)

NFS scores	m1	m2	m3
m1	--		
m2	,24	--	
m3	,21	,01	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

Table 63 indicates a moderate positive relationship between LOC scores of the first and second measurement.

**Table 63:** Correlation results (Spearman's rho) of the LOC mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=32; m2: post measurement, N=29; m3: follow up measurement, N=11)

LOC scores	m1	m2	m3
m1	--		
m2	,49**	--	
m3	,04	,22	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

Table 64 shows that the last (summarized) O+LOC score shows relative low correlations between the three measurements.

**Table 64:** Correlation results (Spearman's rho) of the O+LOC mean scores between the three measurements of the intervention group from 2005 (m1: pre measurement, N=30; m2: post measurement, N=25; m3: follow up measurement, N=10)

O+LOC scores	m1	m2	m3
m1	--		
m2	,27	--	
m3	,32	,34	--

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

A final result of the data of 2005 is shown in Table 65. The table shows correlations between different STSQ scores. Except result number one and six shown in Table 65, almost all results are based on the post measurement (post1) sample. All shown results in Table 65 indicate reasonable relationships between different STSQ scores and can therefore be interpreted as supportive indicators for construct validity (Crocker & Algina, 1986).

**Table 65: Correlations between STSQ scores**

Nr.	STSQ mean scores	Measurement	N	Correlation (Spearman's rho)
1.	STS O	Post Follow up	10	+0,73*
2.	RA 2 O	Post Post	28	+0,44
3.	RA 1 NFS	Post Post	30	+0,5**
4.	STS LOC	Post Post	29	-0,54**
5.	STS O+LOC	Post Post	28	+0,52**
6.	STS O+LOC	Follow up Post 3	10	+0,83**

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\*. Correlation is significant at the 0.01 level (2-tailed).

The following results are based on the data from 2006 sample. Table 84 (appendix 9.7) gives an overview over all measured mean score results of the different STSQ score, overall and country specific. The chart shown in Figure 42 illustrates STSQ mean score differences between the different groups. Figure 43 illustrates the results between different measurements.

All statistically significant results and marginal significant results between IG group results and CG measurements are listed in Table 66. Firstly, the table shows the results of all respondents (not separated by countries). Secondly, the results of country specific analyses follow in the same table. Figure 43 illustrate the measured group differences graphically and points out interesting patterns which will be discussed in the next sub-chapter.

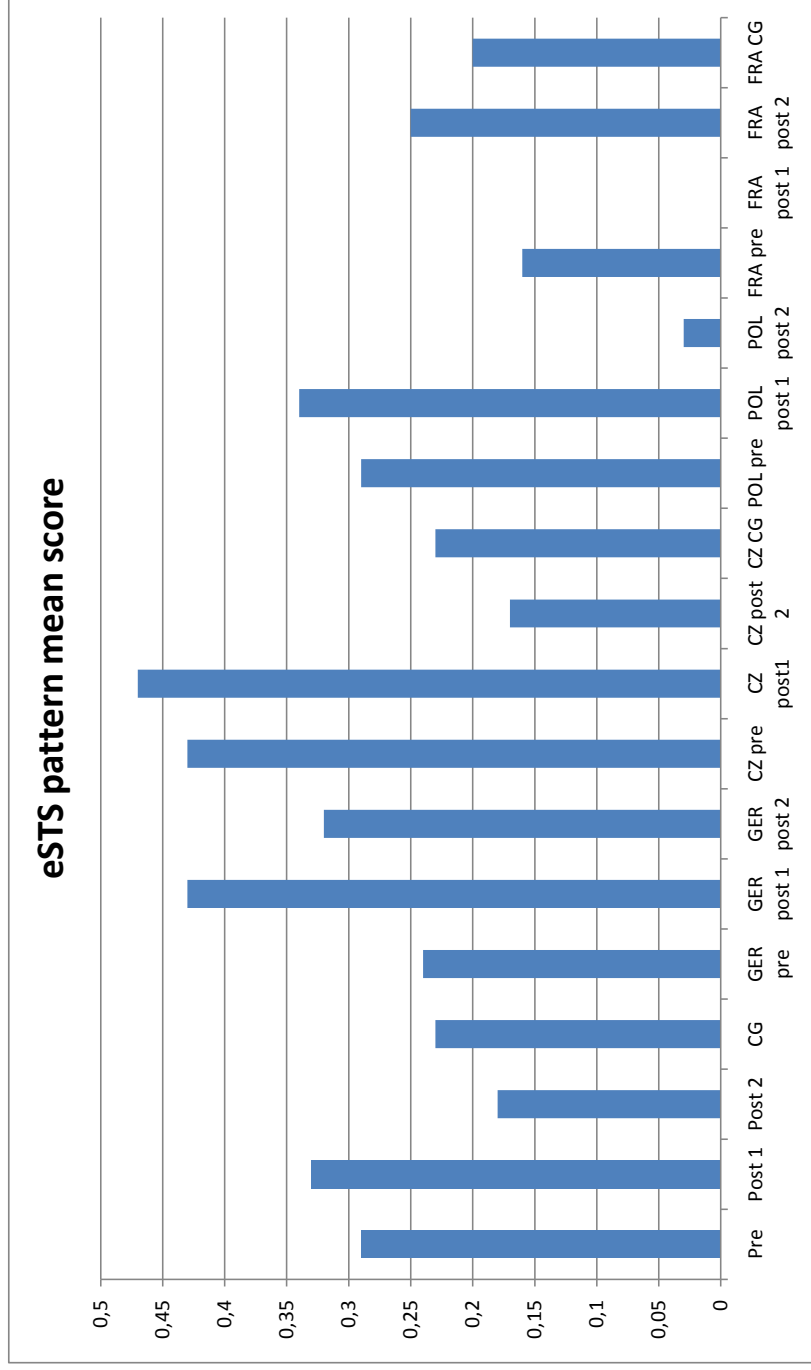


Figure 42: Bar chart of STS mean scores whole 2006 sample and country specified

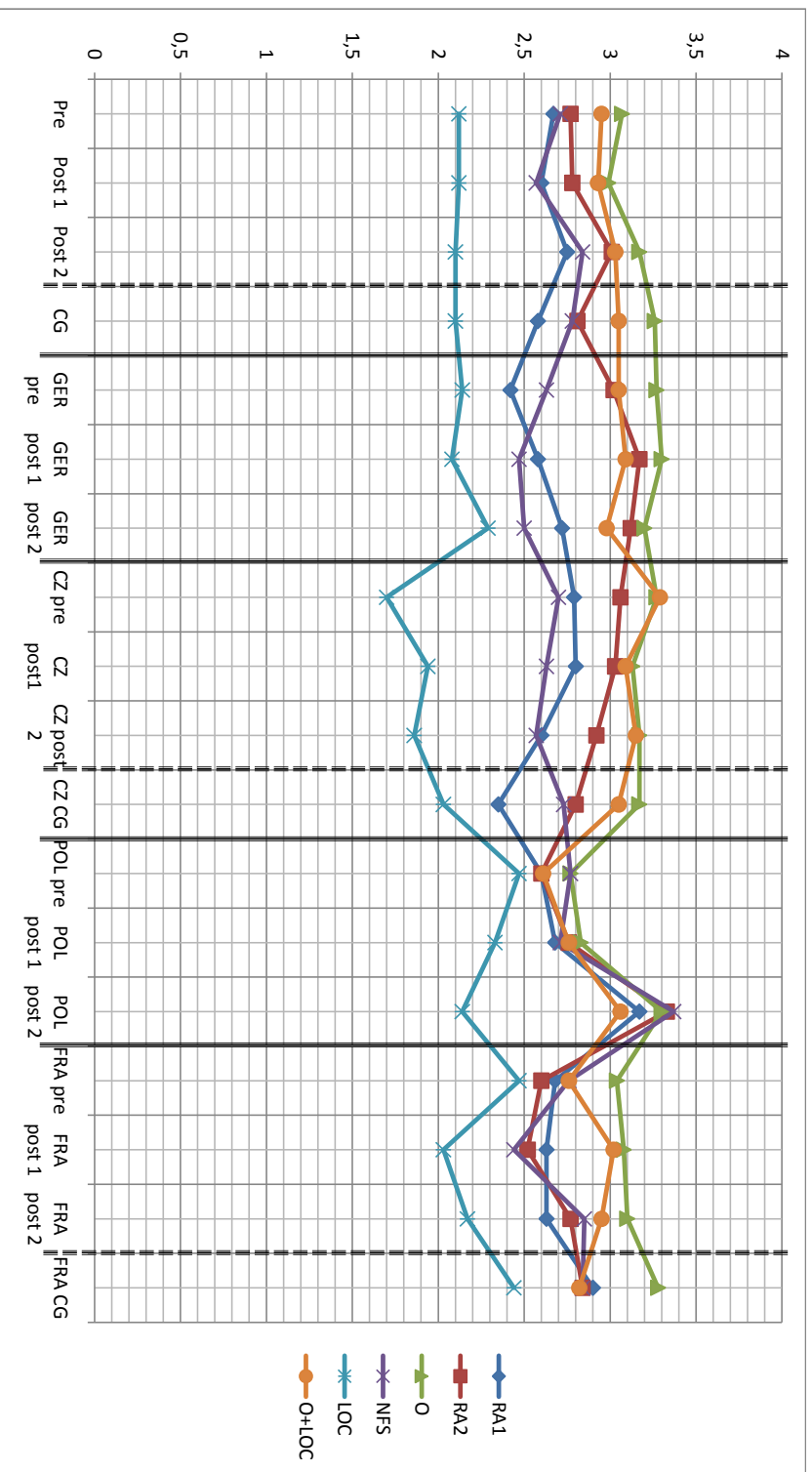


Figure 43: STSQ scores of all measurements of the 2006 sample (overall and country specified)<sup>62</sup>

<sup>62</sup> The black demarcation lines between IG and CG and between different countries demarcate the respective different types of samples.

**Table 66: Significant and marginal significant group differences of the 2006 sample<sup>63</sup>**

Score	Sample	Measurement	N	Mean	SD	p
eSTS pattern	ALL	Post	24	0,33	0,29	0,05*
		Follow up	21	0,18	0,22	
RA 2	ALL	Pre	23	2,80	0,36	0,02*
		Follow up	20	3,07	0,32	
RA 2	ALL	Follow up	20	3,07	0,32	0,03*
	ALL	CG	12	2,81	0,25	
NFS	ALL	Pre	24	2,72	0,39	0,02*
		Post	24	2,58	0,44	
eSTS pattern	CZ	Pre	6	0,43	0,23	0,07
		Follow up	6	0,17	0,08	
eSTS pattern	CZ	Post	6	0,47	0,21	0,02*
		Follow up	6	0,17	0,08	
RA 1	CZ	Pre	6	2,79	0,35	0,03*
		Follow up	6	2,60	0,31	
RA 1	CZ	Post	6	2,81	0,20	0,04*
		Follow up	6	2,60	0,31	
RA 1	POL	Pre	7	2,60	0,25	0,03*
		Follow up	6	3,17	0,22	
RA 1	POL	Post	6	2,68	0,23	0,04*
		Follow up	6	3,17	0,22	
RA 2	POL	Pre	7	2,56	0,24	0,04*
		Post	5	2,76	0,26	
RA 2	POL	Pre	7	2,56	0,24	0,03*
		Follow up	6	3,33	0,16	
O	POL	Pre	7	2,77	0,29	0,04*
		Follow up	6	3,30	0,11	
O	POL	Post	6	2,83	0,23	0,04*
		Follow up	6	3,30	0,11	
NFS	POL	Pre	7	2,77	0,37	0,04*
		Follow up	6	3,37	0,32	
NFS	POL	Post	7	2,71	0,34	0,06
		Follow up	6	3,37	0,32	
LOC	POL	Pre	6	2,47	0,20	0,07
		Follow up	6	2,14	0,29	
O+LOC	POL	post	6	2,76	0,20	0,07
		Follow up	6	3,06	0,16	
O+LOC	POL	Pre	6	2,61	0,21	0,04*
		Follow up	6	3,06	0,16	
eSTS pattern	GER	Pre	5	0,24	0,26	0,06
		Post	6	0,43	0,29	
LOC	GER	Post	6	2,08	0,27	0,06
		Follow up	4	2,29	0,25	
RA2	FRA	pre	5	2,60	0,31	0,07
		Post	5	2,53	0,24	
RA2	FRA	Post	5	2,53	0,24	0,07
		Follow up	4	2,77	0,26	
NFS	FRA	Pre	5	2,76	0,26	0,07
		Post	5	2,44	0,50	

\*. Mean difference is significant at the 0.05 level

\*\*. Mean difference is significant at the 0.01 level

<sup>63</sup> Control group data is only received by The Czech Republic and France

#### 6.9.6. Discussion and conclusions

I mentioned earlier that the design of this intervention did not meet optimal conditions regarding an improvement of STS and control of different variables. This intervention can be described as quasi-experimental because it neither used a random design nor sufficient control group criteria (cf. Shadish et al., 2002). The control group measurements could have been made parallel to the intervention group. But such circumstances are often difficult to construct in practice. In addition, the EU-project was designed as an explorative study as indicated at the beginning of this sub-chapter. However, I tried to use the given circumstances as far as possible because the theoretical relationship between my project and the EU-project suggested that the EU-project could gain useful information about the STSQ measuring characteristics even though the design may not allow strong inferences about the effects of the conducted training weeks within this EU-project. According to Shadish, Cook and Campbell (2002:103ff.) it is possible to gain meaningful conclusions about the measuring results within a quasi-experimental design even though the design does not provide optimal control of circumstance or variables. Lacking control variables within a design can to some extent be “compensated” by stricter statistical testing. The theory-based procedure and application of non-parametric procedure were supposed to reduce a potential risk of a confirmation bias. As I argued previously, using non-parametric procedures for my analyses was a kind of more conservative and critical statistical testing of the results. In addition, the results need be interpreted in relation to the theoretical background. The theoretical pre-assumptions can in this sense be understood as a further control instance by comparing theoretical expectations with empirical results. But all interpretations of the results underlie more uncertainty in comparison to better controlled designs. Therefore, my interpretations refer to the sample and hardly be generalized to a larger population (external validity).

However, some of the described results of the data collections within the Comenius project 2.1 seemed to reveal some hints about the STSQ’s capability to differentiate between groups. I will discuss the result in the same order as described in the previous result chapter.

Table 57 shows the clearest measuring results of the STSQ within the first intervention in 2005. Figure 41 illustrates measured changes of part I (STS). The respondents of the intervention group scored significantly lower after the intervention activity. The previous results of stability over time indicate that the eSTS pattern score showed a significant stability over time regarding an interval of four weeks. The post measurement result was lower than the pre measurement results and in addition nearly the same level as the CG result. The lower measure at

the post measurement seem to reflect that the arrival and possible excitement of the first day led to the higher STS scores in comparison to the parallel CG. The measured STS score after a week of intercultural classes and theory based sport activities may have balanced out to a “normal” level in comparison to the control group. After ca. six months, the STSQ measured a higher STS score which was nearly the same as the pre measurement and distinguishes significantly from the CG score. The differences between the first and the follow up measurement of the emotional STS pattern were not statistically significant. But the results may indicate that the intervention activity show some positive long term effects with respect to an increase of eSTS pattern score after six month<sup>64</sup>. The intervention could at the beginning also lead to some confusion because the intercultural endeavors focus inter alia towards critical self-reflection which could be perceived as confusing in the first place. After some further considerations, this confusion might develop into some deeper insights of intercultural understandings which could be indicated by a higher STS.

LOC difference between the first and the follow up measurement supported this latter interpretation because the measured LOC score within the IG was lower at the follow up measurement. A decrease of LOC and increase of STS was a consistent tendency according to theoretical assumptions. However, interpretations between the first and the follow up measurement needed to be made carefully because the stability over time reference was based on a time interval of only four weeks<sup>65</sup>. The facets of part III are assumed as stable personal characteristics and consequently difficult to intervene. A lower LOC score indicated by the follow up measurement compared to the post measurement could be interpreted as a positive development of this facet of STS.

Analyses of collected qualitative data (diaries of the participants) within an external evaluation of this EU-project support the latter interpretation. A major conclusion of this external evaluation was that the participants indicated that “they found strangeness and differences less terrifying than before” (Fossgård, 2007:67). This result can to some extent be interpreted in consistence with the measured LOC difference which indicated a lower “Loss of control” score at the follow up measurement in comparison to the post measurement.

---

<sup>64</sup> However, the evaluation of the intervention activity was not the aim of this project. Nevertheless, a reasonable differentiation by STSQ scores between the different groups was useful information with respect to construct validity (Crocker & Algina, 1986; Lienert and Raatz 1998:226ff.).

<sup>65</sup> Considering the fact that the STSQ is still in development, the results were interpreted carefully in general.



