

Lisa Christina Schneider

Das MAI-Tool als Untersuchungsinstrument von Lösungsprozessen beim mathematischen Modellieren

MOREMEDIA



Springer Spektrum

Mathematikdidaktik im Fokus

Reihe herausgegeben von

Rita Borromeo Ferri, FB 10 Mathematik, Universität Kassel, Kassel, Deutschland

Andreas Eichler, Institut für Mathematik, Universität Kassel, Kassel, Deutschland

Elisabeth Rathgeb-Schnierer, Institut für Mathematik, Universität Kassel, Kassel, Deutschland

In dieser Reihe werden theoretische und empirische Arbeiten zum Lehren und Lernen von Mathematik publiziert. Dazu gehören auch qualitative, quantitative und erkenntnistheoretische Arbeiten aus den Bezugsdisziplinen der Mathematikdidaktik, wie der Pädagogischen Psychologie, der Erziehungswissenschaft und hier insbesondere aus dem Bereich der Schul- und Unterrichtsforschung, wenn der Forschungsgegenstand die Mathematik ist.

Die Reihe bietet damit ein Forum für wissenschaftliche Erkenntnisse mit einem Fokus auf aktuelle theoretische oder empirische Fragen der Mathematikdidaktik.

Lisa Christina Schneider

Das MAI-Tool als Untersuchungsinstrument von Lösungsprozessen beim mathematischen Modellieren

Lisa Christina Schneider
Fachbereich Mathematik
Technische Universität Kaiserslautern
Kaiserslautern, Deutschland

Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor philosophiae naturalis, Dr. phil. nat.) genehmigte Dissertation, 2022

D 386

Erstgutachter: Prof. Dr. Stefan Ruzika

Erstgutachterin: Prof. Dr. Rita Borromeo Ferri

Zweitgutachter: Prof. Dr. Werner Blum

Tag der Disputation: 09. Dezember 2022

ISSN 2946-0174

ISSN 2946-0182 (electronic)

Mathematikdidaktik im Fokus

ISBN 978-3-658-41731-4

ISBN 978-3-658-41732-1 (eBook)

<https://doi.org/10.1007/978-3-658-41732-1>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2023

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Marija Kojic

Springer Spektrum ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

„Nein“, rief er [Klaus Heinrich], „heute dürfen Sie keine Algebra treiben, oder im luftleeren Raum spielen, wie Sie es nennen! Sehen Sie doch die Sonne! ... Darf ich ... ?“ Und er trat zum Tischchen und nahm das Kollegheft zur Hand. Was er sah, war sinnverwirrend. In einer krausen, kindlich dick aufgetragenen Schrift, die Imma Spoelmanns besondere Federhaltung erkennen liess, bedeckte ein phantastischer Hokusfokus, ein Hexensabbat verschränkter Runen die Seiten. Griechische Schriftzeichen waren mit lateinischen und mit Ziffern in verschiedener Höhe verkoppelt, mit Kreuzen und Strichen durchsetzt, ober- und unterhalb waagrechter Linien bruchartig aufgereiht, durch andere Linien zeltartig überdacht, durch

*Doppelstrichelchen gleichgewertet,
durch runde Klammern zu grossen
Formelmassen vereinigt. Einzelne
Buchstaben, wie Schildwachen
vorgeschoben, waren rechts oberhalb der
umklammerten Gruppen ausgesetzt.
Kabbalistische Male, vollständig
unverständlich dem Laiensinn, umfassten
mit ihren Armen Buchstaben und Zahlen,
während Zahlenbrüche ihnen
voranstanden und Zahlen und
Buchstaben ihnen zu Häupten und Füßen
schwebten. Sonderbare Silben,
Abkürzungen geheimnisvoller Worte
waren überall eingestreut, und zwischen
den nekromantischen Kolonnen standen
geschriebene Sätze und Bemerkungen in
täglicher Sprache, deren Sinn so hoch
über allen menschlichen Dingen war,
dass man sie lesen konnte, ohne mehr
davon zu verstehen, als von einem
Zaubergemurmel.*

