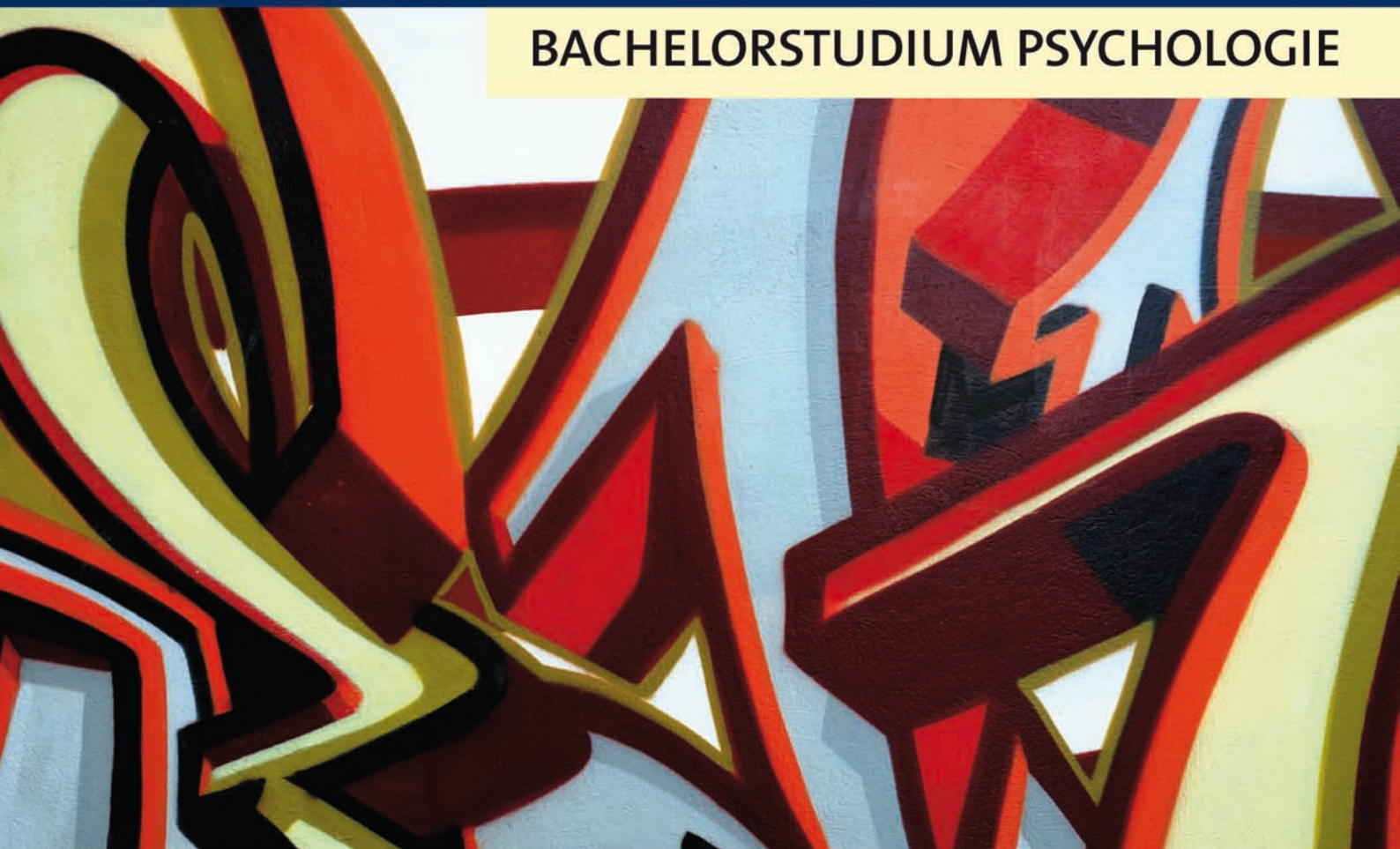


Heinz Holling · Günther Gediga

Statistik – Wahrscheinlichkeitstheorie und Schätzverfahren

BACHELORSTUDIUM PSYCHOLOGIE



Statistik – Wahrscheinlichkeitstheorie und Schätzverfahren

Bachelorstudium Psychologie

Statistik – Wahrscheinlichkeitstheorie und Schätzverfahren von Prof. Dr. Heinz Holling und PD Dr. Günther Gediga