The measuring results with part one of the STSQ (emotional STS) demonstrated that this emotional facet of STS appeared to be most sensible in measuring group differences. I tried to interpret the results based on the so far gained knowledge of STS and the given design of the intervention. The eSTS pattern score (emotion STS), seemed to react to the intervention. The RA 2 score measured mean score differences between the pre and the follow up measurement. But it is difficult to relate this difference to the intercultural event because the time interval between the measurements is too large for relating the measuring results to each other with some certainty.

The correlations of the RA2 score between all first measurement correlated significantly with the second and follow up measurement (cf. Table 60). Assuming that these two measurement results could be related to each other, the higher RA2 score at the second measurement would indicate that the respondents showed higher awareness over attributions and reasoning styles. Consequently, this result would indicate an increase of a STS.

The results of correlation analyzes showed that the xenophobia pattern score, RA 1 and LOC mean score correlated significantly with each other which seemed reasonable according to the theoretical assumptions of STS. The xenophobia pattern was assumed to indicate a negative sympathy assignment towards perceived difference. RA1 was refereeing to more stigmatizing attributions and reasoning styles. The correlation with the xenophobia pattern score supported the idea that the RA1 items measured personal attribution styles indirectly and not only an awareness of hypothetically existing attributions and reasoning styles. The positive correlation with LOC mean score indicated further that both RA1 and xenophobia pattern score can be related towards perceptions of uncertainty or LOC.

The NFS mean score showed a negative correlation but not significantly with one item of the additional questionnaire iii.4 "*How do you estimate the probability of success of endeavours to find peaceful solutions to conflicts between social groups?*". This indicated relationship (no statistically significant) seemed reasonable. A sub-item analysis showed that the NFS4.1 item (*When engaging in an activity, I prefer receiving clear instructions to open suggestions*) correlated highest with this item of the additional questionnaire (item iii.4;  $\rho=0,77$ ;  $p<0,01$ ). This result can be interpreted as support for NFS facet's intended measuring validity because the uncertainty indicated by NFS score was expected to counteract the optimistic attitude that there is a high chance of success of endeavors to find peaceful solutions to conflicts between social groups.

The stability analysis of all three measurements supported the theoretical expectations. eSTS, RA1, and NFS mean scores showed some instability. But these instabilities were consistent with the measured group differences which were discussed previously. All other scores showed satisfying stability results with respect to the given research design.

The significant correlations shown in Table 65 indicated some validity hints. O score/O+LOC<sup>66</sup> correlated positive with the eSTS pattern score, and negative with the LOC score. This reasonable relationship can be interpreted as a supportive hint of construct validity.

The comparisons between different STSQ scores helped to reveal some further validity hints. The significant correlation between RA 2 score and O score supported the indirect measurement of a more rational reasoning style indicated by the RA 2 items. In addition, the correlation of RA1 with NFS score showed at first that the RA1 score and RA2 score measured different facets because each score correlated with different STSQ score. The NFS items are supposed to indicate a need for security. Consequently, a positive correlation between RA1 and NFS mean scores supported a reasonable relationship between more rigid or stereotyping argumentation style (RA1) and inflexibility towards changes (indicated by NFS items).

The results of the STSQ application within the second **intervention in 2006** helped to gain further hints regarding the STSQ's capability to distinguish between groups. The comparison of measured mean-score differences was considered according to reasonability towards the theoretical conceptualization (cf. Lienert & Raatz, 1998:226-228). The last version of STSQ was applied in 2006. The group comparisons were therefore conducted more differentiated compared to the early STSQ application in 2005.

*STSQ, part I (emotional STS):*

Table 84 (appendix 9.7) showed a central trend of the *difference-scale* results indicated by mean scores. The results showed a decrease of mean scores between pre, post and follow up measurement. This decrease indicated that the respondents have perceived the shown pictures more familiar from measurement to measurement. This indicated possible remembering effects.

In order to reduce the risk of remembering effects a larger picture item pool should be developed following a similar pattern as the last item (I.16\*) i.e. ambivalent pictures showing persons or situations which are difficult to classify but at the same time provoke meaningful

---

<sup>66</sup> Whereby the LOC-items are inverted/recoded in relation to the O-items as explained earlier.

associations or a subjective meaningfulness. When intending to apply the STSQ part I in a repeated measurement design with short time intervals, a different but parallel pool of pictures for each different measurement should be selected to avoid remembering effects. It is likewise reasonable to mix new items with already used picture items in the repeated measurement as a compromise. Changing the item order may also help to reduce the risk of remembering effects.

Considering the possibility that results were influenced by remembering effects this bias may have led to a reduction of the number of measureable eSTS pattern scores. When a picture was indicated as familiar, it was not possible to measure eSTS patterns. In relation to a pre and post measurement design such a bias might have made it even harder to interpret significant group differences as a supportive indicator for an intervention activity. The remembering effects might therefore have led to underestimations of possible effects within the intervention group.

However, some of the mean score differences between pre, post, and follow up measurement indicated a trend which seemed to be relatable to the intervention activity. As illustrated in Table 84 and Figure 42, the eSTS pattern mean score of post measurement was slightly higher (not significantly) than the pre measurement (of the whole IG sample), and the follow up measurement was significantly lower compared to the post measurement and CG.

A closer look at results of the German sample pointed out this tendency. This sample showed a significant increase of STS mean score between pre and post measurement (cf. Table 84 and Figure 42). Table 84 indicated a similar but not significant trend within the other (country specific) samples. In comparison to the CG results, this difference could carefully be interpreted as a positive effect of the intervention activity regarding an increase of eSTS at least within the German sample.

The commonly measured lower eSTS mean score on the follow up measurement indicated that the intervention may have had an effect on the results. But after a few months the measured mean score was back to almost the same or even lower value as measured at the pre measurement and the measured CG STS score<sup>67</sup>.

---

<sup>67</sup> If the eSTS at the follow up measurement is significantly lower than the pre measurement and the CG score, might this indicate the possibility of remembering effects as well.

*STSQ part II and III:*

Table 57 showed a significant difference of the RA2 measurements between pre and follow up measurement. This result was relatable to the activity because the control group and pre measurement were at the same level and both were significantly lower than the follow up results. This result seemed therefore to indicate a long term effect within the intervention group. This trend appeared mainly in the Polish sample (cf. Figure 43). The German results indicated an increase between pre and post measurement as well.

The significant decrease of the NFS score between pre and post measurement of the whole sample could be interpreted as increase of a STS according to the conceptualization. Regarding this difference, no significant differences were measured compared to the CG. It is consequently more uncertain if the mentioned effect was relatable to the intervention activity. Figure 43 pointed out that the Polish sample may have reduced the more common trend of the NFS score reduction between pre and post measurement.

Table 66 showed that the STSQ counted the highest number of statistically significant differences between the measurements in the Polish sample. Figure 43 illustrated additionally that the STSQ reacts most sensibly in the sample from Poland. Besides the LOC score, all other score results showed an increase between pre and follow up measurement; the LOC score indicated a decrease. Except the mentioned unexpected increase of NFS scores, the other significant changes between the measurements were reasonable and indicated a positive influence of the activity. Even though this result appeared reasonable, it was interpreted carefully because the sample from Poland lacked a CG sample. A comparison to the existing CG data from CZ and FRA showed that the RA1 and NFS CG scores were similar to the pre measurement scores of the Polish results. All other STSQ scores of the CG were lower than the Polish results of the pre measurement. This similarity between CG and the Polish IG made it more certain to relate only the results of RA1 and NFS score of the Polish sample to the intervention.

However, the decrease of the LOC score of the Polish sample between pre and follow up measurement seemed reasonable. If this result could be related to the activity, it could be interpreted as a positive indicator for an improvement of this facet of STS. It would be reasonable to expect that the NFS score showed a similar trend because LOC and NFS were both related to uncertainty. On the basis of the theoretical assumptions an improvement of STS implied a lesser need for security. All other scores measured in the Polish sample changed between the measurements in a consistent way regarding an improvement of STS (except the discussed NFS scores).

---

One could argue that measured group differences can either be related to the intervention activity or errors/instabilities of the measuring instrument. But previous results showed that the measurements of the STS facets specifically RA1, O, LOC and NFS were quite stable over a period of four weeks (chapter 6.7). In addition, the STSQ showed stable results in the repeated measurement design (including three measurements and one control group) of the EU-project in 2005. Therefore, the results could carefully be interpreted as an improvement of STS of the Polish sample.

The last significant result of the Polish sample was a measured increase of O+LOC score between pre and follow up measurement. Pre and post measurement O+LOC scores showed an increase as well but were not statistically significant (cf. Table 66). The increase between first and follow up measurement was twice as much as between the pre and post measurement. The pre measurement result of the O+LOC score (2,61) was also lower than the existing CG value (3,05), but the follow up measurement showed a similar result as the CG value (of the CZ and FRA sample in summary).

The result of the intervention design showed that it was reasonable to analyze measuring results country specific. Figure 43 showed quite clear that analyses of the whole sample at once can cover measuring effects; probably because of the heterogeneously structured sample regarding country of origin. Such an insight suggests conducting country specific validity studies when intending to apply the STSQ in different countries or interpreting the data on the country specific background. Varying competence in English may also bias the measurements within different countries.

In addition to the Polish sample, the German sample showed the most reasonable results considering the construct. This seemed logical probably because the underlying concept of IME was founded by the German research group.



## 7. Discussion and conclusions

The original aim of my thesis was an operationalization of strangeness in the context of IME. The intended concept-based procedure was the development of the STSQ that started with theoretical pre-considerations and resulted in a conceptualization of sensitivity towards strangeness. The concept-based or theory-driven procedure was an important idea for the development and application of the STSQ. The close relationship to the theoretical concept was expected to make the STSQ easier to apply because the theory-driven procedure did not call for high psychometric norms and standards (cf. Erdmann, 1988).

However, the development of a concept-based measuring device was a complex and challenging task which implied certain risks. The major challenge was to handle and consider numerous complexities on different levels at each stage of the process such as theoretical development, design of the instrument, design of the empirical steps, concrete item formulation, statistical procedures etc. All these levels are influenced by potential “errors”. Errors within the theoretical assumptions can be transferred to subsequent stages of development. But difficulties during data collection and sampling procedures made it difficult to collect evidence to prove the measuring instrument’s intended capability to differentiate between groups in a theoretically intended manner.

The aim of my thesis required a close relationship between IME and STS/STSQ. A look through the literature showed that there were numerous theoretical perspectives which could be related directly or indirectly to the phenomenon of strangeness. As indicated in my literature review, most of these related concepts examined the phenomenon of strangeness from different theoretical perspectives and were more or less comprehensive. The essential ideas of these related theoretical approaches also showed that they are relatable in a coherent way to the outlines of IME and my conceptualization of STS. But the background and intended application in the context of movement education made IME and STS different to other related approaches.

The starting point for my operationalization of strangeness was IME. Further considerations around different concepts of strangeness, identity theory, Piaget’s concepts of assimilation and accommodation and classical attribution theory influenced my conceptualization of sensitivity towards strangeness. The different theoretical perspectives showed that strangeness is related to individual experiences which can make perceptions of strangeness quite subjective. But the selected theoretical approaches also showed that we as human beings we are not

that different so that mechanisms and structures leading to perceptions and experiences of strangeness show some general or coherent characteristics. Being aware of the potential diversity and subjectivity of what could be perceived as different and strange, individuals with a similar social and/or cultural background learn to perceive and explain their world around them in more or less comparable ways. Otherwise common sense and communication would not be possible. The focus of my investigations was therefore to point out assumed general or inter-subjective characteristics of strangeness. Developing a measuring instrument required focusing on these inter-subjective facets of strangeness; otherwise it might not be possible to measure such a theoretical concept.

Since I tried to develop a new concept and measuring device, my investigations were dominated by its explorative character. The indicated structure of my operational model shown in Figure 18 may suggest applying Confirmatory Factor Analysis (CFA) or Structural Equation Modelling (SEM) for investigating construct validity of the STSQ. But these complex and advanced procedures are usually applied for testing more established constructs and measuring instruments (Bühner, 2004). My operational model is supposed to structure the item pool in a theoretically reasonable way. Collecting further empirical knowledge about STS and STSQ, SEM may become an important methodological approach for investigating construct validity of the STSQ in the future.

It was challenging to find and formulate items (indicators) which were easy to understand for potential examinees and at the same time represent measurement for complex ideas of the underlying theoretical considerations. Formulating a pool of items which consisted of too obvious and simple formulated items was considered both provoking the risk of measuring trivial facets and the indicated potential risk of errors such as political correctness or social desirability. Following the usual procedure of developing an item pool with a number of parallel items would probably lead to high psychometric measuring standards. But this was not the aim in the context of measuring sensitivity towards strangeness. Regarding the topic of this thesis and the presented theoretical background the intention was to develop a theory-based heterogeneously structured item pool which represents a compromise between sufficient psychometric standards and the representation of complex ideas of the underlying concept. The aim was to measure in band-width and measure group differences and not individual differences.

In order to handle the different complexities related to the aim of my thesis, I needed to demarcate the aims of my project. My focus was therefore the development of a first item pool (STSQ), and the results of my research were understood as a start to establish STS and



---

STSQ. Both, the theoretical concept of STS and the resulting questionnaire (STSQ) are new and require further development and more systematic and advanced testing in the future in particular if it is intended to apply the STSQ in other contexts than IME and STS.

My methodological agenda was to develop the STSQ successively starting with a more or less deductively derived item pool and with help of a number of small-sample data collections. The successive adaptations of the STSQ were supposed to be the thread through my empirical investigations. I also included additional analyses in order to get an overview on certain “classical” statistical references such as Cronbach’s alpha, explorative factor analysis even though an application of these procedures were problematical regarding the needed mathematical prerequisites and sampling procedures.

The dominantly deductive (concept-based) procedure at the beginning was a crucial step in the development. But the development of a new theoretical concept required explorative procedures. Consequently, the operational model and item pool construction has to be understood as a product of an inter-play between theoretical considerations and empirical findings. To follow an initially intended strict deductive procedure was impossible. On the other hand, the theoretical references were needed as guidelines for the development as interpretations of the data otherwise would be impossible and meaningless. It is natural that the interpretations of the STSQ results are more uncertain compared to established measuring instruments. But I tried to reduce this uncertainty by careful theoretical groundwork.

By empirical steps of development I tried to discover hints on validity basically in order to formulate and refine single items and develop summarizing item scores. Most of my empirical investigations were consequently related to the concept of construct validity. But I preferred the term “validity hints” to point out that I basically tried to adapt the more deductively derived items in the theoretically intended manner (IME and STS).

These aims required a lesser claim to sampling procedures so data analyses were in the first place based on more or less convenient samples. The demographic items of the STSQ helped to characterize the sample more specifically in order to gain validity hints by subgroup comparisons. More systematic sample selection procedures would probably have led to clearer results. But the smaller sample sizes and simpler statistical procedures should reduce the risk of inadmissible inferences and randomly discovered significant results as argued in chapter 6.2. The application of non-parametric statistical procedures may have led to underestimations of the results. But an underestimation was not considered as problematical. It was a part of my strategy to avoid risks of self-serving or confirmation bias and circular reasoning.

Applications of dominant non-parametric statistical procedures were more conservative testing in order to reduce the risks of the aforementioned error influences.

My empirical results are the starting point for further development and should provide an initial reference when intending to apply the STSQ in other contexts e.g. stability results and correlations between the different scores of the STSQ with different samples may serve as initial statistical reference for applications of the STSQ. Since it was difficult to estimate reliability and validity of the STSQ scores according to classical test theory it is so far advisable to apply the STSQ on the basis of my presented theoretical background. As I pointed out previously, the lower scale qualities and psychometric standard of the STSQ require a specific theoretical reference and careful considerations of measuring results. Interpretations can only be made on the basis of the respective theoretical baselines of IME, STS and STSQ. An interesting question in the future would be how far the STSQ is applicable in related but different theoretical contexts.

The first application of the STSQ within the theoretically related EU-project showed that the STSQ was capable of measuring group differences. How far the intervention was responsible for measured differences was quite uncertain. The EU-project had not an optimal intervention design. A six day's training week appeared to be too short and unspecific regarding STS and measuring clear effects of the intervention with the STSQ. My personal teaching experiences of STS in the context of teacher education supported my impression that a mediation of STS requires longer, more intensive and more specific intervention periods besides a more controlled intervention and measurement design in order to be in better control of error influences.

