

## Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>1</b>
<b>Part I - (EFL) Teacher Education in Germany: From Developing Competence to Assessing Performance</b>		
<b>2</b>	<b>Structural Elements of Teacher Education in Germany.....</b>	<b>7</b>
2.1	The Goal of Teacher Education .....	7
2.2	The Structure of (EFL) Teacher Education in Germany .....	10
<b>3</b>	<b>Competence of EFL Teachers .....</b>	<b>12</b>
3.1	Approaching the Terminology of Competence .....	13
3.2	Models of Competence in Teacher Education.....	17
3.2.1	Model of Professional Competence .....	18
3.2.2	Competence as Knowledge .....	23
3.2.3	Competence as a Continuum.....	25
3.3	Frameworks for (EFL) Teacher Education in Germany .....	28
3.4	Empirical Findings on Professional Knowledge of Pre-Service EFL Teachers.....	33
3.5	Competence between Knowledge and Practice .....	37
<b>4</b>	<b>A Performance-Oriented Perspective on Constructive Alignment in the First Phase of EFL Teacher Education.....</b>	<b>40</b>
4.1	Constructive Alignment .....	41
4.2	Learning Objectives: The Core Practice Approach in EFL Teacher Education ....	43
4.2.1	The Core Practice Approach .....	43
4.2.2	Core Practices in the EFL Context .....	45
4.3	Teaching and Learning Activities: Simulation-based Learning.....	47
4.3.1	Characteristics of Simulations .....	49
4.3.2	Typology of Simulations.....	51
4.3.3	Perception of Simulations and their Effectiveness for Learning .....	55
4.4	Assessment: RobS in EFL Teacher Education .....	57
4.4.1	Concepts of Assessment in Higher Education .....	57
4.4.2	Assessment in Higher Education: Students' Perceptions and Impact on Learning ..	60
4.4.3	Simulation-based Assessment in Higher and Teacher Education.....	63
4.5	Addressing the Assessment Gap in Teacher Education .....	66
<b>5</b>	<b>Summary .....</b>	<b>68</b>
<b>Part II - Developing a Role-Play-Based Simulation (RobS) to Assess Feedback Competence on Writing in the EFL context</b>		
<b>6</b>	<b>The Relevance of Feedback on Writing in the EFL context.....</b>	<b>71</b>
6.1	Feedback .....	71
6.1.1	On the History and Terminology of Feedback.....	72
6.1.2	Feedback as Literacies.....	75
6.1.3	Relevance of Feedback in (EFL) Teacher Education .....	76
6.2	Writing in the EFL classroom .....	79
6.2.1	Cognitive Process Models for (L1-) Writing.....	80

6.2.2 Approaches to Writing in the EFL classroom .....	83
6.2.3 Statutory Guidelines for Learning and Teaching Writing .....	87
6.3 Synthesis: Feedback on Writing .....	89
<b>7 Developing a Framework for Feedback Competence on Writing in the EFL context.....</b>	<b>93</b>
7.1 Methodological Reflections on Theoretical Research.....	94
7.2 Instruments for the Evaluation of Feedback Performance .....	95
7.3 The Framework.....	97
7.3.1 Functions of Feedback.....	98
7.3.1.1 Aim .....	98
7.3.1.2 Level.....	101
7.3.2 Strategies of Feedback.....	103
7.3.2.1 Communication .....	104
7.3.2.2 L2/FL Writing Pedagogy.....	107
7.3.3 Topics of Feedback.....	112
7.3.3.1 Diagnosis.....	112
7.3.3.2 Information .....	114
7.4 Discussion and Limitations of the Framework.....	116
<b>8 A RobS to Assess Feedback Competence on Writing in the EFL context ....</b>	<b>119</b>
8.1 Preparation .....	119
8.2 Simulation .....	125
8.2.1 The Role of Mia.....	126
8.2.2 Standardization of the Learner .....	127
8.3 Assessment.....	130
8.4 Discussion and Limitations of the RobS .....	140
<b>9 Summary .....</b>	<b>142</b>
<b>Part III - Methodology: Investigating the Validity of the RobS</b>	
<b>10 Criteria for Test Evaluation .....</b>	<b>145</b>
10.1 Overview of Test Criteria .....	145
10.1.1 Secondary Criteria: Practicality, Fairness and Authenticity.....	146
10.1.2 Primary Criteria: Objectivity and Reliability.....	147
10.2 From Validity to Validation .....	148
10.2.1 Traditional Approaches to Validity.....	149
10.2.2 Messick's Approach to Construct Validity .....	151
10.2.3 The Argument-based Approach to Validation.....	156
10.3 Implications for the Validation Process.....	160
<b>11 IUA for the RobS.....</b>	<b>161</b>
11.1 A Classification System for Validation by Schreiber & Gut (2022) .....	161
11.2 Making the Claims: The IUA and Sources of Evidence for the Test .....	164
11.2.1 Claim 1: Model.....	165
11.2.2 Claim 2: Test.....	165
11.2.3 Claim 3: Performance .....	169
11.2.4 Claim 4: Item score.....	169