Thomas Mann, Königliche Hoheit

Vorwort

Mathematisches Modellieren ist eine zentrale mathematische Kompetenz und wird unter anderem in den Bildungsstandards Mathematik genannt und gefordert. Mathematische Modellierung meint den Prozess, ein Abbild der Realität zu bilden, um damit eine realitätsbezogene Fragestellung mit Hilfe der Mathematik (und ggfs. weiterer Disziplinen) beantworten zu können. Mathematische Modellierung ist aber auch eine Schlüsselkompetenz in den MINT-Disziplinen und damit von großer Relevanz, um komplexe Problemstellungen aus unterschiedlichen Domänen „greifbar“ formulieren und lösen zu können. Dennoch kommt die Modellierung von offenen Fragestellungen in der Schule häufig noch zu kurz: Modellierungsaufgaben sind oft zeitlich aufwändig, haben meist keinen eindeutig vorgegebenen Lösungsweg, sie sind komplex und werden oft im Team bearbeitet.

Es gibt eine rege (inter)nationale Community, die sich mit didaktischen Fragestellungen zu diesem Thema beschäftigt. Dennoch existieren noch einige Forschungslücken zum strukturellen und zeitlichen Ablauf des Modellierungsprozesses von Schülerinnen und Schülern. Hier setzt die Arbeit von Lisa Schneider an: Grundlegendes Ziel ist es, den Modellierungsprozess besser zu verstehen. Insbesondere können Modellierungsaktivitäten von Schülerinnen und Schülern systematisch und präzise erfasst und ausgewertet werden.

Existierende Ansätze in diese Richtung sind eher konzeptioneller Art und wurden bislang mehr oder weniger per Hand umgesetzt. Eine Software, die Forschende bei der Dateneingabe, bei der Visualisierung von Aktivitätsverläufen während des Modellierens und bei der Auswertung der Daten unterstützt, fehlte bislang. Das entwickelte Modelling-Activity-Interaction-Tool (MAI-Tool) ermöglicht es nun, Strukturen von Modellierungsverläufen sichtbar zu machen und zu analysieren. Mit dem MAI-Tool käme dann auch Quasi-Standard für die Speicherung und Erhebung von Beobachtungsdaten. Damit könnten also Daten,

die im Verlauf von Studien erzielt wurden, unter den Forschenden ausgetauscht und somit Ergebnisse reproduziert werden. Unter Umständen könnten diese Daten dann sogar für andere Fragestellungen erneut genutzt werden. In der vorliegenden Arbeit stellt Lisa Schneider das MAI-Tool vor, dass all dies ermöglicht.

Mit dieser Software drängen sich dann natürlich auch weitere didaktische Forschungsfragen auf. Gelingt es z. B., diese beobachteten Strukturen und mit dem MAI-Tool erfassten Strukturen durch geeignete Interventionen oder Schulungen zu beeinflussen, so dass gewünschte Änderungen entstehen? Insbesondere ist bislang noch nicht ausreichend erforscht, wie sich (Meta-)Wissen über den Prozess des Modellierens beim Modellieren selbst auswirkt. Lisa Schneider präsentiert dazu in der vorliegenden Arbeit eine Vergleichsstudie, bei der instruierte und nicht-instruierte Schülerinnen und Schüler mithilfe des MAI-Tools auf individueller Ebene und auf Ebene der Gruppe untersucht wurden.

Maschinelles Lernen hat in den letzten Jahren grundlegende Änderungen in der angewandten Mathematik und vielen anderen Disziplinen bewirkt. In der fachdidaktischen Forschung zur mathematischen Modellierung gehören diese Methoden (und allgemein: Methoden der künstlichen Intelligenz) noch nicht zum Standardrepertoire. Hier liegt jedoch ein großes Potential: Basierend auf den mit dem MAI-Tool erzeugten Daten könnten z. B. Klassifikationsalgorithmen Lernende automatisiert und zuverlässig in Gruppen einteilen z. B. basierend auf Kompetenzniveaus oder vorhandenen Fehlvorstellungen. Auch wenn dies aktuell noch Grundlagenforschung darstellt, so lassen sich doch unmittelbar Anwendungen im Unterrichtskontext erkennen bis hin zu einem digitalen Assistenzwerkzeug, das Lehrkräfte unterstützt. Auch hierzu formuliert Lisa Schneider in ihrer Arbeit erste Ansätze, wie maschinelles Lernen im Kontext der mathematischen Modellierung gewinnbringend Anwendung finden kann.

