

Inhalt

1	Einleitung.....	11
2	Der Ko-WADiS-Test.....	15
2.1	Kompetenzbegriff.....	15
2.2	Naturwissenschaftliches Denken	16
2.3	Vorstellung des Testformats	21
2.4	Ausgewählte Ergebnisse im Fach Physik	24
3	Item Response Theory	27
3.1	Testtheorien	28
3.2	Grundannahmen der IRT	31
3.3	Modelle der IRT	32
3.3.1	Das einparametrische logistische Modell.....	33
3.3.2	Das zweiparametrische logistische Modell.....	38
3.3.3	Das dreiparametrische logistische Modell	41
3.3.4	Exkurs: Raschmodell und 1pl-Modell	44
3.4	Parameterschätzung.....	45
3.4.1	Maximum-Likelihood-Verfahren	46
3.4.2	Bayessche Schätzverfahren.....	49
3.4.3	Schätzverfahren im Vergleich	51
4	Adaptive Testverfahren.....	55
4.1	Vergleich papier- und computerbasierter Tests.....	57
4.2	Computeradaptive Tests	60
4.2.1	Testalgorithmus: Itemauswahl während des Tests	61
4.2.2	Testalgorithmus: Itemauswahl vor dem Test	63
4.2.3	Testalgorithmus: Abschlusskriterien	65
4.3	Multistage-Tests.....	67
4.3.1	Testlänge und Stufenzahl	68

4.3.2	Anzahl der Module	70
4.3.3	Scoring und Routing	71
4.4	Vergleiche zwischen linearen und adaptiven Formaten	73
4.4.1	Effizienz	73
4.4.2	Testmotivation	74
5	Problemstellung und Forschungsfragen	75
5.1	Problemstellung	75
5.2	Forschungsfragen	78
6	Auswahl des IRT-Modells (FF 1)	83
6.1	Datenbereinigung	84
6.2	Methodik – Individueller Modellfit	86
6.2.1	Infit und Outfit	87
6.2.2	RMSD	89
6.2.3	Anwendung der Kriterien	91
6.3	Methodik – Modellvergleich	92
6.3.1	AIC und BIC	92
6.3.2	Testinformation und Reliabilität	94
6.4	Ergebnisse	95
6.4.1	Modellfit 1pl	96
6.4.2	Modellfit 2pl	98
6.4.3	Modellfit 3pl	99
6.4.4	Modellvergleich	100
6.5	Beschreibung des Itempools	102
7	Testkonstruktion (FF 2)	105
7.1	Auswahl des Testverfahrens	105
7.2	Regeln für Modul- und Strukturaufbau	109
7.3	Regeln für Scoring und Routing	114

7.4	Auswahl der Teststruktur durch Simulationsstudien	116
7.4.1	Methodik	118
7.4.2	Ergebnisse	122
7.4.3	Diskussion	128
8	Testpilotierung und Vergleichsstudie (FF 3)	133
8.1	Praktische Umsetzung	133
8.2	Pilotierungsstudie	140
8.2.1	Proband*innenauswahl	141
8.2.2	Methodik	142
8.2.3	Ergebnisse und Diskussion	142
8.3	Vergleich zwischen MST und linearem Format	145
9	Zusammenfassung und Gesamtdiskussion	147
9.1	Modellauswahl	147
9.2	Testkonstruktion und Simulationsstudien	150
9.3	Pilotierung	152
9.4	Ausblick	154
	Literaturverzeichnis	159
	Anhang A – Itemparameter und Fitwerte 1pl	173
	Anhang B – Itemparameter und Fitwerte 2pl	179
	Danksagung	185