11.2.5 Claim 5: Test score.....	170
11.2.6 Claim 6: Latent trait.....	170
11.2.7 Claim 7: Interpreted trait.....	171
11.2.8 Claim 8: Use.....	171
<b>12 Summary .....</b>	<b>172</b>
<b>Part IV - Empirical Findings: Results of the Investigations</b>	
<b>13 Research Design and Overview.....</b>	<b>173</b>
<b>14 (Pre-)Pilot Phase and Survey: Actor Training and Material Development .</b>	<b>177</b>
14.1 Pre-Pilot Phase .....	177
14.1.1 Creating and Evaluating Authentic Material.....	177
14.1.2 Actor Training.....	185
14.2 Pilot Survey.....	187
14.2.1 Data Collection and Sample Characteristics .....	188
14.2.1.1 Data Collection.....	188
14.2.1.2 Sample Characteristics .....	193
14.2.2 Investigating the Verbal Triggers.....	194
14.2.2.1 Method: Stimulated Recall Interviews .....	195
14.2.2.2 Results of the SRIs.....	197
14.2.2.3 Discussion and Limitations .....	203
14.2.2.4 Conclusion .....	205
14.2.3 Deciding on the Test Documents .....	206
14.2.4 Deciding on the Research Procedure .....	209
14.3 Summary .....	210
<b>15 Main Survey .....</b>	<b>211</b>
15.1 Data Collection and Sample Characteristics.....	211
15.1.1 Data Collection.....	211
15.1.2 Sample Characteristics .....	213
15.2 Study 1: Pre-service EFL Teachers' Perspective on RobS as Assessment.....	216
15.2.1 Methods .....	217
15.2.2 Results .....	220
15.2.2.1 Authenticity of the RobS .....	220
15.2.2.2 Competences Tested by the RobS.....	223
15.2.2.3 Fairness of the RobS as Assessment .....	226
15.2.2.4 Use of the RobS for Summative Assessment.....	228
15.2.3 Discussion and Limitations.....	231
15.2.4 Summary .....	234
15.3 Study 2: Expert Interviews .....	235
15.3.1 Method: Expert Interviews.....	236
15.3.2 Study Design and Data Processing .....	238
15.3.3 Results of the Qualitative Expert Interviews.....	243
15.3.3.1 Experts' Evaluative Comments on the Performances .....	243
15.3.3.2 Evaluation of the Scoring Instrument.....	248
15.3.4 Discussion of the Results and Modifications to the Scoring Instrument.....	266
15.3.5 Limitations .....	270
15.3.6 Summary .....	270

15.4 Study 3: Scoring Reliability .....	272
15.4.1 Scoring Reliability on the Item Level .....	272
15.4.1.1 Rating of the Full Sample on the Item Level .....	273
15.4.1.1.1 Methods.....	273
15.4.1.1.2 Results .....	275
15.4.1.2 Inter-Rater Reliability on the Item Level .....	282
15.4.1.2.1 Methods.....	283
15.4.1.2.2 Results of Rating I on the Item Level .....	287
15.4.1.2.3 Modifications .....	291
15.4.1.2.4 Results of Rating II on the Item Level.....	298
15.4.1.3 Discussion of Items for the Final Scale.....	302
15.4.1.3.1 Exclusion of Items .....	302
15.4.1.3.2 Inclusion of Items Despite Scoring Frequency.....	304
15.4.1.3.3 Inclusion of Items Despite Insufficient Corrected Item-Total Correlation.....	305
15.4.1.3.4 Inclusion of Items Despite Potentially Insufficient Inter-Rater Reliability .....	306
15.4.1.4 Revised Reliability Values on the Item Level.....	310
15.4.2 Scoring Reliability on the Test Score Level.....	311
15.4.2.1 Descriptive Statistics and Tests on Normal Distribution for the Full Sample .....	311
15.4.2.2 Inter-Rater Reliability on the Test Score Level .....	315
15.4.2.2.1 Methods.....	315
15.4.2.2.2 Results of Rating I on the Test Score Level .....	315
15.4.2.2.3 Results of Rating II on the Test Score Level .....	320
15.4.2.3 Discussion of the Test Score-Focused Reliability .....	325
15.4.3 Limitations and Implications from the Reliability Investigations .....	327
15.4.4 Scoring Independence from Actors .....	329
15.4.4.1 Results .....	329
15.4.4.2 Discussion.....	330
15.4.5 Summary .....	330
15.5 Study 4: The Relationship between Test Score and Other Variables .....	331
15.5.1 Relevant Criteria .....	332
15.5.2 Hypotheses.....	335
15.5.3 Methods .....	336
15.5.3.1 Personal and Academic Data .....	336
15.5.3.2 Areas of the target domain .....	337
15.5.3.2.1 Specific PK: Declarative Feedback Knowledge.....	337
15.5.3.2.2 EFL-specific PCK: FALKO-E.....	339
15.5.3.2.3 CK / LP: The LexTale Test.....	344
15.5.3.3 Expert Ratings.....	346
15.5.4 Results .....	347
15.5.4.1 H1: School-related Experience .....	347
15.5.4.2 H2: Academic Progress and Success.....	349
15.5.4.3 H3: Facets of EFL Teacher Competence.....	352
15.5.4.3.1 Descriptive Statistics, Normal Distribution and General Inferential Statistics .....	352
15.5.4.3.2 Inferential Statistics .....	354
15.5.4.4 H4: Expert Ratings.....	357