Kaiserslautern
im März 2023

Prof. Dr. Stefan Ruzika

Danksagung

Mein größter Dank geht an meine Doktoreltern Prof. Dr. Stefan Ruzika und Prof. Dr. Rita Borromeo Ferri. Ich danke Stefan Ruzika, dass er mir die Möglichkeit einer Promotion eröffnet und mich bei meiner Forschung immer unterstützt hat. Ohne seine Kreativität und Ideen wäre die Arbeit in ihrer aktuellen Form nicht möglich gewesen. Rita Borromeo Ferri danke ich insbesondere für die Unterstützung und hilfreichen Ratschläge zu meinen fachdidaktischen Themen, die meine Arbeit entscheidend geprägt haben. Sie hat mich immer in meiner Forschung bestärkt und motiviert.

Weiterhin danke ich Prof. Dr. Werner Blum für die Bereitschaft und Mühe als Zweitgutachter dieser Arbeit.

Bedanken möchte ich mich bei Dr. Florentine Kämmerer, die mich darin bestärkt hat, eine Promotion anzustreben.

Außerdem möchte ich mich ganz herzlich bei Sebastian Hiller und Mattia Masiero bedanken, die mich bei der Implementierung des MAI-Tools tatkräftig unterstützt haben. Danke für die vielen Nachschichten, ohne die das MAI-Tool nicht in seiner aktuellen Version wäre. Weiterhin geht mein Dank an Dominik Frey für die Unterstützung beim Kodieren der Videos.

Vielen Dank an Kathrin Kennel, Lynn Knippertz, Paul Weber und Corinna Zurloh für die konstruktiven Hinweise und Verbesserungsvorschläge, die diese Arbeit vorangebracht haben. Auch danke ich Dr. Jean-Pierre Stockis für die hilfreichen Anmerkungen zur Statistik.

Ein besonderer Dank gilt meiner Familie sowie meinen Freundinnen und Freunden, die mich besonders in schwierigen Phasen dieser Arbeit immer unterstützt haben. Danke!

Zusammenfassung

Mathematische Modellierung ist in den curricularen Bildungsstandards verankert und somit ein fester Bestandteil des Mathematikunterrichts. Lernende sollen in die Lage versetzt werden, Fragestellungen aus der Realität mit Mathematik zu lösen. Deswegen ist es von Relevanz, das Vorgehen von Lernenden beim Bearbeiten einer Modellierungsaufgabe zu untersuchen und daraus Folgerungen für den Mathematikunterricht abzuleiten.

In der Arbeit wird zunächst das neue Untersuchungsinstrument *Modelling-Activity-Interaction-Tool* (MAI-Tool) präsentiert, das zur Analyse von Lösungsprozessen beim Modellieren entwickelt, implementiert und schließlich validiert wurde. Bekannte Methoden zur Untersuchung von Lösungsprozessen werden detailliert vorgestellt, die im MAI-Tool aufgegriffen und wesentlich erweitert wurden: Neben einer digitalen Erfassung werden die Modellierungsprozesse automatisiert algorithmisch ausgewertet, um quantitative Ergebnisse zu erzielen. Mit dem MAI-Tool ist ein neuer Blickwinkel möglich: Das Vorgehen beim Modellieren wird anhand der Struktur von Lösungsprozessen umfassend untersucht. Strukturmerkmale werden vom Modellierungskreislauf abgeleitet und präzise definiert.

Das MAI-Tool wird anschließend zur Erfassung und Auswertung von individuellen Lösungsprozessen in einer empirischen Studie angewendet. Der Fokus liegt auf dem Einfluss von Wissen über idealtypische Modellierungsprozesse auf die Struktur von individuellen Lösungsprozessen. Zunächst erfolgt eine umfangreiche und detaillierte Analyse der individuellen Lösungsprozesse anhand der im MAI-Tool implementierten Strukturmerkmale. Im Rahmen der quantitativen Auswertung konnte der Einfluss des Wissens auf die individuellen Lösungsprozesse nachgewiesen und daraus ableitend Hürden beim Modellieren identifiziert werden. Zusätzlich konnte ein Zusammenhang der Interaktionen – betrachtet über

den gesamten Lösungsprozess als auch in den Phasen – und der Sympathie sowie Antipathie zwischen den Gruppenmitgliedern verdeutlicht werden.

