

Gütekriterien diktionsbasierter
Textanalysen zur Erfassung
domänenspezifischer Kommunikationsinhalte

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Fakultät für Psychologie, Pädagogik und Sportwissenschaft
der Universität Regensburg

vorgelegt von

Michael Heilemann

Straubing

2016

Regensburg 2016

Einleitung	7
I. Theoretischer Hintergrund	11
1. Diktionärsbasierte Textanalysen	11
1.1 Funktionsweise diktionärsbasierter Textanalysen	11
1.2 Vor- und Nachteile diktionärsbasierter Textanalysen.....	17
2. Bedeutung diktionärsbasierter Textanalysen für die erziehungswissenschaftliche Forschung.....	21
2.1 Einordnung diktionärsbasierter Textanalysen.....	21
2.2 Theoretische Annahmen diktionärsbasierter Textanalysen	25
2.3 Bedeutung domänenspezifischer Inhaltswörter	29
3. Gütekriterien diktionärsbasierter Textanalysen.....	32
3.1 Objektivität	33
3.1.1 Durchführungsobjektivität	33
3.1.2 Auswertungsobjektivität.....	34
3.1.3 Interpretationsobjektivität	35
3.2 Reliabilität.....	37
3.2.1 Paralleltest-Reliabilität.....	37
3.2.2 Split-Half-Reliabilität.....	38
3.2.3 Retest-Reliabilität.....	39
3.3 Validität	41
3.3.1 Inhaltsvalidität.....	41
3.3.2 Konstruktvalidität.....	42
3.3.3 Kriteriumsvalidität	44
II. Ziele und Fragestellungen	47
1. Entwicklung eines Wörterbuchs zur Erfassung von Kommunikation über MINT-Inhalte.....	51
2. Überprüfung der Gütekriterien diktionärsbasierter Textanalysen am Beispiel des erstellten MINT-Wörterbuchs	51
III. Empirische Studien	57
1. Untersuchungssetting CyberMentor.....	57
1.1 Hintergrund und Beschreibung von CyberMentor	57
1.2 Studienteilnehmerinnen und Textmaterial.....	62
2. Studie 1: Erstellung eines MINT-Wörterbuchs und Überprüfung seiner Inhaltsvalidität	65
2.1 Forschungsanliegen der Studie	65
2.2 Vorgehensweise bei der Erstellung des Wörterbuchs.....	66
2.3 Überprüfung der Inhaltsvalidität.....	67
2.4 Zusammenfassung und Diskussion von Studie 1	69
3. Studien 2 und 3: Zur Objektivität diktionärsbasierter Textanalysen.....	71

3.1 Studie 2: Überprüfung der Auswertungsobjektivität des MINT-Wörterbuchs	71
3.1.1 Forschungsanliegen der Studie	71
3.1.2 Methode	72
3.1.3 Ergebnisse	75
3.1.4 Zusammenfassung und Diskussion von Studie 2.....	78
3.2 Studie 3: Erhöhung der Interpretationsobjektivität des MINT-Wörterbuchs	82
3.2.1 Forschungsanliegen der Studie	82
3.2.2 Methode	82
3.2.3 Ergebnisse	85
3.2.4 Zusammenfassung und Diskussion von Studie 3.....	101
4. Studien 4 bis 6: Zur Reliabilität diktionsärsbasierter Textanalysen	104
4.1 Studie 4: Überprüfung der Paralleltest-Reliabilität des MINT-Wörterbuchs	104
4.1.1 Forschungsanliegen der Studie	104
4.1.2 Methode	106
4.1.3 Ergebnisse	107
4.1.4 Zusammenfassung und Diskussion von Studie 4.....	115
4.2 Studie 5: Überprüfung der Split-Half-Reliabilität des MINT-Wörterbuchs	117
4.2.1 Forschungsanliegen der Studie	117
4.2.2 Methode	118
4.2.3 Ergebnisse	119
4.2.4 Zusammenfassung und Diskussion von Studie 5.....	123
4.3 Studie 6: Überprüfung der Retest-Reliabilität des MINT-Wörterbuchs	126
4.3.1 Forschungsanliegen der Studie	126
4.3.2 Methode	127
4.3.3 Ergebnisse	128
4.3.4 Zusammenfassung und Diskussion von Studie 6.....	133
5. Studien 7a bis 8c: Zur Validität diktionsärsbasierter Textanalysen	135
5.1 Studie 7: Überprüfung der Konstruktvalidität des MINT-Wörterbuchs	136
Studie 7a: Vergleich des MINT-Wörterbuchs mit manuellen Kodierungen	136
5.1.1 Forschungsanliegen der Studie	136
5.1.2 Methode	137
5.1.3 Ergebnisse	138
5.1.4 Zusammenfassung und Diskussion von Studie 7a.....	140
Studie 7b: Vergleich des MINT-Wörterbuchs mit Selbstberichten	141
5.1.5 Forschungsanliegen der Studie	141
5.1.6 Methode	142