Herausgeber der Reihe:

Prof. Dr. Eva Bamberg, Prof. Dr. Hans-Werner Bierhoff,
Prof. Dr. Alexander Grob, Prof. Dr. Franz Petermann

Statistik – Wahrscheinlichkeitstheorie und Schätzverfahren

von

Heinz Holling
und Günther Gediga

HOGREFE  GÖTTINGEN · BERN · WIEN · PARIS · OXFORD
PRAG · TORONTO · BOSTON · AMSTERDAM
KOPENHAGEN · STOCKHOLM · FLORENZ

Prof. Dr. Heinz Holling, geb. 1950. 1969–1976 Studium der Mathematik, Psychologie und Soziologie in Würzburg und Berlin. 1974–1987 Wissenschaftlicher Mitarbeiter an der FU Berlin und der Universität Osnabrück. Promotion 1980 (Dr. phil.) und 1987 (Dr. rer. nat.). 1987 Habilitation. 1987–1993 Vertretungsprofessor an den Universitäten Oldenburg, Münster und Mannheim. Seit 1993 Professor für Statistik und Quantitative Methoden am Psychologischen Institut der Universität Münster.

PD Dr. Günther Gediga, geb. 1953. 1971–1979 Studium der Informatik und Mathematik in Dortmund und Osnabrück. 1986 Promotion in Psychologie. 1994 Habilitation. 1979–2000 wissenschaftlicher Mitarbeiter im Fachgebiet Methodenlehre an der Universität Osnabrück. Seit 2001 außerplanmäßiger Professor am Department of Computer Science an der Brock University in St. Catharines, Kanada, und freiberufliche Tätigkeit als Statistical Consultant. Seit 2008 akademischer Rat an der Universität Münster im Institut für Statistik und Methoden.



Informationen und Zusatzmaterialien zu diesem Buch finden Sie unter www.hogrefe.de/buecher/lehrbuecher/psychlehrbuchplus

Wichtiger Hinweis: Der Verlag hat für die Wiedergabe aller in diesem Buch enthaltenen Informationen (Programme, Verfahren, Mengen, Dosierungen, Applikationen etc.) mit Autoren bzw. Herausgebern große Mühe darauf verwandt, diese Angaben genau entsprechend dem Wissensstand bei Fertigstellung des Werkes abzudrucken. Trotz sorgfältiger Manuskripterstellung und Korrektur des Satzes können Fehler nicht ganz ausgeschlossen werden. Autoren bzw. Herausgeber und Verlag übernehmen infolgedessen keine Verantwortung und keine daraus folgende oder sonstige Haftung, die auf irgendeine Art aus der Benutzung der in dem Werk enthaltenen Informationen oder Teilen davon entsteht. Geschützte Warennamen (Warenzeichen) werden nicht besonders kenntlich gemacht. Aus dem Fehlen eines solchen Hinweises kann also nicht geschlossen werden, dass es sich um einen freien Warennamen handelt.

© 2013 Hogrefe Verlag GmbH & Co. KG
Göttingen · Bern · Wien · Paris · Oxford · Prag · Toronto · Boston
Amsterdam · Kopenhagen · Stockholm · Florenz

<http://www.hogrefe.de>

Aktuelle Informationen · Weitere Titel zum Thema · Ergänzende Materialien

Copyright-Hinweis:

Das E-Book einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.

Der Nutzer verpflichtet sich, die Urheberrechte anzuerkennen und einzuhalten.

Umschlagabbildung: © Fitzer – istockphoto.com

Format: PDF

ISBN 978-3-8409-2135-3

Nutzungsbedingungen:

Der Erwerber erhält ein einfaches und nicht übertragbares Nutzungsrecht, das ihn zum privaten Gebrauch des E-Books und all der dazugehörigen Dateien berechtigt.

Der Inhalt dieses E-Books darf von dem Kunden vorbehaltlich abweichender zwingender gesetzlicher Regeln weder inhaltlich noch redaktionell verändert werden. Insbesondere darf er Urheberrechtsvermerke, Markenzeichen, digitale Wasserzeichen und andere Rechtsvorbehalte im abgerufenen Inhalt nicht entfernen.