In dieser Arbeit wird ein bisher noch nicht berücksichtigter Ansatz verfolgt, um die Ergebnisse der empirischen Studie zu untermauern: Die Daten aus der empirischen Studie wurden verwendet, um einen Algorithmus mit Machine Learning zu entwickeln, der die individuellen Lösungsprozesse anhand der Strukturmerkmale klassifiziert. Dabei wurde die Leistung verschiedener Klassifikationsalgorithmen verglichen: Die *Support-Vector-Machine* als auch die *logistische Regression* klassifizieren die Daten am besten.

Insgesamt leistet die Arbeit einen Beitrag zur Grundlagenforschung in mathematischer Modellierung auf technologischer, empirischer und algorithmischer Ebene.

Abstract

Mathematical modelling is integrated in the curriculum and is thus an integral part of mathematics teaching. Students should be enabled to solve problems from reality with mathematics. Therefore, it is relevant to investigate the procedure of students when working on a modelling task and to derive conclusions for mathematics teaching.

The thesis first presents the new research tool *Modelling-Activity-Interaction-Tool* (MAI-Tool), which was developed, implemented and finally validated for the analysis of solution processes during modelling. Well-known methods for the investigation of solution processes are presented in detail, which have been taken up and significantly extended in the MAI-Tool: In addition to a digital collection, the modelling processes are automatically evaluated algorithmically in order to achieve quantitative results. With the MAI-Tool, a new perspective is possible: the modelling procedure is comprehensively examined on the basis of the structure of the solution process. Structural features are derived from the modelling cycle and precisely defined.

The MAI-Tool is then used to capture and evaluate individual solution processes in an empirical study. The focus is on the influence of knowledge about ideal-typical modelling processes on the structure of individual solution processes. First, a comprehensive and detailed analysis of the individual solution processes is carried out using the structural features implemented in the MAI-Tool. Within the framework of the quantitative evaluation, the influence of knowledge on the individual solution processes could be proven and, deriving from this, hurdles in modelling could be identified. In addition, a relationship between the interactions – considered over the entire solution process as well as in the phases – and the sympathy and antipathy between the group members could be clarified.

In this thesis, a previously unconsidered approach is taken to substantiate the results of the empirical study: The data from the empirical study was used to develop a machine learning algorithm to classify individual solution processes based on structural features. The performance of different classification algorithms was compared: *support vector machine* as well as *logistic regression* classified the data best.

Overall, this thesis contributes to basic research in mathematical modelling at the technological, empirical and algorithmic level.

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen	7
2.1	Grundlagen mathematischer Modellierung	7
2.2	Grundlagen der Graphentheorie	15
Teil I Das Modelling-Activity-Interaction-Tool (MAI-Tool)		
3	Grundlagen des MAI-Tools	25
3.1	Vorläufer des MAI-Tools	25
3.2	Empirische Ergebnisse zur Struktur von Lösungsprozessen	33
3.3	Herleitung der Forschungsfragen	36
4	Entwicklung, Implementierung und Validierung des MAI-Tools	41
4.1	Vorstellung des MAI-Tools und Definition der Strukturmerkmale	42
4.2	Einblick in die Implementierung	58
4.3	Überprüfung der Gütekriterien	67
4.4	Diskussion und Beantwortung der Forschungsfragen	74

Teil II Eine empirische Analyse über die Struktur von individuellen Lösungsprozessen und den Zusammenhang zwischen Gruppenstruktur und Interaktionsprozess

5	Gruppenprozesse und Strategien beim Lösen von Modellierungsaufgaben	81
5.1	Gruppenunterricht und Gruppenstruktur	82
5.2	Hürden im Modellierungsprozess	86
5.3	Strategien beim mathematischen Modellieren	88
5.4	Herleitung der Forschungsfragen und Ziele der Untersuchung	94
5.5	Herleitung der Hypothesen	97
6	Methodischer Rahmen	103
6.1	Unterweisung in idealtypische Modellierungsprozesse	104
6.2	Methodische Verortung	108
6.3	Studiendesign: Erhebungsmethoden und Datenerfassung	111
6.4	Stoffdidaktische Analyse der „Tanken“-Aufgabe	123
6.5	Auswertungsmethoden	127
7	Analyse und Vergleich von individuellen Lösungsprozessen	135
7.1	Beschreibung der Struktur individueller Lösungsprozesse	135
7.2	Vergleich individueller Lösungsprozesse	147
8	Zusammenhang der Gruppenstruktur und des Interaktionsprozesses	151
9	Diskussion und Beantwortung der Forschungsfragen	159