5.1.7 Ergebnisse	143
5.1.8 Zusammenfassung und Diskussion von Studie 7b	146
Studie 7c: Vergleich von Texten mit durchschnittlicher bzw. extremer Merkmalsausprägung	148
5.1.9 Forschungsanliegen der Studie.....	148
5.1.10 Methode.....	149
5.1.11 Ergebnisse	149
5.1.12 Zusammenfassung und Diskussion von Studie 7c	150
5.2 Studie 8: Überprüfung der Kriteriumsvalidität des MINT-Wörterbuchs	151
Studie 8a: Zusammenhang der MINT-Kommunikation mit selbstberichteten Mentoring-Outcomes	151
5.2.1 Forschungsanliegen der Studie.....	151
5.2.2 Methode.....	153
5.2.3 Ergebnisse	157
5.2.4 Zusammenfassung und Diskussion von Studie 8a	163
Studie 8b: Zusammenhang der MINT-Kommunikation mit langfristiger Teilnahme am Mentoring-Programm.....	165
5.2.5 Forschungsanliegen der Studie.....	165
5.2.6 Methode.....	165
5.2.7 Ergebnisse	166
5.2.8 Zusammenfassung und Diskussion von Studie 8b.....	170
Studie 8c: Zusammenhang der MINT-Kommunikation mit tatsächlichen Wahlentscheidungen in MINT.....	171
5.2.9 Forschungsanliegen der Studie.....	171
5.2.10 Methode.....	172
5.2.11 Ergebnisse	174
5.2.12 Zusammenfassung und Diskussion von Studie 8c	178
IV. Diskussion	181
1. Zusammenfassung und Diskussion der Hauptbefunde.....	181
2. Diktionsbasierte Textanalysen in der erziehungswissenschaftlichen Forschungspraxis	203
3. Limitationen	206
4. Forschungsausblick und Fazit	207
Literatur.....	211
Anhang	229
A Abbildungsverzeichnis	229
B Tabellenverzeichnis	231
C MINT-Wörterbuch.....	235

B Tabellenverzeichnis

Tabelle 1: Übersicht über die deutsche Version der LIWC-Wörterbücher mit Beispiel-Wörtern und Wörterbuch-Umfang (vgl. Pennebaker et al., 2001; Wolf et al., 2008)	15
Tabelle 2: Überblick über den Fokus der empirischen Studien bezüglich der Gütekriterien diktionsbasierter Textanalysen	50
Tabelle 3: Textbeiträge von Schülerinnen im CyberMentor-Diskussionsforum.....	61
Tabelle 4: Komponenten des MINT-Wörterbuchs	69
Tabelle 5: Überprüfung der Messwertdifferenzen zwischen korrigierten und unkorrigierten Textbeiträgen der Schülerinnen ($N = 69$) auf Normalverteilung ..	75
Tabelle 6: Mittelwerte, Standardabweichungen, Differenzwerte, Mittelwertsunterschiede und Korrelationskoeffizienten für die korrigierten und unkorrigierten Textbeiträge der Schülerinnen ($N = 69$).....	76
Tabelle 7: Die 10 (MINT-)Wörter mit den jeweils größten Häufigkeitsdifferenzen zwischen korrigierten und unkorrigierten Textkorpora	78
Tabelle 8: Überprüfung des MINT-Wörter-Anteils in den Textbeiträgen der Schülerinnen auf Normalverteilung.....	84
Tabelle 9: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen ($N = 533$) im Forum	86
Tabelle 10: Normierung der anteiligen MINT-Kommunikation der Schülerinnen ($N = 533$) im Forum in Prozenträngen und Stanine-Werten.....	86
Tabelle 11: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen im Forum nach Altersstufen	87
Tabelle 12: Normierung der anteiligen MINT-Kommunikation der Schülerinnen im Forum nach Altersstufen in Prozenträngen und Stanine-Werten	88
Tabelle 13: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen in den Unterforen MINT-Talk, Schulisches und Tummelplatz.....	90
Tabelle 14: Normierung der anteiligen MINT-Kommunikation der Schülerinnen in den Unterforen MINT-Talk, Schulisches und Tummelplatz in Prozenträngen und Stanine-Werten	91
Tabelle 15: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen ($N = 613$) im Chat	93
Tabelle 16: Normierung der anteiligen MINT-Kommunikation der Schülerinnen ($N = 613$) im Chat in Prozenträngen und Stanine-Werten.....	93
Tabelle 17: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen im Chat nach Altersstufen.....	94
Tabelle 18: Normierung der anteiligen MINT-Kommunikation der Schülerinnen im Chat nach Altersstufen in Prozenträngen und Stanine-Werten.....	95
Tabelle 19: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen ($N = 1 531$) in den plattform-internen Mail-Nachrichten.....	97