Der Nutzer ist nicht berechtigt, das E-Book – auch nicht auszugsweise – anderen Personen zugänglich zu machen, insbesondere es weiterzuleiten, zu verleihen oder zu vermieten.

Das entgeltliche oder unentgeltliche Einstellen des E-Books ins Internet oder in andere Netzwerke, der Weiterverkauf und/oder jede Art der Nutzung zu kommerziellen Zwecken sind nicht zulässig.

Das Anfertigen von Vervielfältigungen, das Ausdrucken oder Speichern auf anderen Wiedergabegeräten ist nur für den persönlichen Gebrauch gestattet. Dritten darf dadurch kein Zugang ermöglicht werden.

Die Übernahme des gesamten E-Books in eine eigene Print- und/oder Online-Publikation ist nicht gestattet. Die Inhalte des E-Books dürfen nur zu privaten Zwecken und nur auszugsweise kopiert werden.

Diese Bestimmungen gelten gegebenenfalls auch für zum E-Book gehörende Audiodateien.

Inhaltsverzeichnis

1	Über dieses Buch	9
1.1	Zum Inhalt dieses Buches	11
1.2	Danksagung	13
2	Wahrscheinlichkeitsrechnung	15
2.1	Wozu wird die Wahrscheinlichkeitstheorie benötigt?	17
2.1.1	Das Konzept der Wahrscheinlichkeit im Alltag und in der Wissenschaft..	17
2.1.2	Zufall, Ergebnisse und Ereignisse: Ein einführendes Beispiel	19
2.2	Beschreibung von Zufallsvorgängen	21
2.2.1	Ergebnisräume	21
2.2.2	Ereignisräume	24
2.3	Mengenoperationen	25
2.4	Wahrscheinlichkeit	29
2.4.1	Wahrscheinlichkeiten und ihre Interpretation	31
2.5	Folgerungen aus den Axiomen von Kolmogorov	35
2.6	Bedingte Wahrscheinlichkeiten	38
2.7	Stochastische Unabhängigkeit	41
2.8	Rechnen mit bedingten und unbedingten Wahrscheinlichkeiten	43
2.8.1	Der Satz von der totalen Wahrscheinlichkeit	44
2.8.2	Der Satz von Bayes	46
2.9	Kombinatorik	51
2.10	Software	59
2.10.1	R	59
	Zusammenfassung	60
	Zentrale Begriffe	61
	Notation	64
3	Zufallsvariablen	65
3.1	Definition von Zufallsvariablen	68
3.2	Diskrete Zufallsvariablen	76
3.3	Stetige Zufallsvariablen und ihre Verteilungen	81
3.3.1	Wahrscheinlichkeitsverteilung einer stetigen Zufallsvariablen: Stetige Gleichverteilung	86

3.3.2	Normalverteilung	87
3.4	Quantile von Zufallsvariablen	92
3.4.1	Quantile stetiger Zufallsvariablen	92
3.4.2	Quantile diskreter Zufallsvariablen	94
3.5	Erwartungswert und Varianz einer Zufallsvariablen	95
3.5.1	Erwartungswert einer Zufallsvariablen	96
3.5.2	Varianz einer Zufallsvariablen	98
3.6	Erwartungswert und Varianz der Normalverteilung	100
3.7	Erwartungswert und Varianz transformierter Zufallsvariablen	101
3.7.1	Erwartungswert transformierter Zufallsvariablen	101
3.7.2	Varianz transformierter Zufallsvariablen	103
3.7.3	Standardisierung von Zufallsvariablen	104
3.8	Mehrdimensionale Zufallsvariablen	105
3.8.1	Zweidimensionale Zufallsvariablen und deren Verteilung	106
3.8.2	Unabhängigkeit von Zufallsvariablen	113
3.8.3	Berechnungsformeln für ein- und zweidimensionale Zufallsvariablen	114
3.8.4	Kovarianz und Korrelation von Zufallsvariablen	116
3.9	Rechnen mit Erwartungswerten und Varianzen von Zufallsvariablen	118
3.9.1	Erwartungswert der Summe von Zufallsvariablen	118
3.9.2	Varianz der Summe von Zufallsvariablen	118
	Zusammenfassung	122
	Zentrale Begriffe	125
	Notation	128
4	Verteilungen	129
4.1	Spezielle diskrete Wahrscheinlichkeitsverteilungen	131
4.1.1	Die Binomialverteilung	131
4.1.2	Die hypergeometrische Verteilung	138
4.1.3	Die geometrische Verteilung	142
4.1.4	Die negative Binomialverteilung	143
4.1.5	Die Poissonverteilung	144
4.1.6	Die Multinomialverteilung	146
4.2	Spezielle stetige Wahrscheinlichkeitsverteilungen	149
4.2.1	Die stetige Gleichverteilung	149
4.2.2	Die Betaverteilung	150
4.2.3	Die χ^2 -Verteilung	152
4.2.4	Zur Bedeutung der Freiheitsgrade	160
4.2.5	Die t -Verteilung	161
4.2.6	Die F -Verteilung	165
4.2.7	Zusammenhänge zwischen der Normal-, χ^2 -, t -, und F -Verteilung	168