Teil III Klassifikation individueller Lösungsprozesse mit Machine Learning

10	Motivation und Herleitung der Forschungsfrage	171
11	Einführung in Machine Learning	175
12	Bausteine der Klassifikation	181
12.1	Vorbereitung der Daten	181
12.2	Klassifikationsalgorithmen und Hyperparametersuche	182
12.3	Entwicklung und Evaluation des Klassifikationsmodells	195
12.4	Feature selection	201

13 Vergleich der Performance verschiedener Klassifikationsalgorithmen	207
13.1 Implementierung	207
13.2 Ergebnisse	211
13.3 Diskussion und Beantwortung der Forschungsfrage	215
Teil IV Zusammenfassung der Arbeit	
14 Limitationen der Arbeit	221
15 Fazit der Arbeit und Ausblick für Forschung und Praxis	225
Literatur	231

Abbildungsverzeichnis

Abbildung 2.1	Klassifikation von mathematischen Modellen	10
Abbildung 2.2	Einschrittiges Mathematisieren im Kreislauf (Schupp 1988, S. 11)	12
Abbildung 2.3	Zweischrittiges Mathematisieren im Kreislauf (W. Blum 1985, S. 200)	12
Abbildung 2.4	Dreischrittiges Mathematisieren im Kreislauf (W. Blum et al. 2005, S. 19)	13
Abbildung 2.5	Der fünfschrittige Modellierungskreislauf (Kaiser et al. 2013), zitiert nach Stender (2016, S. 31)	13
Abbildung 2.6	Ein ungerichteter (a), ein gerichteter (b) und ein gewichteter (gerichteter) Graph (c)	18
Abbildung 2.7	Eine Kontaktsequenz, bei der die Startzeit der Interaktion als dazugehöriges Kantengewicht angegeben ist	19
Abbildung 2.8	Ein Intervallnetzwerk, bei dem die Dauer der Kontakte im Intervall als Kantengewicht angegeben ist	20
Abbildung 2.9	Zeitveränderliche Netzwerke für $\delta t = 10$ min (a) und $\delta t = 30$ min (b)	21
Abbildung 3.1	Das Modelling Activity Diagram (MAD) (Ärlebäck 2009, S. 346) visualisiert, welche Aktivitäten zu welchen Zeitpunkten beobachtet wurden	27

Abbildung 3.2	Das erweiterte MAD (Ärleböck et al. 2019, S. 6): Zusätzlich zum MAD werden die Teilprobleme der Fermi-Aufgabe durch farbliche Abstufungen (braun) voneinander abgetrennt	28
Abbildung 3.3	Das MCAD, das neben der Darstellung der Aktivitäten über die Zeit auch die Ideen im Lösungsprozess angibt (Hartmann et al. 2019, S. 98)	28
Abbildung 3.4	Das MAD+ (Pla-Castells et al. 2021). Die Zeichen symbolisieren die Fehler: Vereinfachen (Dreieck), Mathematisieren (Raute), Lösen (Kappe) und Validieren (Stern)	29
Abbildung 3.5	Individuelle Modellierungsverläufe (Borromeo Ferri 2011, S. 115)	30
Abbildung 3.6	Ein Modeling Transition Diagram (MTD) (Czocher 2016, S. 90), das einen idealisierten Modellierungsprozess zeigt. Beim MTD wird der Anfangszeitpunkt einer Aktivität eingetragen	31
Abbildung 3.7	Die Lernenden mit der Lehrkraft – dargestellt als Netzwerk (Bokhove 2018, S. 25)	32
Abbildung 3.8	Minikreisläufe (Borromeo Ferri 2011, S. 152): Mathematisches Zwischenergebnis (MZE), Realmodellrückinterpretation (RMR)	35
Abbildung 3.9	Dualer Modellierungskreislauf (Saeki et al. 2013, S. 94)	36
Abbildung 4.1	Der Startbildschirm des MAI-Tools	42
Abbildung 4.2	Unter <i>Beobachtung</i> können der Kode der Gruppe sowie Interaktionseinheiten eingegeben werden	43
Abbildung 4.3	Im <i>Editiermodus</i> kann die Dateneingabe eingesehen und korrigiert werden	44
Abbildung 4.4	Im <i>Editiermodus</i> können neue Interaktionseinheiten hinzugefügt (links) oder bereits eingegebene korrigiert werden (rechts)	45
Abbildung 4.5	Im <i>Editiermodus</i> wird rot markiert, wenn eine Person zeitgleich in zwei Interaktionseinheiten kodiert wurde	46