The STSQ was structured based on two theoretical dimensions EM and CM. Two different measuring approaches were applied to measure the cognitive dimensions: a semi-projective procedure and a statement item pool. A behavioural dimension was not considered for the STSQ. At first, the underlying concept of subjective meaningfulness (Haußer, 1996; 2007) suggested an emotional and cognitive dimension. Secondly, the two dimensions already led to a quite complex structure of the STSQ. A further "big" dimension would increase this complexity even more. Thirdly, I followed an assumption that subjective meaningfulness and attitudes are assumed to influence behaviour in a significant way. In this sense, behavioural facets were included but in an indirect way. It can be discussed in how far subjective meaningfulness and attitudes (which imply a subjective meaningfulness) are coherent with resulting behaviour. Situational influences can probably lead to unexpected and inconsistent reactions

but attitudes are assumed to influence our behaviour in more or less coherent way (Gerrig & Zimbardo, 2002:550ff.).

My empirical investigations of part I of the STSQ (eSTS) revealed two main challenges regarding the picture selection: 1) what kind of pictures were appropriate in order to provoke a perception of difference, and 2) an appropriate way to analyse collected raw data by relating the two scales in a theoretical reasonable way (pattern score construction). The picture selection was characterized by a kind of dilemma. On the one hand, selected pictures needed a certain degree of meaningfulness to the examinees otherwise a possible perceived difference is irrelevant to the individual and a subsequent evaluation of sympathy as well. On the other hand, the pictures should be perceived as different to the individual's own perception. Optimal would be that the picture item provoked a meaningful distinction between the perceived picture and the individual's own perception. Therefore, pictures need to be selected in relation to the respective population's background. Using pictures from a relatively unknown culture (e.g. from a remote cultural group in the Amazon forest) would probably incite more "strangeness" from the perspective of mainstream population from a western industrialised country such as Germany or Norway. But these natives were probably so far away or distant (according Simmel's social dimension) that most of the examinees might evaluate the pictures frivolous because they had no significant relationship to the shown picture. Increasing the number of picture items according the indicated criteria should increase the chance that more items were perceived as different. In addition, the selection of pictures needed to be guided by considerations of the individual, social or cultural context of the respective goal population in order to insure a certain subjective meaningfulness. My analyses of the item pool showed that some of the items were too provocative as item stimulus and consequently showed problematical (skew score) distributions on the sympathy scale. Therefore these items were considered as unsuitable in order to differentiate between groups. For instance, the initial picture item I.5 (two men kissing) was more an indicator for homophobic tendencies than an indicator for STS. In order to increase a more balanced distribution on the sympathy scale, it might be replaced with a more ambiguous picture such as shown in Figure 44.



**Figure 44: More ambiguous item in comparison to the initial item I.5**

A later application of a Norwegian translation of my STSQ<sup>68</sup> by Per Midthaugen used the picture item shown in Figure 44. The results indicated a more balanced distribution of the sympathy scale which also means that the new picture probably did not appear as obvious as the initial picture I.5.

Item I.16 showed an optimal prerequisite for measuring a STS in the theoretically intended way. This item could therefore serve as guideline for a potential extension of this item pool as well.

The other difficulty was related to the transformation of the measured raw scores of the two different scales to one summarizing eSTS-score. The suggested pattern score indicating eSTS (cf. Figure 30) could only function as indicator when the respective picture was perceived as different. An alternative pattern needed to be considered as well. My results indicated that the xenophobia and exoticism pattern scores correlated in a reasonable way. Instead of calculating three different scores I would like to suggest the following procedure by combining the three scores to one summarizing STS score: Each identified response pattern should be coded with a number from 1 to 3 as shown in Table 67. In order to point out that eSTS refers to a more neutral evaluation of strangeness the raw scores are transformed to a scale between -1 and +1. Measuring results around 0 would then indicate eSTS. Results close to +1 or -1 respectively would indicate more exoticism responses or xenophobic responses respectively. Such a scale appears theoretically reasonable. But it needs to be tested how far this theoretical assumption may work empirically. This type of eSTS score would imply more information than my initially constructed eSTS response pattern score which did not include the alternative responses.

---

<sup>68</sup> Per Midthaugen at the Norwegian School of Sport Sciences translated and applied the STSQ in the context of his PhD project with the working title: Handling differences and intercultural learning in Physical Education. Measuring outcomes among students after a PE-teacher training intervention in upper secondary schools in Norway.

**Table 67: Suggested eSTS pattern score**

<b>Pattern</b>	<b>Xenophobia</b>	<b>STS</b>	<b>Exoticism</b>
Coding observed pattern	1	2	3
<i>transformation</i>	-2	-2	-2
<b>Transformed eSTS score</b>	<b>-1</b>	<b>0</b>	<b>+1</b>

Part II of the STSQ (situation item pool) was supposed to measure awareness of different reasoning and argumentation types in relation to described situation. It was initially expected to distinguish between three levels of rationality (RA1-3). But the empirical results did not support the theoretical idea as differentiated as intended. My empirical results pointed out the RA1 scores which seemed to differentiate between groups in a consistent way. The negative correlations with eSTS pattern scores and openness scores indicated that RA1 seemed to measure kind of stigmatizing reasoning in opposite to STS.

To some extent my analyses supported a differentiation between two RA levels: RA1 and some of the RA3 items. Developing further items according the intended meaning of RA3 items may lead to clearer results. In particular, situation three (Yildiz) appeared to be a potentially fruitful item pool for differentiating between RA1 and RA3 level. It was not possible to support this idea clearly with empirical evidence but a rational re-consideration pointed out this item specifically. The RA 3 items of situation 3 are theoretically corresponding with the meaning of STS. If it would be possible to extend the item pool of each RA level of situation 3, a differentiation between RA1 and RA3 might possibly become clearer. Situation one on the other hand, appeared to be a better indicator for RA1 items. Based on my overall consideration of part II, I would suggest focussing on situation one and three in the future and extend the item pool of these situations in the mentioned way. Situation 2 is more in line with situation 1 an indicator for RA1. Situation 4 is more similar to situation 3 and should therefore represent a better indicator for RA3 level.

RA scores were supposed to measure an awareness of hypothetically existing reasoning and argumentation types. Some results indicate a negative relationship between RA1 and RA2 scores. A negative correlation between RA1 and RA2 may indicate that the differentiation reflects indirectly the respondent's personal opinion. Then, the second part of the STSQ differentiates specifically between the assumed two rationality levels of the RA-facet.

The items of part three of the STSQ were also heterogeneously structured and it was therefore difficult to estimate reliability with Cronbach's alpha. The results of test-re-test reliability indicated a moderate stability over time for LOC and NFS score. The heterogeneity of the

items may have led to underestimate the reliability with coefficient alpha (cf. Osburn, 2000). In addition to the heterogeneity, the number of items was also relative low. Both factors reduce Cronbach's alpha results (Weller & Matiaske, 2003). Increasing the number of items of each facet would probably lead to higher Cronbach's alpha values. Correlations indicated a negative relationship between O and LOC score. This result suggested combining the LOC and O score to one summarizing O+LOC score. The result already led to higher Cronbach's alpha values and better stability over a four weeks time interval. Subsequent analysis supported the idea to summarize the O and LOC items.

Correlations between NFS and LOC indicated a reasonable relationship between these two facets. Both facets are related to the concept of uncertainty. The indicated relationship to the openness facet and uncertainty related facets may suggest developing this item pool further with the purpose to develop an openness or uncertainty scale respectively. However, the correlations between the scale items were not always consistent when comparing correlations with different samples. But it is also difficult to achieve consistent results with different samples (cf. Bühner, 2004).

The results of my thesis are only the beginning of developing and establishing my concept of STS and its measurements. My empirical analyses showed that the intended heterogeneity of the STSQ created difficulties regarding reliability and validity according to classical test theory. However, the multi-faceted character of STS and the resulting heterogeneity of the STSQ items were important characteristics of sensitivity. Sensitivity requires awareness and a better understanding about strangeness. An awareness of the complexity related to strangeness will promote openness. Openness requires a willingness to listen and trying to understand. Both the awareness and openness are needed competencies for a STS. Training a STS would focus on developing these two aspects.

Dealing with strangeness is related to identity conceptions. Identity work is a crucial element in the development and improvement of STS. Experiencing acknowledgement and inclusion are stabilizing elements in identity development. This creates a prerequisite for being open-minded towards strangeness because openness towards strangeness implies taking a risk. We do not know the outcome of dealing with strangeness. Promoting a better understanding of difference and strangeness in the context of intercultural learning may help to tolerate the implicit uncertainty better and enable to learn strategies of how to deal constructively with difference and strangeness.







---

## 8. Further research

The STSQ is based on a new concept which requires further critical and specialized investigation and development. My considerations around strangeness pointed out the concept of *uncertainty* as a central factor within identity conceptions. The ways people deal with strangeness are influenced by our identity conception and how we learned to deal with uncertainty. It would be interesting to develop further the concept of uncertainty in relation to strangeness, identity and to investigate more specifically the conditions for dealing constructively with strangeness. A further development of STSQ scales would suggest collecting data with larger sample sizes in order to enable more advanced analyzing techniques.

My application of STSQ within the EU-project indicated that an intervention of STS requires longer and more specified intervention periods. The assumed supportive effects of sport and movement education need to be investigated systematically. Therefore, my research plans in the future are directed at developing an intervention in a quasi-experimental design among teacher students. This aim implies developing teaching material for improving a STS and investigates more systematically how far movement education can promote a STS. In this context, the STSQ will be an instrument for evaluating potential effects of my intervention. Additional qualitative procedures may help to evaluate the intervention and STSQ's measuring capability. A more controlled and specified intervention design will also show how far the STSQ is able to measure a STS. Results are consequently expected to generate more explicit knowledge about the empirical representation of STS and STSQ's capability of measuring this concept and help developing my concept and measuring instrument further.



## 9. Appendix

### 9.1. Descriptive item/scale statistics of the stability data collection

**Table 69: Descriptive item statistics of part I, STSQ, sample of measurement 1, N=69**

Scale	Item	N	Mean	SD	Skew-ness	Std. Error	Item total correlation	Cronbach's Alpha if Item Deleted
<b>Familiarity</b>	<b>I.1</b>	<b>69</b>	<b>2,19</b>	<b>1,18</b>	<b>0,73</b>	<b>0,29</b>	0,39	0,78
	<b>I.2</b>	<b>69</b>	<b>2,28</b>	<b>1,16</b>	<b>0,83</b>	<b>0,29</b>	0,52	0,76
	<b>I.3</b>	<b>69</b>	<b>1,97</b>	<b>1,00</b>	<b>0,79</b>	<b>0,29</b>	0,52	0,77
	I.4	69	3,22	1,24	-0,38	0,29	0,32	0,79
	<b>I.5</b>	<b>69</b>	<b>1,93</b>	<b>1,25</b>	<b>1,16</b>	<b>0,29</b>	0,50	0,77
	<b>I.6</b>	<b>69</b>	<b>1,68</b>	<b>0,98</b>	<b>1,56</b>	<b>0,29</b>	0,67	0,75
	I.7	69	3,58	1,27	-0,50	0,29	0,43	0,77
	I.8	69	<b>4,03</b>	1,14	-1,05	0,29	0,39	0,78
	I.9	67	3,04	1,30	0,00	0,29	0,45	0,77
	I.10	69	2,61	1,32	0,13	0,29	0,29	0,79
	I.11	68	2,84	1,23	0,02	0,29	0,50	0,77
<b>Sympathy</b>	I.1	69	3,17	1,15	-0,17	0,29	0,50	0,72
	<b>I.2</b>	<b>69</b>	<b>4,07</b>	<b>0,81</b>	<b>-0,48</b>	<b>0,29</b>	0,34	0,74
	I.3	69	2,65	1,07	0,07	0,29	0,15	0,76
	I.4	69	2,86	1,24	0,05	0,29	0,53	0,72
	I.5	69	2,54	1,30	0,43	0,29	0,57	0,71
	<b>I.6</b>	<b>69</b>	<b>3,30</b>	<b>1,31</b>	<b>-0,27</b>	<b>0,29</b>	0,14	0,77
	I.7	69	2,75	1,17	0,27	0,29	0,46	0,73
	<b>I.8</b>	<b>69</b>	<b>2,83</b>	<b>1,43</b>	<b>0,04</b>	<b>0,29</b>	0,46	0,73
	I.9	66	3,33	1,29	-0,35	0,30	0,55	0,71
	I.10	69	2,97	1,24	-0,09	0,29	0,31	0,75
	I.11	68	3,07	1,27	-0,01	0,29	0,41	0,73
	Valid N (listwise)	65						

## 9.2. Mean score results of the first data collection in 2005 within the Comenius 2.1-project

**Table 70: The mean score results (Pre, post, follow up and control group [CG ] measurement)**

STSQ score	Pre			Post			Follow up			CG		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
eSTS pattern	32	,26	,34	29	,13	,19	11	,26	,13	49	,14	,19
Xenophobia pattern	32	,09	,15	29	,09	,15	11	,24	,2	49	,11	,22
Exoticism pattern	32	,29	,31	28	,29	,31	11	,07	,14	49	,31	,32
Pity pattern	32	,02	,08	29	,04	,10	11	,04	,08	49	,03	,10
RA 1	32	2,53	,31	30	2,54	,42	10	2,76	,20	47	2,57	,33
RA 2	33	2,8	,3	30	2,8	,4	11	2,84	,3	47	2,79	,24
O	32	3,21	,31	28	3,09	,38	10	3,18	,22	48	3,14	,37
NFS	32	2,95	,28	30	2,91	,3	11	2,84	,27	47	2,85	,34
LOC	32	2,32	,27	30	2,40	,29	11	2,06	,32	49	2,25	,39
O + LOC	30	3,03	,21	28	2,93	,25	10	3,05	,21	48	3,01	,28

**Table 71: The German mean score results (Pre, post, follow up and control group [CG ] measurement)**

STSQ score	Pre			Post			Follow up			CG		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
eSTS pattern	10	0,33	0,35	10	0,20	0,23	--	--	--	22	0,15	0,17
Xenophobia pattern	10	0,10	0,16	10	0,13	0,17	--	--	--	22	0,05	0,12
Exoticism pattern	10	0,13	0,17	10	0,27	0,26	--	--	--	22	0,36	0,32
Pity pattern	10	0,03	0,11	10	0,03	0,11	--	--	--	22	0,08	0,14
RA 1	10	2,38	0,26	11	2,32	0,32	--	--	--	22	2,60	0,43
RA 2	11	2,95	0,26	11	2,92	0,34	--	--	--	22	2,94	0,26
O	11	3,22	0,30	10	3,32	0,25	--	--	--	22	3,18	0,25
NFS	10	2,90	0,27	11	2,80	0,27	--	--	--	20	2,83	0,28
LOC	10	2,28	0,24	11	2,36	0,30	--	--	--	22	2,19	0,40
O + LOC	10	2,99	0,15	10	3,07	0,14	--	--	--	22	3,03	0,25

**Table 72: The French mean score results (Pre, post, follow up and control group [CG ] measurement)**

STSQ score	Pre			Post			Follow up			CG		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
eSTS pattern	7	0,33	0,27	6	0,17	0,18	4	0,20	0,16	10	0,43	0,27
Xenophobia pattern	7	0,19	0,18	6	0,06	0,14	4	0,30	0,26	10	0,13	0,17
Exoticism pattern	7	0,29	0,23	6	0,22	0,34	4	0,10	0,20	10	0,20	0,23
Pity pattern	7	0,05	0,13	6	0,06	0,14	4	0,05	0,10	10	0,03	0,11
RA 1	7	2,82	0,25	6	2,82	0,17	4	2,86	0,27	10	2,74	0,24
RA 2	7	2,70	0,27	6	2,69	0,37	4	2,73	0,30	10	2,57	0,31
O	6	3,30	0,28	6	3,07	0,41	4	3,25	0,25	9	3,20	0,26
NFS	7	2,82	0,35	6	3,04	0,29	4	2,85	0,30	10	2,88	0,29
LOC	7	2,26	0,23	6	2,28	0,31	4	1,96	0,46	10	2,28	0,19
O + LOC	6	3,14	0,10	6	3,02	0,24	4	3,14	0,28	9	3,12	0,08

**Table 73: The Czech Republic mean score results (Pre, post, follow up and control group [CG ] measurement)**

STSQ score	Pre			Post			Follow up			CG		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
eSTS pattern	5	0,13	0,18	5	0,07	0,15	5	0,32	0,11	6	0,17	0,18
Xenophobia pattern	5	0,07	0,15	5	0,13	0,18	5	0,16	0,17	6	0,06	0,14
Exoticism pattern	5	0,27	0,15	5	0,13	0,18	5	0,04	0,09	6	0,22	0,34
Pity pattern	5	0,00	0,00	5	0,07	0,15	5	0,04	0,09	6	0,06	0,14
RA 1	5	2,49	0,08	5	2,64	0,26	4	2,70	0,14	6	2,82	0,17
RA 2	5	2,87	0,35	5	2,88	0,30	5	2,98	0,33	6	2,69	0,37
O	5	3,08	0,27	5	3,00	0,24	5	3,08	0,18	6	3,07	0,41
NFS	5	2,90	0,22	5	2,95	0,21	5	2,92	0,23	6	3,04	0,29
LOC	5	2,23	0,32	5	2,60	0,28	5	2,10	0,25	6	2,28	0,31
O + LOC	5	2,98	0,28	5	2,76	0,29	5	2,98	0,15	6	3,02	0,24