4.3	Software	169
4.3.1	SPSS	169
4.3.2	R	172
	Zusammenfassung	175
	Zentrale Begriffe	175
	Notation	177
5	Stichprobenverteilungen	179
5.1	Stichprobenverteilung einer Statistik	181
5.2	Stichprobenverteilung des Mittelwerts einer normalverteilten Zufallsvariablen	183
5.3	Stichprobenverteilung des Mittelwerts nicht normalverteilter Zufallsvariablen	188
5.4	Stichprobenverteilung von Anteilswerten	193
5.5	Stichprobenverteilung der Varianz	198
5.6	Bootstrap-Verfahren	199
	Zusammenfassung	205
	Zentrale Begriffe	206
	Notation	207
6	Punktschätzung von Parametern	209
6.1	Einleitendes Beispiel	211
6.2	Gütekriterien von Schätzern	214
6.2.1	Erwartungstreue	215
6.2.2	Konsistenz	217
6.2.3	Effizienz	219
6.3	Maximum-Likelihood-Schätzung	221
6.3.1	Schätzung des Parameters einer Binomialverteilung	221
6.3.2	Schätzung des Erwartungswerts einer Normalverteilung	223
6.3.3	Rationale der Maximum-Likelihood-Schätzung	226
6.4	Methode der kleinsten Quadrate	232
6.5	Bayesianische Schätzmethoden	236
6.6	Bootstrap-Verfahren	245
	Zusammenfassung	245
	Zentrale Begriffe	247
	Notation	249

7	Intervallschätzungen	251
7.1	Konfidenzintervalle	254
7.1.1	Interpretation von Konfidenzintervallen	256
7.2	Konfidenzintervalle für Erwartungswerte	257
7.2.1	Konfidenzintervalle für den Erwartungswert μ normalverteilter Variablen mit bekannter Varianz σ^2	258
7.2.2	Konfidenzintervalle für den Erwartungswert μ normalverteilter Variablen mit unbekannter Varianz σ^2	260
7.2.3	Zufallsvariablen mit unbekannter Verteilung	262
7.3	Konfidenzintervalle für Varianzen	264
7.4	Konfidenzintervalle für Anteilswerte	266
7.4.1	Approximative Konfidenzintervalle für Anteilswerte	266
7.4.2	Wilson-Konfidenzintervalle	267
7.5	Konfidenzintervalle für die Regressionskoeffizienten in linearen Modellen	270
7.6	Konfidenzintervalle für die Produkt-Moment-Korrelation	272
7.7	Einseitige Konfidenzintervalle	275
7.8	Bootstrap-Intervalle	277
7.9	Bayesianische Intervallschätzungen	278
7.10	Software	281
7.10.1	SPSS	281
7.10.2	R	283
	Zusammenfassung	286
	Zentrale Begriffe	287
	Notation	289
Anhang		291
	Literatur	293
	Glossar	294
	Sachregister	304
	Normalverteilungstabelle	306
	Quantile der χ^2 -Verteilung	309
	Quantile der t-Verteilung	313
	Quantile der F-Verteilung	315