15.5.5 Discussion .....	360
15.5.6 Limitations .....	366
15.5.7 Summary .....	368
<b>16 Summary .....</b>	<b>369</b>
<b>Part V – Discussion: Evaluating the Validity of the RobS</b>	
<b>17 Evaluating the Claims: The Validity Argument .....</b>	<b>375</b>
17.1 Claim 1: Model .....	375
17.2 Claim 2: Test .....	377
17.3 Claim 3: Performance .....	379
17.4 Claim 4: Item Score .....	381
17.5 Claim 5: Test Score .....	383
17.6 Claim 6: Latent Trait .....	384
17.7 Claim 7: Interpreted Trait .....	386
17.8 Claim 8: Use .....	388
17.9 Conclusion: To use, or not to use? .....	389
<b>18 Looking Back – Looking Forward .....</b>	<b>390</b>
18.1 Summary .....	390
18.2 Limitations .....	393
18.3 Implications and Future Perspectives .....	397
<b>19 References .....</b>	<b>401</b>
<b>List of Abbreviations .....</b>	<b>448</b>
<b>List of Figures .....</b>	<b>452</b>
<b>List of Tables .....</b>	<b>454</b>
<b>Appendix</b>	
<b>A. Test Material .....</b>	<b>458</b>
A.1. M0: Vignette .....	458
A.2. M1: Learner text .....	459
A.3. M2: The Task for the Text .....	460
A.4. M3: Excerpt from the Series of Lessons .....	461
A.5. M4: Sample Text from the Series of Lessons .....	462
A.6. M5: Empty Sheet for Notes .....	463
A.7. Role description of Mia (German) .....	464
A.8. Text Sample A from the Pilot Survey .....	466
<b>B. Rating manual .....</b>	<b>467</b>
<b>C. Demographic and Academic Data Collected in the Written Survey .....</b>	<b>524</b>
<b>D. Interview Guidelines .....</b>	<b>526</b>
D.1. Pre-pilot Phase: Expert Interviews .....	526
D.2. Pre-pilot Phase: PSETs Interviews .....	526
D.3. Pilot Survey .....	528
D.4. Main Survey .....	530
D.5. Expert Interview .....	531
<b>E. Transcription Guidelines .....</b>	<b>541</b>

E.1. Follow-Up Interviews and Simulations (Main Survey).....	541
E.2. Stakeholder and Expert Interviews.....	542
<b>F. Coding Manuals.....</b>	<b>543</b>
F.1. Stimulated Recall .....	543
F.2. Guideline Interview .....	549
F.3. Expert Interviews.....	608
<b>G. Coding Samples .....</b>	<b>621</b>
G.1. Stimulated Recall .....	621
G.2. Simulation .....	639
G.3. Guideline Interview– Main Survey .....	659
G.4. Expert Interview – Part 1 .....	667
G.5. Expert Interview – Part 2 .....	671
<b>H. Statistical Information.....</b>	<b>686</b>
H.1. Intercoding Sub-Codes PSETs Guideline Interview .....	686
H.2. Discrepancies in Inter-Rating I.....	687
H.3. Inter-Item Correlation Matrix for 27 Items.....	690
H.4. Test Score Frequencies.....	693
H.5. Normal Distribution in the Actor-specific Groups.....	694
H.6. Further Statistical Information on SPK, PCK and CK /LP .....	694
<b>I. Digital Supplement.....</b>	<b>696</b>