**Table 74: The Polish mean score results (Pre, post, follow up and control group [CG ] measurement)**

STSQ score	Pre			Post			Follow up			CG		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
eSTS pattern	8	0,13	0,35	6	0,00		--	--	--	7	0,24	0,25
Xenophobia pattern	8	0,00	0,00	6	0,06	0,14	--	--	--	7	0,14	0,26
Exoticism pattern	8	0,50	0,47	5	0,53	0,38	--	--	--	7	0,29	0,30
Pity pattern	8	0,00	0,00	6	0,00	0,00	--	--	--	7	0,00	0,00
RA 1	8	2,53	0,33	6	2,58	0,74	--	--	--	6	2,68	0,21
RA 2	8	2,65	0,28	6	2,53	0,54	--	--	--	6	2,68	0,13
O	8	3,10	0,32	5	2,68	0,36	--	--	--	6	3,43	0,34
NFS	8	3,16	0,23	6	2,96	0,43	--	--	--	7	2,71	0,39
LOC	8	2,50	0,32	6	2,44	0,27	--	--	--	7	2,31	0,42
O + LOC	7	2,96	0,27	5	2,65	0,16	--	--	--	6	3,02	0,23

## 9.3. STSQ mean score results of different sample types

Table 75: STSQ scores of different sample types

Sample	N	eSTS pattern		RA1		RA2		O		NFS		LOC	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1. Sport Gymnasium (high school students)	224	0,25	0,27	1,49	0,21	1,69	0,16	3,02	0,34	2,87	0,36	2,14	0,31
2. EU project 2006 (international intervention sample, post measurement)	24	0,33	0,29	2,68	0,28	2,89	0,34	3,09	0,3	2,58	0,43	2,11	0,29
3. Master students of social work	10	0,36	0,31	2,56	0,22	2,99	0,26	3,2	0,31	2,53	0,32	2,0	0,16
4. Students of physics (4. semester)	41	0,28	0,21	2,33	0,42	2,84	0,2	3,18	0,3	2,56	0,36	1,95	0,34
5. Master students of multicultural and international studies (international sample)	10	0,2	0,16	2,42	0,53	2,89	0,33	3,28	0,25	2,66	0,4	1,96	0,34
6. Master students of information technology	22	0,33	0,21	2,59	0,38	2,83	0,36	3,19	0,24	2,52	0,32	1,92	0,35
7. International summer school	22	0,13	0,15	2,41	0,35	2,93	0,35	3,01	0,42	2,82	0,48	2,02	0,46

## 9.4. STSQ inter-score correlations of different sample types

**Table 76: STSQ inter score correlation (Spearman's rho), Sport Gymnasium (high school students), N=224**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	,07	--				
<b>RA2</b>	,08	,27**	--			
<b>O</b>	0	0,4	0	--		
<b>NFS</b>	,03	,02	,11	,13	--	
<b>LOC</b>	-,07	0	-,04	-,19**	-,11	--

\*\* . Correlation is significant at the 0.01 level (2-tailed).

**Table 77: STSQ inter score correlation (Spearman's rho), EU project 2006 (intervention sample, post measurement), N=21**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	-,4	--				
<b>RA2</b>	-,48	,45	--			
<b>O</b>	-,3	,01	,54*	--		
<b>NFS</b>	-,46*	,61**	,34	,16	--	
<b>LOC</b>	,02	,43	,4	-,13	,18	--

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

**Table 78: STSQ inter score correlation (Spearman's rho), master students of social work, N=10**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	,42	--				
<b>RA2</b>	-,46	,11	--			
<b>O</b>	-,05	,57	,29	--		
<b>NFS</b>	,08	,13	-,04	,56	--	
<b>LOC</b>	-,09	-,3	-,02	,18	-,12	--

**Table 79: STSQ inter score correlation (Spearman's rho), physic students, N=41**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	,09	--				
<b>RA2</b>	-,02	,54**	--			
<b>O</b>	-,05	-,26	,02	--		
<b>NFS</b>	,2	,03	,19	-,11	--	
<b>LOC</b>	,07	,02	-,14	-,24	,16	--

\*\* . Correlation is significant at the 0.01 level (2-tailed).



**Table 80: STSQ inter score correlation (Spearman's rho), master students of multicultural and international studies (international sample), N=10**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	,09	--				
<b>RA2</b>	,62	,32	--			
<b>O</b>	,37	,3	,2	--		
<b>NFS</b>	,12	-,22	-,35	-,17	--	
<b>LOC</b>	-,5	,03	-,69*	-,21	,65*	--

\*. Correlation is significant at the 0.05 level (2-tailed).

**Table 81: STSQ inter score correlation (Spearman's rho), master students of information technology, N=22**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	,19	--				
<b>RA2</b>	,15	,06	--			
<b>O</b>	,35	-,03	,3	--		
<b>NFS</b>	,04	-,14	-,10	-,09	--	
<b>LOC</b>	,19	,33	-,03	-,24	-,02	--

**Table 82: STSQ inter score correlation (Spearman's rho), international summer school Oslo, N=24**

<i>STSQ score</i>	<b>eSTS pattern</b>	<b>RA1</b>	<b>RA2</b>	<b>O</b>	<b>NFS</b>	<b>LOC</b>
<b>STS pattern</b>	--					
<b>RA1</b>	,41	--				
<b>RA2</b>	-,27	,15	--			
<b>O</b>	,43	,48*	,16	--		
<b>NFS</b>	-,09	,07	,02	,22	--	
<b>LOC</b>	,3	-,55*	-,25	-,37	,04	--

9.5. Difficulty index of the items of part II according to Guttman (1950)

The following figures (Figure 45 - 49) show the items of part II of the STSQ arranged according to increase of difficulty.

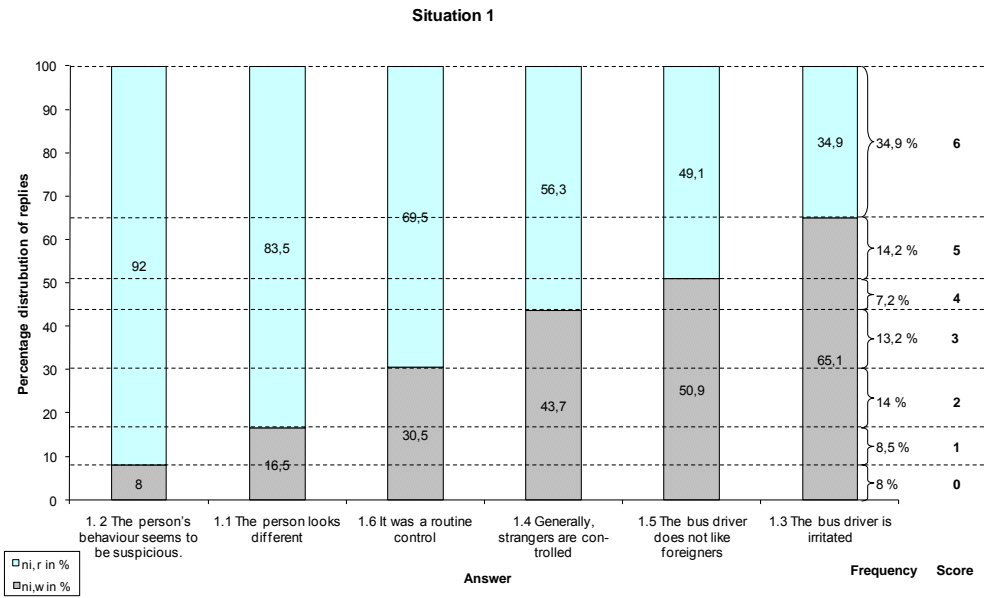


Figure 45: Item order of situation 1 according to the item's difficulty

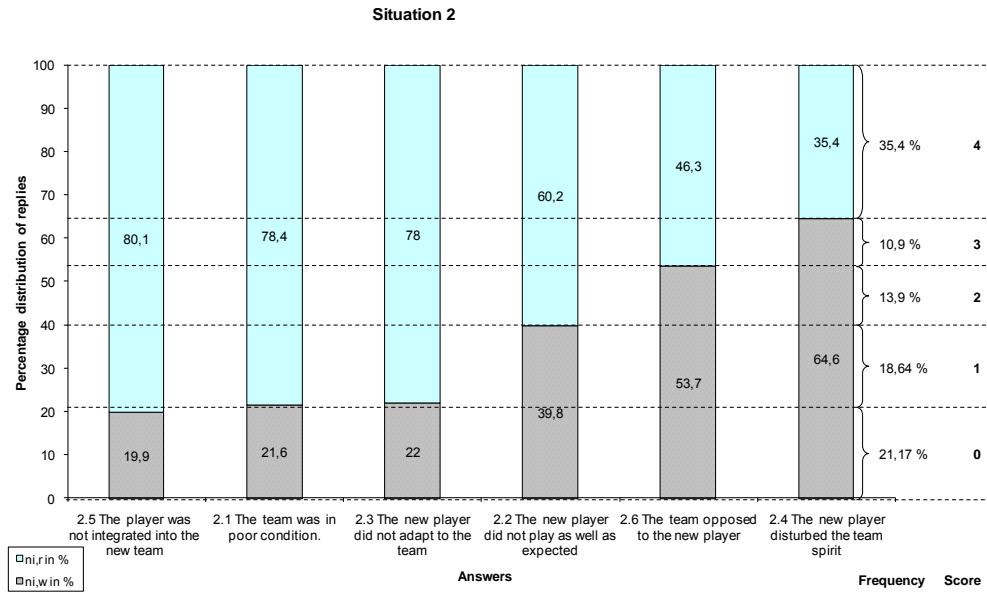


Figure 46: Item order of situation 2 according to the item's difficulty

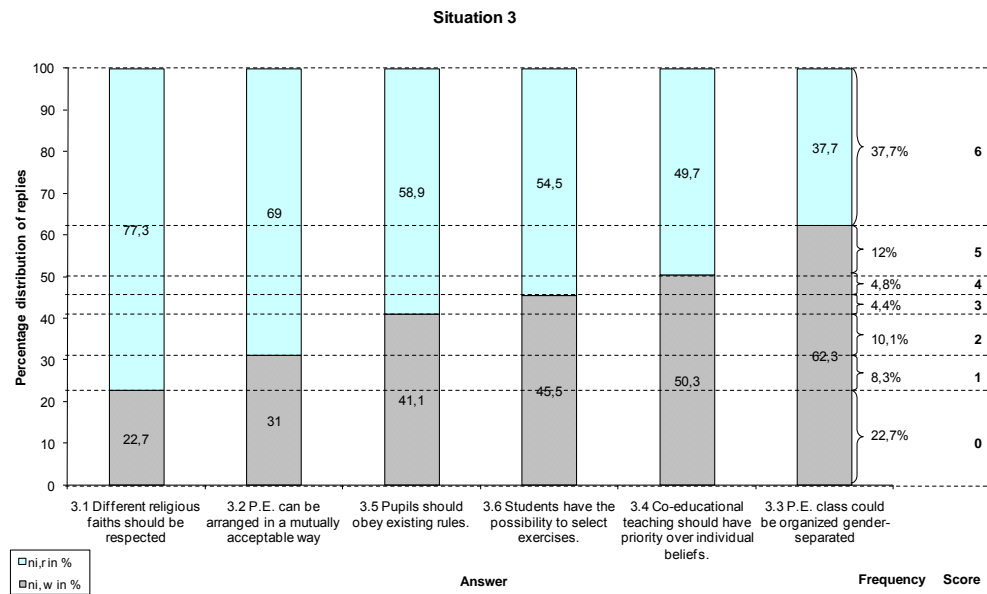


Figure 47: Item order of situation 3 according to the item's difficulty

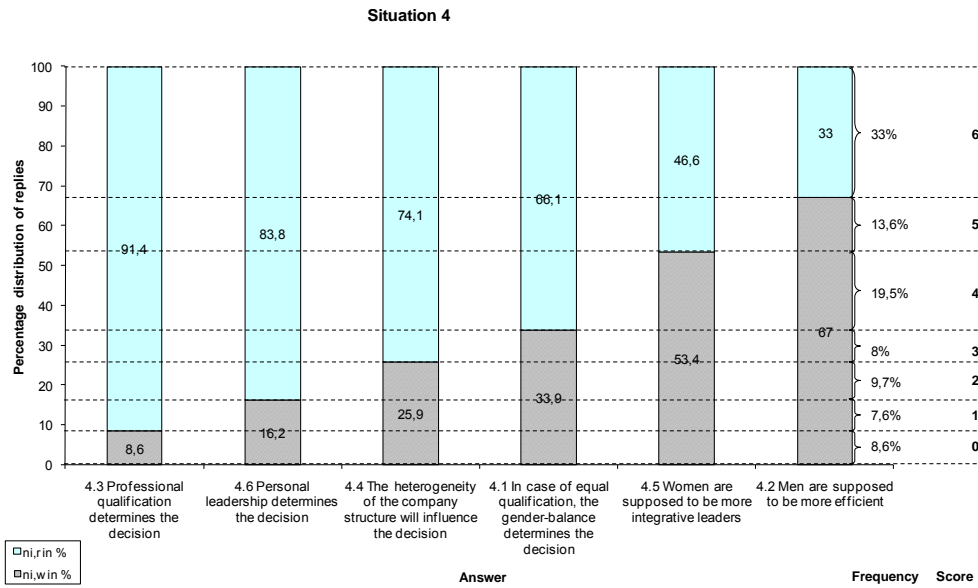


Figure 48: Item order of situation 4 according to the item's difficulty

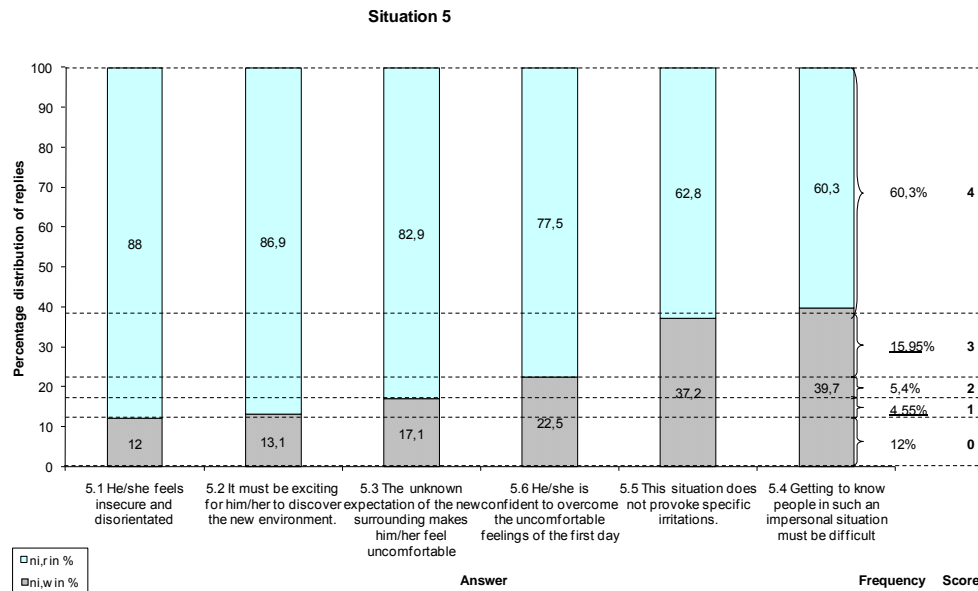


Figure 49: Item order of situation 5 according to the item's difficulty

## 9.6. Cronbach's alpha of the STSQ scores separated by high school years

**Table 83: Cronbach's alpha of the STSQ scores separated by high school years**

STSQ score	Sample size N*	N of items	Cronbach's alpha		
			1. year	2. year	3. year
eSTS pattern score	77/70/67	3	0,41	0,11	0,37
RA 1	72/70/66	12	0,55	0,64	0,27
RA 2	72/70/64	11	0,38	0,54	0,44
O	78/70/67	5	0,46	0,22	0,18
NFS	70/70/65	4	0,13	-0,1	0,04
LOC	69/66/62	6	0,27	0,25	0,06
O+LOC	68/65/61	11	0,5	0,29	0,26

\*Listwise deletion based on all variables in the procedure.