Abbildung 4.6	Das Maß der Übereinstimmung beider kodierenden Personen wird für jedes Individuum mit <i>Cohen's Kappa</i> berechnet. Außerdem wird die Tabelle angezeigt, die Basis für die Berechnung ist	46
Abbildung 4.7	Vergleich der Snapshots zur Überprüfung der Übereinstimmung der kodierenden Personen auf Gruppenebene	47
Abbildung 4.8	Das MAIN zu drei Zeitpunkten, wobei jede Kante nach der jeweiligen Modellierungsphase gefärbt ist	48
Abbildung 4.9	Die Auswertungsmaske im MAI-Tool	49
Abbildung 4.10	Das MAD nach Ärlebäck (2009) wird ebenfalls im MAI-Tool erstellt – statt Aktivitäten werden Phasen dargestellt	50
Abbildung 4.11	Darstellung der chronologischen Abfolge der Phasen	50
Abbildung 4.12	Auswertung der Phasen nach deren Auftreten (links) und Dauer (rechts) in den Interaktionen	51
Abbildung 4.13	Ein individueller Lösungsprozess ist in vier Blöcke unterteilt. Die Bedeutung der Farben ist in Abbildung 4.11 nachzulesen	52
Abbildung 4.14	Auswertung der Phasen nach Auftreten (a) und Dauer (b) in jedem der vier Blöcke	53
Abbildung 4.15	Klassifikation der Phasenwechsel: benachbart (grün), benachbart rückgängig (blau), Sprung (rot) und außerhalb (schwarz)	54
Abbildung 4.16	Auswertung der Phasenwechsel im MAI-Tool	55
Abbildung 4.17	Klassifizierung und Darstellung der Muster am Graphen: Ein Teilkreislauf der Länge 2 (rot) und der Länge 3 (orange), ein verschachtelter Teilkreislauf (grün), ein alternierender erweiterter Teilkreislauf (blau) und ein erweiterter Teilkreislauf Dreieck (pink)	56
Abbildung 4.18	Chronologische Abfolge der Phasen mit Hervorhebungen der Muster	57

Abbildung 4.19	Darstellung der Muster in einer Tabelle: Neben der Abfolge der Phasen wird der Mustertyp, die Länge, die absolute Häufigkeit sowie die Farbe der Kanten angegeben	57
Abbildung 4.20	Die detektierten Muster werden separat zur besseren Veranschaulichung in Graphen dargestellt	57
Abbildung 4.21	Ausschnitt des Quellcodes aus dem MAI-Tool: Aus den Eingabedaten der Gruppe werden die der Individuen extrahiert	59
Abbildung 4.22	Ausschnitt des Quellcodes aus dem MAI-Tool: Für jedes Individuum werden die absoluten und relativen Häufigkeiten bzw. Dauer berechnet, wie oft bzw. wie lange die Phase im Modellierungsprozess in Interaktionen kodiert wurde	60
Abbildung 4.23	Ausschnitt des Quellcodes aus dem MAI-Tool: Für jedes Individuum wird der Endzeitpunkt und somit die Gesamtdauer des Modellierungsprozesses bestimmt	60
Abbildung 4.24	Ausschnitt des Quellcodes aus dem MAI-Tool: Jeder individuelle Modellierungsprozess wird in vier Blöcke unterteilt, in die die Interaktionseinheiten zugeteilt werden	61
Abbildung 4.25	Ausschnitt des Quellcodes aus dem MAI-Tool: Die Kanten werden den jeweiligen Phasenwechseln zugeordnet und in der jeweiligen Farbe gefärbt	62
Abbildung 4.26	Ausschnitt des Quellcodes aus dem MAI-Tool: Es werden Kreisläufe in den Abfolgen der Phasen individueller Modellierungsprozesse gespeichert, bei denen die Anfangs- und Endphase gleich sind	63
Abbildung 4.27	Ausschnitt des Quellcodes aus dem MAI-Tool: Klassifizierung der Muster als Teilkreisläufe, verschachtelte Teilkreisläufe und erweiterte Teilkreisläufe	64