Kapitel 1

Über dieses Buch

Inhaltsübersicht

1.1	Zum Inhalt dieses Buches	11
1.2	Danksagung	13

1.1 Zum Inhalt dieses Buches

Der vorliegende Band ist der zweite Teil eines insgesamt dreibändigen Lehrbuchs der Statistik. Der gesamte Stoff ist gegliedert in die deskriptive (beschreibende) Statistik (Band 1), Wahrscheinlichkeitstheorie und Schätzverfahren (Band 2) und statistische Testverfahren (Band 3). Jedoch setzt dieser Band nicht unbedingt die Lektüre des ersten Bandes voraus. Leserinnen und Leser¹ können diesen Band unabhängig vom ersten Band rezipieren, wenn sie über elementare Kenntnisse der deskriptiven Statistik verfügen, wie sie in den gängigen Statistiklehrbüchern vermittelt werden.

Grundlegendes Ziel dieses Bandes, wie auch der beiden übrigen Bände, ist es, die Statistik umfassend und verständlich darzustellen. Es wird sehr viel Wert darauf gelegt, den Sinn der hier behandelten wahrscheinlichkeitstheoretischen und statistischen Konzepte aufzuzeigen und ihren Nutzen für die Gewinnung neuer Erkenntnisse darzulegen. Dazu werden die einzelnen Themen anschaulich und fundiert vermittelt, stets wird der Bezug zu inhaltlichen Problemen hergestellt.

Da Studierende der Psychologie eine wesentliche Zielgruppe dieses Buches bilden, sind viele Beispiele aus dem Bereich der Psychologie entnommen. Nichtsdestoweniger ist dieser Band ebenso für Studierende anderer Studiengänge aus den Natur-, Sozial- und Erziehungswissenschaften als grundlegende Einführung geeignet.

In diesem Band erfolgt zunächst eine ausführliche Darstellung der Wahrscheinlichkeitsrechnung. Sie bildet nicht nur die Voraussetzung für die dann abgehandelten Verfahren zur Schätzung von Parametern, sondern stellt selbst eine wichtige Grundlage für viele inhaltliche Themen dar, wie z. B. zur Entscheidungsforschung. Daher erfolgt eine umfassende Einführung in wesentliche Konzepte der Wahrscheinlichkeitstheorie.

Ebenso gründlich werden Zufallsvariablen hergeleitet, die einen grundlegenden Bestandteil der gesamten Statistik bilden. Bei der Darstellung der einzelnen Wahrscheinlichkeitsverteilungen wird sogleich viel Wert darauf gelegt, ihren Beitrag bzw. Stellenwert für die Lösung statistischer Probleme herauszustellen.

¹Wir werden jedoch im Folgenden immer dann, wenn beide Geschlechter gemeint sind, lediglich die männliche Form benutzen. Damit wird u. E. die Darstellung sehr vereinfacht. Wenn wir explizit die weibliche Form nutzen, sind nur Personen weiblichen Geschlechts gemeint.

In diesem Band werden die Verfahren zur Parameterschätzung ausführlicher behandelt als in vielen anderen Textbüchern zur Statistik. Diese Verfahren stellen auch in anderen Disziplinen, wie z. B. der psychologischen Testtheorie, eine wesentliche Grundlage dar. So wird das Maximum-Likelihood-Schätzverfahren recht umfassend dargestellt. Zudem werden aber auch die in jüngster Zeit immer populärer werdenden Bayes'schen Verfahren zur Parameterschätzung in diesem Band abgehandelt.

Ein Lehrbuch der Statistik ist einfacher zu lesen, wenn es einheitlich und übersichtlich strukturiert ist. Dazu haben wir völlig analog zum ersten Band dieser Lehrbuchreihe die folgenden Gestaltungsrichtlinien angewendet.

Um die Darstellung der statistischen Konzepte nicht zu überfrachten, wird auf die Darstellung mathematischer Ableitungen und Beweise, die für das grundlegende Verständnis nicht unbedingt erforderlich sind, verzichtet. Für besonders interessierte Leserinnen und Leser werden diese Beweise auf der folgenden Website zur Verfügung gestellt:

<http://www.hogrefe.de/buecher/lehrbuecher/psychlehrbuchplus>



Im Text wird jeweils auf diese Website hingewiesen, zusätzlich werden die entsprechenden Stellen durch das „Internetsymbol“ am Seitenrand gekennzeichnet.