Tabelle 20: Normierung der Anteiligen MINT-Kommunikation der Schülerinnen ($N = 1531$) in den plattform-internen Mail-Nachrichten in Prozenträngen und Stanine-Werten.....	97
Tabelle 21: Mittelwerte und Standardabweichungen zur Kommunikation der Schülerinnen in den plattform-internen Mail-Nachrichten nach Altersstufen	98
Tabelle 22: Normierung der Anteiligen MINT-Kommunikation der Schülerinnen in den plattform-internen Mail-Nachrichten nach Altersstufen in Prozenträngen und Stanine-Werten.....	99
Tabelle 23: Deskriptive für die Kommunikation der Schülerinnen in Forum, Chat und Mail	108
Tabelle 24: Paralleltest-Reliabilitätskoeffizienten für die MINT-Kommunikation der Schülerinnen in Forum, Chat und Mail.....	111
Tabelle 25: Paralleltest-Reliabilitätskoeffizienten für die MINT-Kommunikation der 8- bis 13-jährigen Schülerinnen in Forum, Chat und Mail	112
Tabelle 26: Paralleltest-Reliabilitätskoeffizienten für die MINT-Kommunikation der 14- bis 15-jährigen Schülerinnen in Forum, Chat und Mail	113
Tabelle 27: Paralleltest-Reliabilitätskoeffizienten für die MINT-Kommunikation der 16- bis 18-jährigen Schülerinnen in Forum, Chat und Mail	114
Tabelle 28: Mittelwerte und Standardabweichungen für die Kommunikation der Schülerinnen in Forum, Chat und Mail	121
Tabelle 29: Split-Half-Reliabilitätskoeffizienten und Cronbachs Alphas für die MINT-Kommunikation der Schülerinnen in Forum, Chat und Mail	123
Tabelle 30: Mittelwerte und Standardabweichungen für die Kommunikation der Schülerinnen in der jeweils zweiten bis dritten Woche eines Monats	129
Tabelle 31: Retest-Reliabilitätskoeffizienten (Intraklassenkorrelationen) für die MINT-Kommunikation der Schülerinnen in Forum, Chat und Mail	130
Tabelle 32: Retest-Reliabilitätskoeffizienten (Intraklassenkorrelationen) für die MINT-Kommunikation der 8- bis 13-jährigen Schülerinnen in Forum, Chat und Mail	131
Tabelle 33: Retest-Reliabilitätskoeffizienten (Intraklassenkorrelationen) für die MINT-Kommunikation der 14- bis 15-jährigen Schülerinnen in Forum, Chat und Mail	132
Tabelle 34: Retest-Reliabilitätskoeffizienten (Intraklassenkorrelationen) für die MINT-Kommunikation der 16- bis 18-jährigen Schülerinnen in Forum, Chat und Mail	132
Tabelle 35: Mittelwerte und Standardabweichungen für die automatisch sowie manuell kodierte Kommunikation der Schülerinnen ($N = 125$) im Forum	139
Tabelle 36: Korrelationskoeffizienten für den Zusammenhang zwischen automatischer und manueller Koderung der MINT-Kommunikation der Schülerinnen ($N = 125$) im Forum	140
Tabelle 37: Mittelwerte und Standardabweichungen für die Kommunikation der Schülerinnen in Forum, Chat und Mail während der dreiwöchigen Erhebung...	144