## 9.7. Mean score results of the second data collection in 2006 within the Comenius 2.1.1-project

Table 84: Mean score results for the whole sample and country specified

Group/ measurement	N	Difference mean score	Sympathy mean score	eSTS pattern mean score	RAI	RA2	O	NFS	LOC	O+LOC
Pre	23	3,80	2,91	0,29	2,67	2,77	3,07	2,71	2,12	2,95
Post	24	3,43	3,07	0,33	2,6	2,78	2,99	2,57	2,12	2,93
Follow up	21	3,32	2,89	0,18	2,75	3,01	3,17	2,84	2,1	3,03
CG	13	3,77	2,66	0,23	2,58	2,81	3,26	2,78	2,1	3,05
GER pre	6	3,83	2,64	0,24	2,42	3,02	3,27	2,63	2,14	3,05
GER post	6	3,9	2,93	0,43	2,58	3,17	3,3	2,47	2,08	3,09
GER follow up	5	3,64	2,80	0,32	2,72	3,12	3,2	2,5	2,29	2,98
CZ pre	6	3,97	3,33	0,43	2,79	3,06	3,27	2,7	1,7	3,29
CZ post1	6	3,8	2,83	0,47	2,8	3,03	3,13	2,63	1,94	3,09
CZ follow up	6	3,33	2,97	0,17	2,6	2,92	3,17	2,57	1,86	3,15
CZ CG	6	3,57	2,57	0,23	2,35	2,8	3,17	2,73	2,03	3,05
POL pre	7	3,91	2,80	0,29	2,6	2,6	2,77	2,77	2,47	2,61
POL post	7	3,31	3,14	0,34	2,68	2,76	2,83	2,71	2,33	2,76
POL follow up	6	3,2	2,63	0,03	3,17	3,33	3,3	3,37	2,14	3,06
FRA pre	5	3,4	2,84	0,16	2,68	2,6	3,04	2,76	2,47	2,76
FRA post	5	2,6	3,4	-	2,63	2,52	3,08	2,44	2,03	3,02
FRA follow up	4	3,1	3,25	0,25	2,63	2,77	3,1	2,85	2,17	2,95
FRA CG	5	4,0	2,76	0,20	2,9	2,84	3,28	2,84	2,44	2,82

9.8. Initial and last (adapted) version of the STSQ

9.8.1. Initial version of the STSQ

## **The initial version of the STSQ**

### Introduction

People experience various situations differently. This device deals with associations, feelings, and attitudes about assorted topics. You will be presented with different assessments, which will be explained each separately.

All of the information you provide will be treated as confidential. No information about your identity will be requested at any stage.

- Please read all the instructions carefully, and answer the questions as honestly and accurately as possible. There are no "right", or "wrong" answers!
- Please answer all items.

### I. Please answer each of the following questions to your person!

1. Age \_\_\_\_\_ years

2. Sex  male  female

### II. Have a closer look at each of the following pictures!

1 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

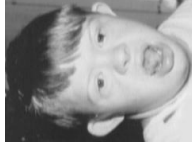
1 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.

2 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

2 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel / ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.

3 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

3 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel / ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.



- 4 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

- 4 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

I feel/ ...

1. close	1	2	3	4	5
2. dissimilar	1	2	3	4	5
3. sympathetic	1	2	3	4	5

... towards the picture above.

- 5 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

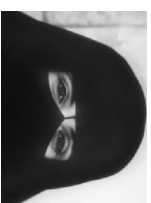
- 5 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

I feel/ ...

1. close	1	2	3	4	5
2. dissimilar	1	2	3	4	5
3. sympathetic	1	2	3	4	5

... towards the picture above.

- 6 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

- 6 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

I feel/ ...

1. close	1	2	3	4	5
2. dissimilar	1	2	3	4	5
3. sympathetic	1	2	3	4	5

... towards the picture above.

- 7 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

- 7 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

I feel/ ...

1. close	1	2	3	4	5
2. dissimilar	1	2	3	4	5
3. sympathetic	1	2	3	4	5

... towards the picture above.

- 8 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

- 8 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.

- 9 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

- 9 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.

- 10 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

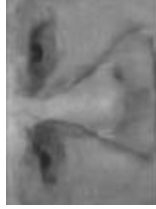
- 10 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.

- 11 a) What comes first to your mind when you see the following picture?  
Please describe your initial associations!



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

- 11 b) Please evaluate your feelings towards the picture above on each of the following scales!  
Indicate your subjective relationship by crossing out the respective number:  
5 = highest value; 1 lowest value!

/ feel ...

1.	close	1	2	3	4	5
2.	dissimilar	1	2	3	4	5
3.	sympathetic	1	2	3	4	5

... towards the picture above.

### III. Put yourself into the following situations!

1) *Imagine people are entering a bus. All passengers pass by the conductor. One person was asked to show his/her ticket.*

**What reasons could the bus driver have to stop this person?  
Please evaluate the probability of each suggested reason!**

- |   |                      |                 |               |                    |
|---|----------------------|-----------------|---------------|--------------------|
| 1. The person looks different                     | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 2. The person's behaviour seems to be suspicious. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 3. The bus driver is irritated.                   | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 4. Generally, strangers are controlled.           | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 5. The bus driver does not like foreigners.       | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 6. It was a routine control.                      | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |

Further reasons \_\_\_\_\_

2) *Imagine a basketball game. Team A is playing with a new team member who is expected to lead the team to the top. The match is over, and they have lost the game.*

**How could the coach explain the result?  
Please evaluate the probability of each suggested reason!**

- |   |                      |                 |               |                    |
|---|----------------------|-----------------|---------------|--------------------|
| 1. The team was in poor condition.                  | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 2. The new player did not play as well as expected. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 3. The new player did not adapt to the team.        | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 4. The new player disturbed the team spirit.        | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 5. The player was not integrated into the new team. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 6. The team opposed to the new player.              | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |

Further reasons \_\_\_\_\_

- 3) *Yildiz is a student at a high school for boys and girls. Outside of her home, she wears wide clothes and a head scarf. Last week Yildiz's family applied to exempt her from the co-educational Physical Education (P.E.) class because her family's Islamic faith does not permit girls to participate in sport together with boys.*

**How could the principal of the school argue?  
Please evaluate the probability of each suggested argument!**

1. Different religious faiths should be respected.
 

1	2	3	4
most improbable	improbable	probable	most probable
2. P.E. can be arranged in a mutually acceptable way.
 

1	2	3	4
most improbable	improbable	probable	most probable
3. P.E. class could be organized gender-separated.
 

1	2	3	4
most improbable	improbable	probable	most probable
4. Co-educational teaching should have priority over individual beliefs.
 

1	2	3	4
most improbable	improbable	probable	most probable
5. Pupils should obey existing rules.
 

1	2	3	4
most improbable	improbable	probable	most probable
6. Students have the possibility to select exercises.
 

1	2	3	4
most improbable	improbable	probable	most probable

Further alternatives

---

- 4) *Imagine a company has to decide between two final applicants for a leadership position. The final candidates are a man and a woman.*

**How could the company argue?  
Please evaluate the probability of each suggested argument!**

1. In case of equal qualification, the gender-balance determines the decision.
 

1	2	3	4
most improbable	improbable	probable	most probable
2. Men are supposed to be more efficient.
 

1	2	3	4
most improbable	improbable	probable	most probable
3. Professional qualification determines the decision.
 

1	2	3	4
most improbable	improbable	probable	most probable
4. The heterogeneity of the company structure will influence the decision.
 

1	2	3	4
most improbable	improbable	probable	most probable
5. Women are supposed to be more integrative leaders.
 

1	2	3	4
most improbable	improbable	probable	most probable
6. Personal leadership determines the decision.
 

1	2	3	4
most improbable	improbable	probable	most probable

Further alternatives

---

5) *Imagine a student begins to study at the university. Neither people, nor structures are familiar to him/her.*

**Please evaluate the probability of each suggested feeling/reason!**

- |   |                      |                 |               |                    |
|---|----------------------|-----------------|---------------|--------------------|
| 1. He/she feels insecure and disorientated.   | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 2. It must be exciting for him/her to discover the new environment.                 | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 3. The unknown expectation of the new surrounding makes him/her feel uncomfortable. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 4. Getting to know people in such an impersonal situation must be difficult.        | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 5. This situation does not provoke specific irritations.                            | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 6. He/she is confident to overcome the uncomfortable feelings of the first day.     | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |

Further alternatives

**IV. Please evaluate each of the following statements?**

- |   |                        |               |            |                     |
|---|------------------------|---------------|------------|---------------------|
| 1. Other people are not as open-minded as I am.                                       | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 2. My personal way of life should be a model for other people.                        | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 3. It seems that different people have different values.                              | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 4. In spite of individual differences people do not differ substantially.             | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 5. To interact with different kinds of people is enjoyable.                           | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 6. Sitting in a group of strangers provokes uncomfortable feelings.                   | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 7. In spite of its apparently continuous change, culture contains some stable values. | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 8. It is important to forward cultural values to the next generations.                | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 9. Unforeseen events upset me.  | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 10. It is necessary to plan ahead in order to avoid surprises.                        | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |

11. The open result is responsible for the excitement of a competition.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree
12. There is no freedom without rules.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree
13. Being obliged to make decisions provokes uncertainty.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree
14. For activities, I prefer precise instructions to open suggestions.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree
15. The number of friends I make depends only on me and my behaviour.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree
16. Fate determines whether I have more or fewer friends.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree
17. Most of the events in my life are determined by other people.
- 1 strongly disagree      2 disagree      3 agree      4 strongly agree

**V. Now, here are the final questions!**

1. Your native language \_\_\_\_\_
2. Further languages \_\_\_\_\_
3. Your country of residence \_\_\_\_\_
4. Have you ever lived in other countries? YES NO  
 If YES, in which countries and how long?  
 country \_\_\_\_\_ from \_\_\_\_\_ till \_\_\_\_\_  
 country \_\_\_\_\_ from \_\_\_\_\_ till \_\_\_\_\_  
 country \_\_\_\_\_ from \_\_\_\_\_ till \_\_\_\_\_  
 country \_\_\_\_\_ from \_\_\_\_\_ till \_\_\_\_\_
5. From which countries do your parents originate?  
 mother \_\_\_\_\_ father \_\_\_\_\_

**Thank you for your cooperation!**

9.8.2. Last (adapted) version of the STSQ

## **The last (adapted) version of the STSQ**

Code: 

--	--	--	--

**Introduction**

People experience various situations differently. This device deals with associations, feelings, and attitudes about assorted topics. You will be presented with different assessments, which will be explained each separately.

All of the information you provide will be treated as confidential. No information about your identity will be requested at any stage.

- Please read all the instructions carefully, and answer the questions as honestly and accurately as possible. There are no "right", or "wrong" answers!
- Please answer all items.

**Please answer each of the following questions to your person!**

1. Age \_\_\_\_\_ years
2. Sex  male  female

**Part I: Please have a closer look at each of the following pictures!**

1 a) Please give a headline to the following picture!



1 b) Please cross out the respective number!

1. How different do you perceive yourself compared to the person illustrated in the picture above?

*low*

1	2	3	4	5
---	---	---	---	---

*high*

2. How sympathetic do you feel, when you imagine the person illustrated above?

*low*

1	2	3	4	5
---	---	---	---	---

*high*

2 a) Please give a headline to the following picture!



2 b) Please cross out the respective number!

1. How different do you perceive yourself compared to the person illustrated in the picture above?

*low*

1	2	3	4	5
---	---	---	---	---

*high*

2. How sympathetic do you feel, when you imagine the person illustrated above?

*low*

1	2	3	4	5
---	---	---	---	---

*high*

3 a) Please give a headline to the following picture!



3 b) Please cross out the respective number!

1. How different do you perceive yourself compared to the person illustrated in the picture above?

*low*

1	2	3	4	5
---	---	---	---	---

*high*

2. How sympathetic do you feel, when you imagine the person illustrated above?

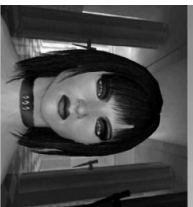
*low*

1	2	3	4	5
---	---	---	---	---

*high*



4 a) Please give a headline to the following picture!



\_\_\_\_\_

\_\_\_\_\_

4 b) Please cross out the respective number!

1. How different do you perceive yourself compared to the person illustrated in the picture above?

low 

1	2	3	4	5
---	---	---	---	---

 high

2. How sympathetic do you feel, when you imagine the person illustrated above?

low 

1	2	3	4	5
---	---	---	---	---

 high

5 a) Please give a headline to the following picture!



\_\_\_\_\_

\_\_\_\_\_

5 b) Please cross out the respective number!

1. How different do you perceive yourself compared to the person illustrated in the picture above?

low 

1	2	3	4	5
---	---	---	---	---

 high

2. How sympathetic do you feel, when you imagine the person illustrated above?

low 

1	2	3	4	5
---	---	---	---	---

 high

Part II: Put yourself into the following situations!

1) *Imagine people are entering a bus. All passengers pass by the conductor. One passenger was asked to show his/her ticket.*

**What reasons could the bus driver have to stop this passenger? Please indicate the probability of each suggested reason!**

- |   |                      |                 |               |                    |
|---|----------------------|-----------------|---------------|--------------------|
| 1. The passenger appears suspiciously.              | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 2. The passenger looks different.                   | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 3. It was a routine control.                        | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 4. Strangers are often asked to show their tickets. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 5. The bus driver does not like foreigners.         | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 6. The bus driver is irritated.                     | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |

Further reasons

\_\_\_\_\_

2) *Imagine a basketball game. Team A is playing with a new team member who is expected to lead the team to the top. The match is over, and they have lost the game.*

**How could the coach explain the result?**

**Please indicate the probability of each suggested reason!**

1. The new player was not yet integrated into the new team. 1 most improbable    2 improbable    3 probable    4 most probable
2. The new player did not adapt to the team. 1 most improbable    2 improbable    3 probable    4 most probable
3. The new player did not play as well as expected. 1 most improbable    2 improbable    3 probable    4 most probable
4. The team opposed to the new player. 1 most improbable    2 improbable    3 probable    4 most probable
5. The new player disturbed the team spirit. 1 most improbable    2 improbable    3 probable    4 most probable
6. The team has to work on a new game strategy. 1 most improbable    2 improbable    3 probable    4 most probable

Further reasons

---

3) *Yildiz is a student at a high school for boys and girls. Outside of her home, she wears wide clothes and a head scarf. Last week Yildiz's family applied to exempt her from the co-educational Physical Education class because her family's Islamic faith does not permit girls to participate in sport together with boys.*

**How could the principal of the school argue?**

**Please indicate the probability of each suggested argument!**

1. Different religious faiths should be respected. 1 most improbable    2 improbable    3 probable    4 most probable
2. Physical Education can be arranged in a mutually acceptable way. 1 most improbable    2 improbable    3 probable    4 most probable
3. Students should obey existing rules. 1 most improbable    2 improbable    3 probable    4 most probable
4. Students are given the opportunity to choose between exercises in Physical Education class. 1 most improbable    2 improbable    3 probable    4 most probable
5. Co-educational teaching should have priority over individual beliefs. 1 most improbable    2 improbable    3 probable    4 most probable
6. Physical Education class could be organized gender-separated. 1 most improbable    2 improbable    3 probable    4 most probable

Further alternatives

---

4) *Imagine a company has to decide between two final applicants for a leadership position. The final candidates are a man and a woman.*

**How could the company argue?  
Please indicate the probability of each suggested argument!**

- |  |                      |                 |               |                    |
|--|----------------------|-----------------|---------------|--------------------|
| 1. Professional qualification determines the decision.   | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 2. Personal leadership determines the decision.  | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 3. In case of equal qualification, the balance of gender within the company determines the decision. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 4. Women are supposed to be more integrative leaders.  | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 5. Men are supposed to be more efficient.  | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 6. The heterogeneity within the company influences the decision.                                     | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |

Further alternatives

---

5) *Imagine a student begins to study at the university. Neither people, nor structures are familiar to him/her.*

**Please indicate the probability of each suggested feeling/reason!**

- |   |                      |                 |               |                    |
|---|----------------------|-----------------|---------------|--------------------|
| 1. The new surrounding makes him/her feel uncomfortable.                        | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 2. He/she feels insecure and disorientated.                                     | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 3. It is difficult to discover the expectations of the new surrounding.         | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 4. He/she is confident to overcome the uncomfortable feelings of the first day. | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 5. This situation does not provoke specific irritations.                        | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |
| 6. Getting to know people in such an impersonal situation must be difficult.    | 1<br>most improbable | 2<br>improbable | 3<br>probable | 4<br>most probable |

Further alternatives

---

**Part III: Please evaluate each of the following statements?**

- |  |                        |               |            |                     |
|--|------------------------|---------------|------------|---------------------|
| 1. Open-mindedness helps understanding.  | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 2. People differ substantially.  | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 3. To meet different kinds of people is enjoyable.   | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 4. Sitting in a group of strangers provokes uncomfortable feelings.                          | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 5. Life brings changes continuously.   | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 6. Culture has some stable values.   | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 7. It is important to forward cultural values to the next generations.                       | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 8. I like unforeseen events.   | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 9. To plan ahead helps to provide security.  | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 10. The uncertain outcome is responsible for the excitement of a competition.                | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 11. Freedom does not mean the absence of rules.  | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 12. Having to make decisions makes me feel uncertain.  | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 13. When engaging in an activity, I prefer receiving clear instructions to open suggestions. | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 14. Fate determines whether I have more or fewer friends.                                    | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 15. Most of the events in my life are determined by other people.                            | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |
| 16. Having responsibility makes me feel uncertain.   | 1<br>strongly disagree | 2<br>disagree | 3<br>agree | 4<br>strongly agree |

---

**Now, here are the final questions!**

1. Your native language

\_\_\_\_\_

2. Further languages

\_\_\_\_\_

3. Your country of residence

\_\_\_\_\_

4. Have you ever lived in other countries?  
If YES, in which countries and how long?

YES

NO

country from \_\_\_\_\_ till \_\_\_\_\_

country from \_\_\_\_\_ till \_\_\_\_\_

country from \_\_\_\_\_ till \_\_\_\_\_

country from \_\_\_\_\_ till \_\_\_\_\_

5. From which countries do your parents originate?

mother \_\_\_\_\_

father \_\_\_\_\_

**Thank you for your cooperation!**



## 9.9. Pre-study 1

### **Pre-study 1: Writing task**

### Introduksjon

Du blir nå deltaker av en vitenskapelig undersøkelse. All informasjon behandles helt anonymt. Oppgavene er selvfølgelig frivillig, og det er viktig for mitt forskningsprosjekt at du svarer så ærlig som mulig. Det finnes ingen gale svar på spørsmålene.