Abbildung 4.28	Ausschnitt des Quellcodes aus dem MAI-Tool: Die längste Dauer von beiden Modellierungsprozessen wird als Endzeit festgesetzt	65
Abbildung 4.29	Ausschnitt des Quellcodes aus dem MAI-Tool: Interaktionen und Akte werden für jede Sekunde gespeichert	66
Abbildung 4.30	Ausschnitt des Quellcodes aus dem MAI-Tool: Für nichtaktive Individuen werden Schleifen erstellt	66
Abbildung 4.31	Ausschnitt des Quellcodes aus dem MAI-Tool: Vergleich der Kanten beider Graphen und Berechnung der sekundenweisen Übereinstimmung beider kodierenden Personen auf der Gruppenebene	67
Abbildung 4.32	Die Benchmarks nach Landis et al. (1997) zur Einordnung von Cohen's Kappa	72
Abbildung 4.33	Tabelle zur Berechnung von κ : Neben den Phasen wird auch die (Nicht-) Übereinstimmung bestimmt, wenn ein Individuum <i>nicht aktiv</i> (NA) ist	72
Abbildung 4.34	Klassifizierung der Interaktionen: Identische Kodierung (grün), Interaktionseinheit liegt im Toleranzintervall (blau), Interaktionseinheit liegt außerhalb des Toleranzbereichs (rot), es wurden unterschiedliche Phasen kodiert (pink) oder es wurde eine unterschiedliche Anzahl an Interaktionseinheiten kodiert (lila)	73
Abbildung 5.1	Der fünfschrittige Lösungsplan (Beckschulte 2019, S. 79)	92
Abbildung 6.1	Das erste Durchlaufen des Modellierungskreislaufs beim Färben der Deutschlandkarte	106
Abbildung 6.2	Das zweite Durchlaufen des Modellierungskreislaufs beim Färben der Deutschlandkarte	108
Abbildung 6.3	Der Fragebogen aus der ersten Erhebungsphase, indem die Vorerfahrungen zum Modellieren abgefragt werden	114

Abbildung 6.4	Die Lernenden erstellen auf jedem Fragebogen einen anonymisierten Personalcode	114
Abbildung 6.5	Im Post-Fragebogen wurde mithilfe soziometrischer Fragen die Gruppenstruktur erfasst	116
Abbildung 6.6	Das Studiendesign der empirischen Studie. G1 stehen für die Gruppen, die eine Unterweisung erhalten haben, G2 für die Gruppen ohne Unterweisung	117
Abbildung 6.7	Das Kodierschema der Studie	123
Abbildung 6.8	Ein Soziogramm einer Gruppe. Positive Beziehungen werden durch eine gerichtete Kante, negative Beziehungen mit einer gestrichelten gerichteten Kante dargestellt	131
Abbildung 6.9	Auswertung der Gruppenstruktur und der Interaktionen	132
Abbildung 7.1	Vergleich der Strukturmerkmale von Instruierten (pink) und Nicht-Instruierten (orange). Es werden die relativen Anteile bzw. die relative Dauer der Mediane gezeigt	149
Abbildung 8.1	Vergleich des relativen Anteils und der relativen Dauer (jeweils Medianwerte) von Interaktionen von Lernenden, die eine Person am liebsten (1; blau), neutral (0; lila) und ungern (−1; pink) in der Freizeit treffen	152
Abbildung 8.2	Vergleich des relativen Anteils und der relativen Dauer (jeweils Medianwerte) von Interaktionen von Lernenden, die am liebsten (1; blau), neutral (0; lila) und ungern (−1; pink) neben einer Person im Unterricht nebeneinandersitzen	153
Abbildung 8.3	Vergleich des relativen Anteils und der relativen Dauer (jeweils Medianwerte) von Interaktionen von Lernenden, die am liebsten (1; blau), neutral (0; lila) und ungern (−1; pink) mit einer Person im Unterricht zusammenarbeiten	153