Um die Inhalte dieses Lehrbuchs übersichtlich darzustellen, verwenden wir im Text verschiedene Arten von Kästen. Diese jeweils unterschiedlich gekennzeichneten Kästen enthalten:

- Beispiele,
- Anwendungen von Software (SPSS und R),
- Zusammenfassungen.

Wichtige Formeln werden ebenfalls durch eine farbige Hinterlegung hervorgehoben, damit sie unmittelbar als solche zu erkennen sind. Bei der Einführung neuer Begriffe werden diese Begriffe im Text kursiv hervorgehoben und zudem am Seitenrand aufgeführt. Am Seitenrand befindet sich des Weiteren das bereits erwähnte Internetsymbol.

Die Vermittlung von Statistik ist heutzutage ohne die Verwendung von EDV-Programmen undenkbar und es gibt mittlerweile eine Vielzahl sehr leistungsfähiger Statistikprogrammpakete. Im vorliegenden Lehrbuch nutzen wir die Softwarepakete SPSS und R.

Das Programmpaket IBM SPSS Statistics, kurz SPSS genannt, ist eine seit jeher sehr weit verbreitete Statistiksoftware. In vielen Organisationen, in denen Natur-, Sozial- und Erziehungswissenschaftler beschäftigt

sind, wird dieses Programmpaket verwendet. SPSS hat allerdings den Nachteil, dass es kostenpflichtig ist.

Das Open-Source-Programm R ist ebenfalls weit verbreitet, hingegen kostenlos verfügbar. R ist allerdings nicht ganz so benutzerfreundlich wie SPSS. Jedoch bietet auch R eine Menüsteuerung mittels des sogenannten R-Commanders an, der für Anfänger eine einfache Benutzung der wesentlichen Funktionen dieses Programms erlaubt. Ein großer Vorteil von R besteht darin, neueste statistische Verfahren aus zuverlässigen Quellen einzubinden. Ebenso ist es mit R im Vergleich zu vielen anderen Statistikprogrammen relativ einfach möglich, eigene statistische Programme zu entwickeln.

Im vorliegenden Buch behandeln wir am Ende der Kapitel jeweils die notwendigen Schritte zur Datenanalyse mittels SPSS und R. Für SPSS sowie für R stellen wir die jeweiligen Menübefehle dar. Da der Funktionsumfang des R-Commanders begrenzt ist, werden zusätzlich die wesentlichen Befehle der R-Syntax aufgeführt. Studierende, die diese Programmsysteme nicht kennen, finden auf der oben genannten Website Hinweise auf eine Einführung in R sowie Anleitungen zur Benutzung von R. Für die Einführung in die Arbeit mit SPSS möchten wir auf das Buch von Leonhart (2010) verweisen.

Auf der Website befinden sich neben den oben bereits angesprochenen Inhalten viele weitere Ressourcen zu diesem Lehrbuch. Dabei handelt es sich um die in diesem Buch benutzten Datensätze und die entsprechenden Kommandos für die Programme SPSS und R. Somit können die in diesem Buch dargestellten statistischen Analysen eigenständig nachvollzogen werden. Des Weiteren enthält die Website ergänzende Inhalte und Literaturhinweise.

1.2 Danksagung

Zur Entstehung dieses Buches haben nicht nur die Autoren, sondern zahlreiche weitere Personen beigetragen. Viele wertvolle Kommentare und Überarbeitungshinweise erhielten wir nach dem Studium einer vorläufigen Endversion von Herrn Paul-Christian Bürkner, Herrn Dr. Philipp Doeblner, Herrn Dipl.-Psych. Boris Forthmann, Frau Ruth Helmich, Frau Lea Johannsen, Frau Inken Kirschbaum, Herrn Dr. Jörg-Tobias Kuhn, Frau Miriam Miesen, Frau Jana Scharfen und Herrn Philipp Schäpers. Dafür möchten wir ihnen an dieser Stelle unseren großen Dank aussprechen. Ein ganz besonderer Dank gebührt Frau Dr. Anna Doeblner. Ihre sorgfältige Überprüfung des gesamten Textes in formaler und inhaltlicher Hinsicht war sehr hilfreich.