Temaområdet er **forskjellighet mellom mennesker og hva det kan bety for deg hvis du legger merke til at noe eller noen er forskjellig (fremmedhet).**

For å kunne beskrive fenomenet "fremmed" mye bedre, trenger jeg deres hjelp. Vær så snill svar på følgende spørsmål

### (A) Personlige opplysninger

<b>1. Alder</b>			<input type="text"/>
<b>2. Kjønn</b>			
	kvinne		<input type="checkbox"/>
	mann		<input type="checkbox"/>
<b>3. Er du født i Norge?</b>			
	ja		<input type="checkbox"/>
	nei		<input type="checkbox"/>
<b>4. Er begge foreldrene dine av utlandsk opprinnelse?</b>			
	ja, begge er født og oppvokst i et annet land		<input type="checkbox"/>
	nei, en eller begge av foreldrene er født og oppvokst i Norge.		<input type="checkbox"/>
<b>5. Er foreldrene dine fra et vestlig eller ikke-vestlig land?</b>			
	De er begge fra vestlig land (Vest-Europa (unntatt Tyrkia), Nord-Amerika, Australia)		<input type="checkbox"/>
	De er begge fra et ikke-vestlig land (Øst-Europa, Asia, Afrika, Sør- og Mellom-Amerika og Tyrkia)		<input type="checkbox"/>
	En fra et vestlig og en fra et ikke-vestlig land		<input type="checkbox"/>
	fra et annet land	_____	



**(B) Oppgaver**

**(1) Hvem passer ikke inn i rekken? Hvorfor? (Prøv å begrunne svaret ditt!)**




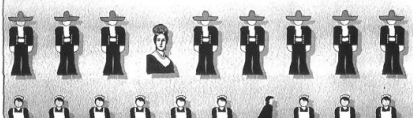

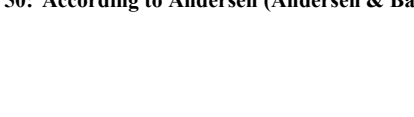
1.		fordi _____
2.		fordi _____
3.		fordi _____
4.		fordi _____
5.		fordi _____
6.		fordi _____

Figure 50: According to Andersen (Andersen & Bali, 2000; Andersen & Bali, 2002)

**(2) Hva betyr fremmed for deg? Prøv å beskrive hva fremmedhet er!**

**(3) Fortell om situasjoner hvor du følte deg fremmed!**

(a) Beskriv kort situasjoner **(gjerne flere!)**

*(du kan også bruke baksiden av arket)*

(b) Hva følte du i situasjonen?

(c) Hvorfor følte du deg fremmed? Nevne noen grunner!

9.10. Pre-study 2

**Pre-study 2.1: Validation of collected pictures and situations**

## Introduksjon

Du blir nå deltaker av en vitenskaplig undersøkelse. Alle informasjonen behandles absolutt fortrolig og anonym. Oppgaver er selvfølgelig frivillig, men det ville være veldig hjelpsom for mitt forskningsprosjekt at du svarer ærlig og så god som mulig. Du kan ikke gir et feil svar på spørsmålene - det finnes ingen riktig eller feil. Det er bare veldig viktig at du bearbeider oppgavene alvorlig.

På forhånd tusen takk!

## Personlige opplysninger

1. Alder

2. Kjønn

kvinne

mann

3. Er du født i Norge?






ja



nei





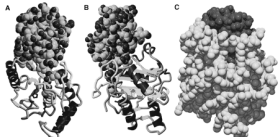

**Part I: Classify pictures with given terms**

- a) Please arrange the following terms to the pictures.
- b) If you do not find a suitable word, try to find an alternative.

male (*mannlig*), educated (*utdannet*), religious (*religiøs*), devalued (*nedvurderende*), athletic (*atletisk*), overweight (*overvekt*), old (*gammel*), female (*kvinnelig*), handicapped (*handikappet*), complex (*kompleks*), uninformed (*uvitende*), immoral (*umoralsk*), dark skin tone (*mørk hudfarge*), appearance (*utseende*), homosexual (*homoseksuell*), culture (*kultur*), abstract (*abstract*), white (*hvit*), difficult (*vanskelig*), fat (*fet*), gay (*homofil*), stupid (*dum*), fundamentalist (*fundamentalistisk*), sporty (*sporty*), sexy (*sexy*), man (*mann*), professor (*professor*), leader (*leder*), teacher (*lærer*), intelligent (*intelligent*), written (*skriftlig*), different (*forskjellig*), strange (*fremmed*), valuable (*verdig*), knowledge (*kunnskap*), young (*ung*), domestically (*huslig*), skin colour (*hudfarge*), pure (*ren*), erotic (*erotisk*), traditional (*tradisjonell*), communication (*kommunikasjon*)

Nr.	Item	Term	Better alternative
a)			
b)			
c)			
d)			
e)			

f)			
g)			
h)			
i)			
j)			

k)			
l)			
m)			
n)			
o)			
p)			



**Part II: Please assign the terms to the described situations.**

Nr.	Item	Term	Better alternative
1.	<i>Nur bei mir schaute der Busfahrer genauer auf meinen Fahrausweis. Alle anderen konnten einfach so durchgehen.</i>		
2.	<i>Stellen Sie sich die folgende Situation in einer Sportunterrichtsstunde in der: Johanna wird beim Auswählen der Mannschaften immer zu letzt gewählt.</i>		
3.	<i>Stellen Sie sich die folgende Situation vor: Frau Meier ist Lehrerin an einer Schule. Meist hat sie guten Kontakt zu ihren Schülern. Seltsamerweise verhält sie sich nur den beiden Mädchen mit Kopftuch irgendwie seltsam gegenüber.</i>		
4.	<i>Marco ist ein Basketballspieler. Er bekommt die Möglichkeit im Ausland zu trainieren. Nach seiner Rückkehr berichtet er, dass es zu Beginn ein wenig unwohl oder sogar ängstlich in seinem neuen Team gefühlt hat.</i>		
5.	<i>Yousaf ist im Ausland aufgewachsen. Obwohl er schon sehr lange im Ausland lebt, fühlt er sich oft fremd.</i>		
6.	<i>Am Ende eines Arbeitstages fehlten 100 Euro in der Kasse. An diesem Tag haben zwei Leute an dieser Kasse gearbeitet, Anna und Achmed.</i>		
7.	<i>Der Lehrer zitiert ein lustiges Sprichwort in der Schulklasse. Alle lachen, bis auf eine Schülerin.</i>		
8.	<i>Stellen Sie sich folgende Situation vor: Im Skiurlaub hat Namal oft das Gefühl, dass ihn häufig Leute auf der Piste verwundert anstarren.</i>		



## **Pre-study 2.2: Cross-validation of collected pictures and situations**

## Pre-study 2.2

### Introduksjon

Du blir nå deltaker av en vitenskaplig undersøkelse. Alle informasjonen behandles absolutt fortrolig og anonym. Oppgaver er selvfølgelig frivillig, men det ville være veldig hjelpsom for mitt forskningsprosjekt at du svarer ærlig og så god som mulig. Du kan ikke gir et feil svar på spørsmålene - det finnes ingen riktig eller feil. Det er bare veldig viktig at du bearbeider oppgavene alvorlig.


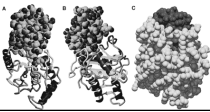

På forhånd tusen takk!

### Personlige opplysninger

1. Alder		<input type="text"/>
2. Kjønn	kvinne	<input type="checkbox"/>
	mann	<input type="checkbox"/>
3. Er du født i Norge?	ja	<input type="checkbox"/>
	nei	<input type="checkbox"/>

**Part I: Classify pictures with given terms**

- a) Please arrange the following pictures and situations to the best fitting words.
- b) If you do not find a suitable picture, draw or describe an alternative.

			
<b>B</b>	<b>O</b>	<b>E</b>	<b>D</b>
			
<b>F</b>	<b>K</b>	<b>H</b>	<b>N</b>
			
<b>C</b>	<b>L</b>	<b>M</b>	<b>I</b>
			
<b>P</b>	<b>J</b>	<b>G</b>	<b>A</b>

Nr.	Item
I	<i>Nur bei mir schaute der Busfahrer genauer auf meinen Fahrausweis. Alle anderen konnten einfach so durchgehen.</i>
II	<i>Stellen Sie sich die folgende Situation in einer Sportunterrichtsstunde in der: Johanna wird beim Auswählen der Mannschaften immer zu letzt gewählt.</i>
III	<i>Stellen Sie sich die folgende Situation vor: Frau Meier ist Lehrerin an einer Schule. Meist hat sie guten Kontakt zu ihren Schülern. Seltsamerweise verhält sie sich nur den beiden Mädchen mit Kopftuch irgendwie seltsam gegenüber.</i>
IV	<i>Marco ist ein Basketballspieler. Er bekommt die Möglichkeit im Ausland zu trainieren. Nach seiner Rückkehr berichtet er, dass es zu Beginn ein wenig unwohl oder sogar ängstlich in seinem neuen Team gefühlt hat.</i>
V	<i>Yousaf ist im Ausland aufgewachsen. Obwohl er schon sehr lange im Ausland lebt, fühlt er sich oft fremd.</i>
VI	<i>Am Ende eines Arbeitstages fehlten 100 Euro in der Kasse. An diesem Tag haben zwei Leute an dieser Kasse gearbeitet, Anna und Achmed.</i>
VII	<i>Der Lehrer zitiert ein lustiges Sprichwort in der Schulklasse. Alle lachen, bis auf eine Schülerin.</i>
VIII	<i>Stellen Sie sich folgende Situation vor: Im Skiurlaub hat Namal oft das Gefühl, dass ihn häufig Leute auf der Piste verwundert anstarren.</i>



Nr.	Item	Picture (bokstave).	Situations (røm. tall)
1.	male ( <i>mannlig</i> ),		
2.	educated ( <i>utdannet</i> )		
3.	religious ( <i>religiøs</i> )		
4.	devalued ( <i>nedvurderende</i> )		
5.	athletic ( <i>atletisk</i> )		
6.	overweight ( <i>overvekt</i> )		
7.	old ( <i>gammel</i> )		
8.	female ( <i>kvinnelig</i> )		
9.	handicapped ( <i>handikappet</i> )		
10.	complex ( <i>kompleks</i> )		
11.	uninformed ( <i>uvitende</i> )		
12.	immoral ( <i>umoralsk</i> )		
13.	dark skin tone ( <i>mørk hudfarge</i> )		
14.	appearance ( <i>utseende</i> )		
15.	homosexual ( <i>homoseksuell</i> )		
16.	culture ( <i>kultur</i> )		
17.	abstract ( <i>abstract</i> )		
18.	white ( <i>hvit</i> )		
19.	difficult ( <i>vanskelig</i> )		
20.	fat ( <i>fet</i> )		
21.	gay ( <i>homofil</i> )		
22.	stupid ( <i>dum</i> )		
23.	fundamentalist ( <i>fundamentalistisk</i> )		
24.	sporty ( <i>sparty</i> )		
25.	sexy ( <i>sexy</i> )		

26.	man ( <i>mann</i> )		
27.	professor ( <i>professor</i> )		
28.	leader ( <i>leder</i> )		
29.	teacher ( <i>lærer</i> )		
30.	intelligent ( <i>intelligent</i> )		
31.	written ( <i>skriftlig</i> )		
32.	different ( <i>forskjellig</i> )		
33.	strange ( <i>fremmed</i> )		
34.	valuable ( <i>verdig</i> )		
35.	knowledge ( <i>kunnskap</i> )		
36.	young ( <i>ung</i> )		
37.	domestically ( <i>huslig</i> )		
38.	skin colour ( <i>hudfarge</i> )		
39.	pure ( <i>ren</i> )		
40.	erotic ( <i>erotisk</i> )		
41.	traditional ( <i>tradisjonell</i> )		
42.	communication ( <i>kommunikasjon</i> )		



## 9.11. Summary of item changes and adaptations during the development and empirical testing of the STSQ

The table above should help to follow the main changes of items, and emphasize the successively development of the STSQ.

STSQ versions	Part	Changes/adaptations <sup>69</sup>																																			
1 → 2	I	<p><b>initial version</b></p> <p>1 a) What comes first to your mind when you see the following picture? Please describe your initial associations!</p>  <p>_____</p> <p>_____</p> <p>_____</p> <p>1 b) Please evaluate your feelings towards the picture above on each of the following scale! Indicate your subjective relationship by crossing out the respective number: 5 = highest value; 1 lowest value!</p> <p>/ feel ...</p> <table border="0"> <tr> <td>1. close</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>2. dissimilar</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>3. sympathetic</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> </table> <p>...towards the picture above.</p>	1. close	1	2	3	4	5	2. dissimilar	1	2	3	4	5	3. sympathetic	1	2	3	4	5	<p><b>subsequent version</b></p> <p>1 a) Please give a headline to the situation shown in the following picture!</p>  <p>_____</p> <p>_____</p> <p>_____</p> <p>1 b) Please indicate how familiar (1), and how sympathetic (2) you feel imagining the given situation in the picture above. Please cross out the respective number:</p> <table border="0"> <tr> <td>1. Familiarity:</td> <td>low</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>high</td> </tr> <tr> <td>2. Sympathy:</td> <td>low</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>high</td> </tr> </table>	1. Familiarity:	low	1	2	3	4	5	high	2. Sympathy:	low	1	2	3	4	5	high
1. close	1	2	3	4	5																																
2. dissimilar	1	2	3	4	5																																
3. sympathetic	1	2	3	4	5																																
1. Familiarity:	low	1	2	3	4	5	high																														
2. Sympathy:	low	1	2	3	4	5	high																														
	III	<p><b>Other people are not as open-minded as I am.</b></p> <p>My personal way of life should be a model for other people.</p> <p>In spite of individual differences people do not differ substantially.</p> <p>To interact with different kinds of people is enjoyable.</p> <p>It is necessary to plan ahead in order to avoid surprises.</p>																																			
		<p><b>Open-mindedness helps understanding.</b></p> <p>--</p> <p>People differ substantially.</p> <p>To meet different kinds of people is enjoyable.</p> <p>To plan ahead helps to provide security.</p>																																			

<sup>69</sup> Each chapter of the reported studies contains an argumentation for the adaptations made.