Abbildung 8.4	Vergleich des relativen Anteils und der relativen Dauer (jeweils Medianwerte) von Interaktionen von Lernenden, die am liebsten (1; blau), neutral (0; lila) und ungern (−1; pink) mit einer Person im Mathematikunterricht zusammenarbeiten	154
Abbildung 11.1	ML benutzt Eingabedaten und lernt mit Algorithmen, um ein Modell zu erstellen	176
Abbildung 11.2	Veranschaulichung von Under- (a) und Overfitting (b) sowie eine geeignete Anpassung (c). Die Datenpunkte sind blau und die Modellanpassung rot gefärbt	178
Abbildung 12.1	Veranschaulichung des KNN für $k = 3$	185
Abbildung 12.2	Die Hyperebene, die die Datenpunkte in zwei Klassen einteilt, ist (zunächst) nicht eindeutig	187
Abbildung 12.3	Die Supportvektoren der Hyperebene H sind x_1 und x_2	188
Abbildung 12.4	Die soft margin SVM, bei der Datenpunkte im Datenstreifen liegen können	192
Abbildung 12.5	Die SVM mit nichtlinearen Datenpunkten und nichtlinearer Trennlinie	193
Abbildung 12.6	Die logistische Funktion für $\beta_0 = 0$ und $\beta_1 = 3$	194
Abbildung 12.7	Die Train/Test-Split-Methode: Die Datenpunkte N werden in zwei Folds N_t (Trainingsdaten) und N_v (Testdaten) unterteilt, wobei das Modell mit N_t entwickelt wird	196
Abbildung 12.8	k -Kreuzvalidierung: Der Datensatz wird in k Folds aufgeteilt, wobei jeweils ein Fold zum Testen verwendet wird. Zuerst erfolgt die Modellentwicklung, danach die Validierung	199
Abbildung 12.9	Die <i>verschachtelte Kreuzvalidierung</i> : In der inneren Validierung wird das Modell trainiert und in der äußeren Validierung getestet	201
Abbildung 12.10	Überblick über die Verfahren der Dimensionsreduktion	202
Abbildung 12.11	Ablauf bei der Filtermethode: Zuerst wird die Anzahl der Features reduziert, bevor das Modell trainiert und bewertet wird	203

Abbildung 12.12	Ablauf der Wrappermethode: Die Features werden während des Trainings des Modells ausgewählt	204
Abbildung 12.13	Ablauf der eingebetteten Methode: Die Feature Selection ist Bestandteil des Klassifikationsalgorithmus	204
Abbildung 13.1	Darstellung der wiederholten stratifizierten verschachtelten Kreuzvalidierung. Als Klassifikationsalgorithmen werden KNN, SVM und logReg implementiert	209
Abbildung 13.2	Bei der Implementierung der Klassifikation (hier: KNN) werden zu Beginn Spalten des Datensatzes der Features und die Spalte der Label festgelegt. Außerdem werden die Anzahl der Folds und Wiederholungen der äußeren Validierung definiert. In der inneren Validierung wird die Menge der Anzahl der Features und der Hyperparameter festgelegt sowie die Gridsearch definiert	210
Abbildung 13.3	Die wiederholte stratifizierte verschachtelte Kreuzvalidierung im Programmcode. Die Wahl des Parameters $p = 2$ und der Minkowski-Metrik ergeben die euklidische Distanz	211
Abbildung 13.4	Die Vorhersagemetriken werden für die Testdaten ermittelt	211
Abbildung 13.5	Vergleich der Klassifikationsalgorithmen KNN (grün), SVM (rot) und logReg (blau) anhand der Vorhersagemetriken Accuracy (a) und Precision (b): Bewertung der Klassifikationsgenauigkeit für jede Anzahl von Features	213
Abbildung 13.6	Vergleich der Klassifikationsalgorithmen KNN (grün), SVM (rot) und logReg (blau) anhand der Vorhersagemetriken Accuracy (a) und Precision (b): Bewertung der Klassifikationsgenauigkeit für jede Anzahl von Wiederholungen in der äußeren Validierung	214