Tabelle 38: Mittelwerte und Standardabweichungen der selbstberichteten Kommunikation der Schülerinnen während der Erhebung.....	145
Tabelle 39: Korrelationskoeffizienten für den Zusammenhang zwischen der automatisch kodierten und der selbstberichteten MINT-Kommunikation der Schülerinnen	146
Tabelle 40: Mittelwerte, Standardabweichungen sowie Mittelwertsunterschiede für die Kommunikation der Schülerinnen ($N = 75$) in den Unterforen MINT-Talk und Tummelplatz	150
Tabelle 41: Übersicht über die in den Online-Fragebogen verwendeten Messinstrumente zu Beginn (T_1) und am Ende (T_2) eines Mentoring-Jahres	154
Tabelle 42: Mittelwerte und Standardabweichungen für die Kommunikation der Schülerinnen in Forum, Chat und Mail.....	158
Tabelle 43: Mittelwerte und Standardabweichungen der Fragebogen-Skalen zu beiden Messzeitpunkten (T_1 und T_2) sowie t-Tests für verbundene Stichproben zwischen beiden Messzeitpunkten ($N = 425$)	159
Tabelle 44: Korrelationskoeffizienten für den Zusammenhang zwischen der MINT-Kommunikation der Schülerinnen während des Mentoring-Jahres und den selbstberichteten Mentoring-Outcomes nach Ende des Mentoring-Jahres (T_2) .	160
Tabelle 45: Ergebnisse der schrittweisen Regressionen für die anteilige MINT-Kommunikation in den Textbeiträgen der Schülerinnen ($N = 126$)	162
Tabelle 46: Ergebnisse der schrittweisen Regressionen für die absolute MINT-Wörteranzahl in den Textbeiträgen der Schülerinnen ($N = 126$).....	163
Tabelle 47: Mittelwerte und Standardabweichungen für die Kommunikation der Schülerinnen während der Mentoring-Runden 2010/11 und 2011/12.....	167
Tabelle 48: Korrelationskoeffizienten für den Zusammenhang zwischen der MINT-Kommunikation der Schülerinnen während des Mentoring-Jahres und einer Wiederanmeldung für das folgende Mentoring-Jahr	168
Tabelle 49: Ergebnisse der schrittweisen logistischen Regressionen für die anteilige MINT-Kommunikation in den Textbeiträgen der Schülerinnen zur Vorhersage der Wiederanmeldung der Schülerinnen am Ende des Mentoring-Jahres ($N = 1369$)	169
Tabelle 50: Ergebnisse der schrittweisen logistischen Regressionen für die absolute MINT-Wörteranzahl in den Textbeiträgen der Schülerinnen zur Vorhersage der Wiederanmeldung der Schülerinnen am Ende des Mentoring-Jahres ($N = 1369$)	170
Tabelle 51: Mittelwerte und Standardabweichungen für die Kommunikation der Schülerinnen in Forum, Chat und Mail.....	175
Tabelle 52: Korrelationskoeffizienten für den Zusammenhang zwischen der MINT-Kommunikation der Schülerinnen während des Mentorings und in Telefoninterviews erhobenen Selbstberichten nach ihrer Teilnahme am Mentoring ($N = 35$).....	176
Tabelle 53: Ergebnisse der schrittweisen logistischen Regressionen für die anteilige MINT-Kommunikation in den Textbeiträgen der Schülerinnen zur Vorhersage von Interviewfrage (2) „Durch CyberMentor weiß ich ziemlich sicher, was ich für einen Beruf/Studium machen möchte“ ($N = 35$).....	177

Tabelle 54: Ergebnisse der schrittweisen logistischen Regressionen für die anteilige MINT-Kommunikation in den Textbeiträgen der Schülerinnen zur Vorhersage von Interviewfrage (3) „Wofür hast du dich entschieden?“ ($N = 35$)	178
Tabelle 55: Ergebnisse der schrittweisen logistischen Regressionen für die absolute MINT-Wörterzahl in den Textbeiträgen der Schülerinnen zur Vorhersage von Interviewfrage (3) „Wofür hast du dich entschieden?“ ($N = 35$)	178
Tabelle 56: Übersicht über die Wörter bzw. Wortstämme des MINT-Wörterbuchs und ihre Zugehörigkeit zu den drei Wörterbuch-Komponenten	235