		There is no freedom without rules.	Freedom does not mean the absence of rules.
2 → 3	I	Item I.1 – I.11	I.1, I.3, I.9
	II	Initial item pool	<i>no changes</i>
	III	It seems that different people have different values. In spite of its apparently continuous change, culture contains some stable values. Unforeseen events upset me. For activities, I prefer precise instructions to open suggestions. The number of friends I make depends only on me and my behaviour.	<i>excluded from the item pool</i> Life brings continuously changes. Culture has some stable values. I like unforeseen events. When engaging in an activity, I prefer receiving clear instructions to open suggestions. <i>excluded from the item pool</i>
3 → 4	I	Instructions: “Please give a headline to the situation shown in the following picture!” Instructions “Please indicate how familiar (1.), and how sympathetic (2.) you feel imaging the given situation in the picture above.” Four point interval scale level Systematic structure/sequence of items	Please give a headline to the following picture!  Please indicate how familiar (1.), and how sympathetic (2.) you feel, when you imagine the situation illustrated above.  Reduction to a dichotomous scale (improbable/probable) for all items of part II new sequence of items, sorted/structured by increasing difficulty index
	II		
	..... Sit. 5	The unknown expectation of the new surrounding makes him/her feel uncomfortable.	The new surrounding makes him/her feel uncomfortable. It is difficult to discover the expectations of the new surroundings. <i>no changes</i>
	III		

4 → 5	I	<p>Please indicate how familiar (1), and how sympathetic (2) you feel, when you imagine the situation illustrated above. Please cross out the respective number:</p> <p>1. Familiarity:      low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>2. Sympathy:      low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>Reduced item pool: I.1, I.3, I.9</p>	<p>Please cross out the respective number!</p> <p>1. How different do you perceive yourself compared to the person shown in the picture above? low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>2. How sympathetic do you feel, when you imagine the person illustrated above? low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>Extended item pool: I.1, I.3, I.9 + I.12 + I.13 + I.14 + I.15 + I.16</p>
	II		<i>no changes</i>
	III		<i>no changes</i>
5 → 6	I	<p>Please cross out the respective number!</p> <p>1. How different do you perceive yourself compared to the person shown in the picture above? low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>2. How sympathetic do you feel, when you imagine the person illustrated above? low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>Item pool: I.1, I.3, I.9, I.12, I.13, I.14, I.15, I.16</p>	<p>Please cross out the respective number!</p> <p>1. How different do you perceive yourself compared to the person illustrated in the picture above? low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>2. How sympathetic do you feel, when you imagine the person illustrated above? low <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> high</p> <p>Reduction of item pool: I.1, I.3, I.9, I.12, I.16</p>
	II	Dichotomous scale level (probable or improbable)	Reduction of item pool: I.1, I.3, I.9, I.12, I.16 Changed back to initial 4 point interval scale level (from most improbable to most probable)
	Sit. I	<p>“Imagine people are entering a bus. All passengers pass by the conductor. One person was asked to show his/her ticket.”</p> <p>“What reasons could the bus driver have to stop this person? [...]”</p>	<p>“Imagine people are entering a bus. All passengers pass by the conductor. One passenger was asked to show his/her ticket.”</p> <p>“What reasons could the bus driver have to stop this passenger? [...]”</p>
		The person’s behaviour seems to be suspicious.	The passenger appears suspiciously.
		The person looks different.	The passenger looks different.
		Generally, strangers are controlled.	Strangers are often asked to show their tickets.

	Sit. 2	The player was not integrated into the new team. The team was in poor condition.	The new player was not yet integrated into the new team. The team has to work on a new game strategy.
	Sit. 3	P.E. can be arranged in a mutually acceptable way.	Physical Education can be arranged in a mutually acceptable way.
		Pupils should obey existing rules.	Students should obey existing rules.
		Students have the possibility to select exercises.	Students are given the opportunity to choose between exercises in Physical Education class.
	Sit. 4	P.E. class could be organized gender-separated.	Physical Education class could be organized gender-separated.
		In case of equal qualification, the gender-balance determines the decision.	In case of equal qualification, the balance of gender within the company determines the decision.
		The heterogeneity of the company structure will influence the decision.	The heterogeneity within the company influences the decision.
	III	The open result is responsible for the excitement of a competition.	The uncertain outcome is responsible for the excitement of a competition.
		Being obliged to make decisions provokes uncertainty.	Having to make decisions makes me feel uncertain.
		--	Added (parallel) item: Having responsibility makes me feel uncertain.

---

## 10. Reference List

- Andersen, E. F. & Bali, B. (2000). *Anflug.de 3 - Textbuch*. Oslo: Cappelen.
- Andersen, E.F. & Bali, B. (2002). *@nflug.de3: Arbeitsbuch (und Textbuch) 2. klasse, B-språk, 3. klasse C-språk*. Cappelen.
- Atherton, J. S. (2009). Learning and teaching: Assimilation and Accommodation [online], <http://www.learningandteaching.info/learning/assimacc.htm>, 27.5.2009.
- Auernheimer, G. (1995). *Einführung in die interkulturelle Erziehung*. (2 ed.) Darmstadt: Wissenschaftliche Buchgesellschaft.
- Auernheimer, G. (2003). *Einführung in die interkulturelle Pädagogik. 5./6. Schuljahr*. Wissenschaftliche Buchgesellschaft.
- Auernheimer, G. (2005). *Interkulturelle Kommunikation, vierdimensional betrachtet* [online], <http://www.uni-koeln.de/ew-fak/paedagogik/interkulturelle/publikationen/muenchen.html>, downloaded 12.4.2005.
- Baumann, Z. (1995). *Moderne und Ambivalenz. Das Ende der Eindeutigkeit*. Frankfurt am Main: Fischer.
- Bender-Szymanski, D. (2004). Die schwierige Toleranz. Ein Planspiel zu religiös-weltanschaulichen Konflikten in der Schule. *Praxis Schule 5-10*, (6), 30-37.
- Bender-Szymanski, D. (2006). Die schwierige Toleranz. Ein empirisch gestützter Erfahrungsbericht über die Vorstellung einer Lehr-Lernsequenz zu einem religiös-weltanschaulichen Konflikt im Sportunterricht. In D. Blecking & P. Gieß-Stüber (Eds.), *Sport bewegt Europa. Beiträge zur interkulturellen Verständigung* (pp. 186-216). Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Bennet, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for intercultural experience* (pp. 21-72), Yarmouth, Maine: Intercultural Press.

- 
- Bibliographisches Institut & F.A.Brockhaus AG (2008). Der Brockhaus multimedial 2008 premium (Version 10) [Computer Software]. Bibliographisches Institut & F.A.Brockhaus AG.
- Blecking, D. (2008). „Homo migrans’ plays: migration and sport in historical perspective. In: Gieß-Stüber, P. and Blecking, D. (Eds.), Sport, Integration, Europe: widening horizons in intercultural education. Hohengehren: Schneider Verlag, 51-55.
- Bogardus, E. S. (1925). Measuring social distance. *Journal of Applied Sociology*, (9), 299-308.
- Borkenau, P., Egloff, B., Eid, M., Henning, J., Kersting, M., Neubauer, A. C. et al. (2005). Persönlichkeitspsychologie: Stand und Perspektiven. *Psychologische Rundschau*, 54, 271-290.
- Bortz, J. (2005). *Statistik: Für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation*. Heidelberg: Springer Medizin.
- Bortz, J. & Lienert, G. A. (2003). *Kurzgefaßte Statistik für die klinische Forschung*. Berlin, Heidelberg, New York: Springer.
- Brennan, R. L. (2001). *Generalizability Theory*. New York: Springer.
- Bruner, J. S. & Postman, L. (1949). Perception, cognition and behavior. *Journal of Personality*, 18, 14-31.
- Bröskamp, B. (1994). Körperliche Fremdheit. Zum Problem der interkulturellen Begegnung im Sport. Sankt Augustin: Academia.
- Bröskamp, B. (2008). Bodily strangerhood(s) revisited. In: Gieß-Stüber, P. and Blecking, D. (Eds.), Sport, Integration, Europe: widening horizons in intercultural education. Hohengehren: Schneider Verlag, 213-227.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.

- 
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Buhr, K. & Dugas, M. J. (2002). The intolerance of uncertainty scale: psychometric properties of the English version. *Behaviour Research and Therapy*, 40, 931-945.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56, 81-105.
- Chen, G. M. & Starosta, W. J. (2000). The Development and Validation of the Intercultural Sensitivity Scale. *Human Communication*, 3, 2-14.
- Chen, G. M. & Starosta, W. J. (2002). Intercultural Awareness. In L. A. Samovar & R. E. Porter (Eds.), *Intercultural communication: A reader*. (10 ed., pp. 344-353). Belmont, CA: Wadsworth.
- Chen, G.-M. & Starosta, W. J. (1997). A Review of the Concept of Intercultural Sensitivity. *Human Communication*, 1, 1-16.
- Clason, D.J. & Dormody, T.J. (1994). Analyzing Data Measured by Individual Likert-Type Items, *Journal of Agricultural Education*, 35 (4), 31-35.
- Cooley, C.H. (1902). *Human nature and the social order*. New York: Charles Scribner's sons.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.
- Cronbach, L. J. & Shavelson, R. J. (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, 64, 391-418.
- Der Duden (2001). *Duden Herkunftswörterbuch*. (vols. 7). Mannheim: Bibliographisches Institut.
- Downing, S. M. & Haladyna, T. M. (1997). Test item development: Validity evidence from quality assurance procedures. *Applied Measurement in Education*, 10, 61-82.
- Earley, P.C. & Ang, S. (2003). *Cultural intelligence: individual interactions across cultures*. Stanford: University Press.

- 
- Erdheim, M. (1988). *Die Psychoanalyse und das Unbewußte in der Kultur*. Frankfurt am Main: Suhrkamp.
- Erdmann, R. (1979). Ein Verfahren zur Erfassung von sozialisiertem Einflusstreben (sE) - theoretische Überlegungen und erste Befunde. Köln : Deutsche Sporthochschule.
- Erdmann, R. (1988). Die Bedeutung empirischer Studien mit kleinen Stichproben für die Theoriebildung im sozialwissenschaftlichen Bereich. *Sportwissenschaft, 18*, 270-283.
- Erdmann, R. (1991). Weniger auf die Methode, sondern auf das Ziel kommt es an. Zur Abgrenzung qualitativer und quantitativer Forschungsansätze. In R. Singer (Ed.), *Sportpsychologische Forschungsmethodik - Grundlagen, Probleme, Ansätze*. (pp. 63-68). Köln: bps.
- Erdmann, R. (1996). Empirical Sport Pedagogy. In P. G. Schempp (Ed.), (pp. 174-202). Münster, New York: Waxmann.
- Erdmann, R. (1999a). Ich bin Ich und Du bist Du. Zur Bedeutung der Identität für die Interkulturelle Bewegungserziehung. In R. Erdmann (Ed.), *Interkulturelle Bewegungserziehung* (pp. 61-80). Sankt Augustin: Academia.
- Erdmann, R. (1999b). *Interkulturelle Bewegungserziehung*. Sankt Augustin: Academia.
- Erdmann, R. (2001). Intercultural Movement Education - Facets of Identity, oral presentation (16th Annual Congress of the European College of Sport Science in Cologne). In Cologne: Sport und Buch Strauss.
- Erdmann, R. (2002). Need to be uncertain? [Unpublished Work]
- Erdmann, R. (2005). Erkenne die Differenz – Gedanken zur Unsicherheit, der Identitätsentwicklung und interkultureller Erziehung. In P. Gieß-Stüber (Ed.), *Interkulturelle Erziehung im und durch Sport* (pp. 48-59). Münster, Berlin, Hamburg: LIT Verlag.
- Erdmann, R. (2008). Leistungen fördern, beurteilen und beraten. In: Lange, H. & S. Sinning (Ed.). *Handbuch: Didaktik des Sportunterrichts* (155-173). Balingen: Spitta-Verlag.



- 
- Erdmann, R. & Schulz, H. G. (1999). Differenz und Bewegung – Fragen an Prof. Dr. Auernheimer. In Erdmann, R. (Ed.), *Interkulturelle Bewegungserziehung* (pp. 8-21). Sankt Augustin: Academia.
- Esser-Noethlichs, M. (2008). Sensitivity towards strangeness: conceptualization, operationalization and practical implications within the field of physical education, oral presentation, European Conference of Educational Research (ECER), Göteborg, 10.09.-12.09.2008.
- Esser-Noethlichs, M. (2010). Learning and improving a swimming technique within an interdisciplinary and experience based teaching approach with two 7th grade school classes. *XIth International Symposium on Biomechanics and Medicine in Swimming*; Norwegian School of Sport Sciences, 16.6.2010 - 19.6.2010.
- Faulenbach, J., Hesse, C., & Klaeren, J. (2001). *Vorurteile - Stereotype - Feindbilder. Informationen zur politischen Bildung*, 271, (2). [Full journal]
- Fischer, T. & Ziegenspeck, J. (2000). *Handbuch Erlebnispädagogik. Von Ursprüngen bis zur Gegenwart*. Bad Heilbrunn: Klinkhardt.
- Fiske, S. T. & Taylor, S. E. (1991). *Social cognition*. New York, London, Toronto: McGraw-Hill.
- Fisseni, J. (1997). *Lehrbuch der psychologischen Diagnostik* (2 ed.). Göttingen, Bern, Toronto: Hogrefe.
- Fossgård, Å. M. F. (2007). *Intercultural competence: evaluation of diaries from a Comenius 2.1 project*. Master thesis in higher education/comparative and international education, University of Oslo.
- Frey, H.P.; Kumpf, M.; Ochsmann, R.; Rost Schaude, E. & Sauer, C. (1977). Theorie der kognitiven Kontrolle. In Tack, W.H. (Ed.), *Bericht über den 30. Kongreß der Deutschen Gesellschaft für Psychologie in Regensburg*, Bd. 1(pp. 105-107). Göttingen: Hogrefe.

- 
- Frey, H.P. & Haußer, K. (1987). Entwicklungslinien sozialwissenschaftlicher Identitätsforschung. In H.P. Frey & K. Hausser (Eds.), *Identität: Entwicklungen psychologischer und soziologischer Forschung* (pp. 3-26). Stuttgart: Enke.
- Fritz, W., Mollenberg, A., & Chen, G. M. (2002). Measuring intercultural sensitivity in different cultural context. *Intercultural Communication Studies*, 11, 165-176.
- Fritz-Gerald, H. (2002). *How Different Are We?* Clevedon, Toronto, Sydney: Multilingual Matters.
- Gadamer, H.G. (1975). *Truth and method*. London: Continuum.
- Gadenne, V. (1994a). Theoriebewertung. In T. Herrmann & W. H. Tack (Eds.), *Methodologische Grundlagen der Psychologie*. (pp. 389-418). Göttingen, Bern, Toronto: Hogrefe.
- Gadenne, V. (1994b). Theorien. In T. Herrmann & W. H. Tack (Eds.), *Methodologische Grundlagen der Psychologie* (pp. 295-342). Hogrefe: Göttingen, Bern Toronto.
- Gawronski, B. & Conrey, F. R. (2004). Der Implizite Assoziationstest als Maß automatisch aktivierter Assoziationen: Reichweite und Grenzen. *Psychologische Rundschau*, 55, 118-126.
- Gerrig, R. J. & Zimbardo, P. G. (2002). *Psychology and Life*. (16 ed.). Boston, London, Toronto: Allyn & Bacon.
- Gerrig, R. J. & Zimbardo, P. G. (2004). *Psychologie*. München, Boston, San Francisco: Pearson.
- Gieß-Stüber, P. (1999). Der Umgang mit Fremdheit – Interkulturelle Bewegungserziehung jenseits von Ausgrenzung oder Vereinnahmung. In R. Erdmann (Ed.), *Interkulturelle Bewegungserziehung* (pp. 42-60). Sankt Augustin: Academia.
- Gieß-Stüber, P. (2000a). *Gleichberechtigte Partizipation im Sport? Ein Beitrag zur geschlechtsbezogenen Sportpädagogik*. Butzbach-Griedel: AFRA.
- Gieß-Stüber, P. (2000b). Sport für alle? Zum Umgang mit Fremdheit im Sport. (Manuskript Antrittsvorlesung zum Erwerb der Venia Legendi "Sportwissenschaft unter besonderer Berücksichtigung der Sportpädagogik"). [ Unpublished Work ]

- 
- Gieß-Stüber, P. (2003a). Fremde und Fremdes erleben. *Sportpädagogik*, 27, 4-8.
- Gieß-Stüber, P. (2003b). Fremdheit erleben. *Sportpädagogik*, 27, (6), [Full journal].
- Gieß-Stüber, P. (2005a). Der Umgang mit Fremdheit - Interkulturelle Bewegungserziehung jenseits von Ausgrenzung oder Vereinnahmung. In P. Gieß-Stüber (Ed.), *Interkulturelle Erziehung im und durch Sport* (pp. 67-75). Münster: LIT.
- Gieß-Stüber, P. (2005b). *Interkulturelle Erziehung im und durch Sport*. Münster, Berlin, Hamburg: LIT-Verlag.
- Gieß-Stüber, P. (2006). Für einen produktiven Umgang mit Fremdheit im Schulsport. In D. Blecking & P. Gieß-Stüber (Eds.), *Sport bewegt Europa. Beiträge zur interkulturellen Verständigung* (pp. 151-168). Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Gieß-Stüber, P. & Blecking, D. (Eds.) (2008). *Sport - Integration - Europe. Widening horizons in intercultural education*. Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Gieß-Stüber, P. (2008). Travelling through Europe. Estrangement and return. In Gieß-Stüber, P. & Blecking, D. (Eds.). *Sport - Integration - Europe. Widening horizons in intercultural education* (pp. 36-50). Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Gieß-Stüber, P. (2010a). Development of intercultural skills through sport and physical education in Europe. In Gasparini, W. & Cometti, A. (Eds.). *Sport facing the test of cultural diversity* (pp. 23-30). Strasbourg: Council of Europe.
- Gieß-Stüber, P. (2010b). Einblicke in die Comenius Praktika. Unpublished report, 6.2.2010.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1461-1480.
- Grimminger, E. (2008). Promoting intercultural competence in the continuing education of physical education teachers. In P. Gieß-Stüber & D. Blecking (Eds.), *Sport - Integra-*

- tion - Europe. Widening horizons in intercultural education* (pp. 304-315).  
Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Grimminger, E. (2009). Interkulturelle Kompetenz im Schulsport. Evaluation eines Fortbildungskonzepts. Erlangen: Schneider Verlag Hohengehren.
- Groeben, N., Schreier Margrit, & Christmann, U. (1993). Fairneß beim Argumentieren: Argumentationsintegrität als Wertkonzept einer Ethik der Kommunikation. *Linguistische Berichte*, 355-382.
- Grupe, O. & Krüger, M. (1996). Sport Pedagogy: Anthropological perspectives and traditions. In P. G. Schmepp (Ed.), *scientific development of sport pedagogy* (pp. 155-173). Münster, New York: Waxmann.
- Grupe, O. & Krüger, M. (2002). *Einführung in die Sportpädagogik*. Schorndorf: Hofmann.
- Gudykunst, W. B. (1987). A model of uncertainty reduction in intercultural encounters. *Journal of language and social psychology*, 4, 79-98.
- Gudykunst, W.B. & Kim, Y.Y. (2003). *Communication with strangers*. New York, London, Sydney: McGraw Hill.
- Guttadin, F. (1993). Die Relevanz des hermeneutischen Verstehens für eine Soziologie des Fremden. In T. Jung & S. Müller-Doohm (Ed.), *„Wirklichkeit“ im Deutungsprozeß* (pp. 458-481). Frankfurt am Main: Fischer.
- Guttman, L. (1950). The basis of scalogram analysis. In Special committee of the Social Science Research Council (Ed.), *Studies in social psychology in World War II* (pp. 60-90). Princeton: Princeton University Press.
- Habermas, J. (2002). "Wann müssen wir tolerant sein? Über die Konkurrenz von Weltbildern, Werten und Theorien" Festvortrag zum Leibniztag der Berlin-Brandenburgischen Akademie der Wissenschaften am 29. Juni 2002.
- Hagtvet, K. A. (2006). Using preconceptions in quantitative analysis of empirical data. In G. Imsen, E. Skaalvik, & A. Oftedal Telhaug (Eds.), *Frontlinjer i pedagogisk forskning. [Frontlines in educational research]*. (pp. 27-56). Trondheim: Norwegian Center for Child Research/ Norwegian Research Council.

- 
- Hahn, A. (1994). Die soziale Konstruktion des Fremden. In W.M. Sprondel (Ed.), *Die Objektivität der Ordnung und ihre Kommunikative Konstruktion*. (pp. 140-163). Frankfurt a. M.: Suhrkamp.
- Hahn, A. (1997). "Partizipative" Identitäten. In H. Münkler & B. Ladwig (Eds.), *Furcht und Faszination: Facetten der Fremdheit*. (pp. 115-159). Berlin: Akademie.
- Hahn, A. (2000). The social construction of the stranger. *Alexander von Humboldt-Mitteilungen*, 75, 11-23.
- Hall, R. E. (2005). From psychology of race to issue of skin color: western trivialization and peoples of African descent. *International Journal of Psychology and Psychological Therapy*, 5, 125-134.
- Hammer, M. R., Bennet, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations*, 27, 421-443.
- Haußer, K. (1997). Identitätsentwicklung - vom Phasenuniversalismus zur Erfahrungsverarbeitung. In H. Keupp & R. Höfer (Eds.), *Identitätsarbeit heute* (pp. 120-134). Frankfurt am Main: Suhrkamp.
- Haußer, K. (1995). *Identitätspsychologie*. Berlin, Heidelberg, New York: Springer.
- Haußer, K. (2007). Assessing identity as the meaningfulness of meanings – the Flensburg identity status interview (FISI). In Watzlawik, M. & Born, A. (Eds.), *Capturing identity: quantitative and qualitative methods* (pp. 15-22). Lanham: University Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hirschfeld, L. A. (1993). Discovering Social Differences: The Role of Appearance in the Development of Racial Awareness. *Cognitive Psychology*, 25, 317-350.
- Holzbrecher, A. (1997). *Wahrnehmung des Anderen. Zur Didaktik interkulturellen Lernens*. Leske + Budrich: Opladen.
- Holzbrecher, A. (2004). *Interkulturelle Pädagogik*. Berlin: Cornelsen.

- 
- Hotz, A. (1997). *Qualitatives Bewegungslernen*. Bern: SVSS.
- Keupp, H. (1997). Diskursarena Identität: Lernprozesse in der Identitätsforschung. In H. Keupp & R. Höfer (Eds.), *Identitätsarbeit heute* (pp. 11-39). Frankfurt/M: Suhrkamp.
- Keupp, H.; Ahbe, T.; Gmür, W.; Höfer, R.; Mitzscherlich, B.; Kraus, W. & Strauss, F. (2008). *Identitätskonstruktionen. Das Patchwork der Identitäten in der Spätmoderne*. Hamburg: Rowohlt.
- Kleven, T. A. (2008). Validity and validation in qualitative and quantitative research. *Nordisk pedagogikk*, (3), 219-233.
- Krampen, G. (1981). *IPC-Fragebogen zu Kontrollüberzeugungen (locus of control)*. Göttingen, Bern, Toronto: Hogrefe.
- Kriz, J., Lück, H. E., & Heidbrink, H. (1996). *Wissenschafts- und Erkenntnistheorie. Eine Einführung für Psychologen und Humanwissenschaftler* (3 ed.). Opladen: Leske + Budrich.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Analyse*. (6 ed.) Weinheim: Belz.
- McDonald, R. P. (1999). *Test theory: a unified treatment*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.
- Meinberg, E. (1996). *Hauptprobleme der Sportpädagogik. Eine Einführung*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Merleau-Ponty, M. (1996). *Kroppens fenomenologi*. Oslo: Pax forlag.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50, 741-749.
- Michels, H.; Schulz, H.G. (1999). Sportorganisation und interkulturelle Begegnung – eine kommentierte Befragung. In R. Erdmann (Ed.), *Interkulturelle Bewegungserziehung* (pp. 130-152). St. Augustin: Academia.
- Midthaugen, P. (in preparation.). Handling differences and intercultural learning in Physical Education. Measuring outcomes among students after a PE-teacher training interven-

tion in upper secondary schools in Norway. PhD project at the Norwegian School of Sport Sciences in Oslo.

Miriam-Webster (2009). *Miriam-Webster Online Search*. Miriam-Webster, Incorporated, [online dictionary], <http://www.merriam-webster.com/netdict.htm>

Nakamura, Y. (2000). *Xenosophie. Bausteine für eine Theorie der Fremdheit*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Nieke, W. (1995). *Interkulturelle Erziehung und Bildung. Wertorientierungen im Alltag*. Opladen: Leske + Budrich.

Noethlichs, M. (2001). Interkulturelle Bewegungserziehung - Ansatzpunkte für eine theoriegeleitete Praxis. In Roll, A.; Brinks, P.; Francescon, G.; Ribler, A. (Ed.), *Sport gegen Rassismus - Rassismus im Sport!?* (pp. 13-18). Frankfurt am Main: Sportjugend Hessen.

Noethlichs, M. (2003). Integrieren im Sport – spielend kämpfen. *Sportpädagogik*, 27, 16-19.

Noethlichs, M. (2005a). Die Sensitivität gegenüber Fremdheit (STS) und Interkulturelles Lernen im Sport. In P. Gieß-Stüber (Ed.), *Interkulturelle Erziehung im und durch Sport* (pp. 38-47). Münster: LIT.

Noethlichs, M. (2005b). Judo, ein konstruktiver Weg im Untergang mit Fremdheit. In P. Gieß-Stüber (Ed.), *Interkulturelle Erziehung im und durch Sport* (pp. 76-81). Münster: LIT.

Noethlichs, M. & Schulz, A. (2007). Physik und Schwimmen (gleichzeitig) lernen - ein Beispiel für eine fächerverbindende Unterrichtsreihe. *Praxis der Naturwissenschaften: Physik in der Schule*, 56, 35-43.

Neuber, N. (1999). Vom Gewohnten zum Ungewöhnlichen – Ansatzpunkte für eine Praxis der interkulturellen Bewegungserziehung. In R. Erdmann (Ed.), *Interkulturelle Bewegungserziehung* (pp. 112-129). St. Augustin: Academia.

Nosek, B., Banaji, M., & Greenwald, T. (2005). Implicit Association Test (IAT), online demo version [Computer software], <https://implicit.harvard.edu/implicit/demo/index.jsp>.

- Osburn, H.G. (2000). Coefficient Alpha and Related Internal Consistency Reliability Coefficients, *Psychological Methods*, 5 (3), 343-355.
- Oxford Dictionary (2009). *Oxford Advanced Learner's Dictionary* [online]. Oxford: Oxford University Press, <http://www.oup.com/elt/catalogue/teachersites/oald7/?cc=no>.
- Paige, R. M. (1993). On the nature of intercultural experiences and intercultural education. In R. M. Paige (Ed.), *Education for the intercultural experience* (pp. 1-20). Yarmouth, Maine: Intercultural Press.
- Passer, M. W. & Smith, R. E. (2004). *Psychology. The science of mind and behavior*. (2 ed.) New York, London, Toronto: Mc Graw Hill.
- Pihl, J. (2000). Interkulturell og flerkulturell undervisning. *Norsk pedagogisk tidsskrift*,(4/5), 308-319.
- Popper, K. R. (2002). *The logic of scientific discovery*. London and New York: Routledge classics.
- Porter, J. D. R. (1973). *Black child, white child. The development of racial attitudes*. Cambridge, Massachusetts: Harvard University Press.
- Raaheim, K. & Brun, W. (1985). Task novelty and intelligence. *Scandinavian Journal of Psychology*, 26, 35-41.
- Rehm, J. & Strack, F. (1994). Kontrolltechniken. In T. Herrmann & W. H. Tack (Eds.), *Methodologische Grundlagen der Psychologie* (pp. 508-555). Göttingen, Bern, Toronto: Hogrefe.
- Reuter, J. (2002). *Ordnungen des Anderen. Zum Problem des Eigenen in der Soziologie des Fremden*. Bielefeld: transcript.
- Reviere, U. (1998). Ansätze und Ziele interkulturellen Lernens in der Schule. Ein Leitfaden für die Sekundarstufe. Frankfurt am Main: IKO.
- Rommelspacher, B. (1997). Zur Internalisierung von Diskriminierung und Dominanz. In R. Höfer & H. Keupp (Eds.), *Identitätsarbeit heute. Klassische und aktuelle Perspektiven der Identitätsforschung* (pp. 251-269). Frankfurt am Main: Suhrkamp.



- 
- Rommelspacher, B. (1998). *Dominanzkultur – Texte zu Fremdheit*. (2 ed.) Berlin: Orlanda.
- Rost, J. (2004). *Lehrbuch der Testtheorie, Testkonstruktion*. Bern, Göttingen, Toronto, Seattle: Huber.
- Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement, *Psychological Monographs*, 80, 1-28.
- Schäffter, O. (1991). Modi des Fremderlebens. Deutungsmuster im Umgang mit Fremdheit. In: Schäffter, O. (Ed.), *Das Fremde. Erfahrungsmöglichkeiten zwischen Faszination und Bedrohung*. Opladen: Westdeutscher Verlag, 11-42.
- Schütz, A. (1974). Grundzüge einer Theorie des Fremdverstehens. In: Schütz, A., *Der sinnhafte Aufbau der sozialen Welt. Eine Einleitung in die Verstehende Soziologie*. Frankfurt am Main: Suhrkamp, 137-197.
- Schwarzer, R. (2000). *Streß, Angst und Handlungsregulation*. Stuttgart, Berlin, Köln: Kohlhammer.
- Selter, R. (1996). *Du bist wie dein Sport: Studien zur Entwicklung von Selbstkonzept und Identität*. Schorndorf: Hofmann.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Simmel, G. (1968). *Soziologie*. Berlin.
- Simmel, G. (1992). Exkurs über den Fremden. In A. Loycke (Ed.), *Der Gast, der beleibt. Dimensionen von Georg Simmels Analyse des Fremdseins*. (pp. 9-16). Frankfurt, New York: Campus Verlag.
- Sonnenschein, W. (1999). Assimilation versus Ethnizität Sport und gesellschaftliche Integration ausländischer Mitbürger. In R. Erdmann (Ed.), *Interkulturelle Bewegungserziehung* (pp. 81-92). Sankt Augustin: Academia.
- Steyer, R. & Eid, M. (2001). *Messen und Testen*. Berlin, Heidelberg, New York: Springer.

- 
- Strauss, F. and Höfer R. (1997). Entwicklungslinien alltäglicher Identitätsarbeit. In Keupp, H. und R. Höfer (Eds.), *Identitätsarbeit heute. Klassische und aktuelle Perspektiven der Identitätsforschung* (pp. 270-307). Frankfurt am Main: Suhrkamp.
- Taylor, C. (1993). *Multikulturalismus und die Politik der Anerkennung*. Frankfurt am Main: Fischer.
- Thiele, J. (1999). Bewegungskulturen im Widerstreit - ein Beitrag zur Begrenzung des Verstehens. In R. Erdmann (Ed.), *Interkulturelle Bewegungserziehung* (pp. 22-41). Sankt Augustin: Academia.
- Thomas, A. (1991). Soziale Wahrnehmung und Kognition. In A. Thomas (Ed.), *Grundriß der Sozialpsychologie. Bd. 1 (Grundlegende Begriffe und Prozesse)* (pp. 163-183). Göttingen, Toronto, Zürich: Hogrefe.
- Thomas, A. (1992). *Grundriß der Sozialpsychologie, Bd.2( Individuum, Gruppe, Gesellschaft)*. Göttingen, Bern, Toronto, Seattle: Hogrefe.
- Valentin, K. (1984). Die Fremden. In M. Schulte (Ed.), *Das Valentin-Buch. Von und über Karl Valentin in Texten und Bildern*. (pp. 488-489). München.
- Van de Vliert, E. & Kabanoff, B. (1990). Toward Theory-Based Measures of Conflict Management, *The Academy of Management Journal*, 33 (1), 199-209
- Van der Zee, K.I. & Van Oudenhoven J.P. (2000). The multicultural personality questionnaire: a multidimensional instrument of multicultural effectiveness, *European Journal of Personality*, 14, 291-309.
- Van der Zee, K.I. & Van Oudenhoven J.P. (2001). The multicultural personality questionnaire: reliability and validity of Self- and other ratings of multicultural effectiveness, *Journal of Research in Personality*, 35, 278-288.
- Voss, A. (2003). *Geschlecht im Sport - sozialkonstruktivistische Lesearten*. Schorndorf: Hofmann.
- Waldenfels, B. (1997). *Topographie des Fremden. Studien zur Phänomenologie des Fremden 1*. Frankfurt am Main: Suhrkamp.

- 
- Waldenfels, B. (2006). *Grundmotive einer Phänomenologie des Fremden*. Frankfurt am Main: Suhrkamp.
- Walseth, K. (2004). En problematisering av idretten som arena for integrasjon av etniske minoriteter. *www.idrottsforum.org*, 27.1.2004.
- Walseth, K. and Fasting, K. (2004). Sport as a means of integrating minority women, *Sport in society*, 7 (1), 109-129.
- Walseth, K. (2006). *Sport and integration - the experience of young Muslim women*. Oslo: The Norwegian School of Sport Sciences.
- Weller, I. & Matiaske, W. (2003). Gütekriterien und faktorielle Struktur des IMC-Gitters zur Messung von Leistungs-, Macht- und Anschlussmotiven. *Berichte der Werkstatt für Organisations- und Personalforschung e.V.* Berlin.
- Welsch, W. (1999). Transculturality - the Puzzling Form of Cultures Today, In: Featherstone, M.; Lash, S. (Ed.), *Spaces of Culture: City, Nation, World*. London: Sage, 194-213.
- Wood, Margret (1934). *The stranger. A study in social relationships*. New York: Columbia University.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik. Ein Lehrbuch zur psychologischen Methodenlehre*. Göttingen, Bern, Toronto: Hogrefe.

