

Lern- und Kooperationsprozesse. Warum sind sie interessant und wie können sie analysiert werden?

1. Lernprozesse beim fallbasierten Lernen

Kooperativem Lernen wird ein großes Potenzial für die Vermittlung von theoriebasiertem Handlungswissen zugeschrieben. Dies gilt insbesondere für kooperatives fallbasiertes Lernen in der Weiterbildung. Damit wird die Hoffnung verknüpft, dass Lernende die Fähigkeit erlangen, die theoretisch relevanten Merkmale einer Situation zu erkennen und wissenschaftliche Erkenntnisse darauf anzuwenden. Diskussionen mit anderen Lernenden über Fallanalysen sollen dies unterstützen. Das Potenzial der Kooperation wird dabei zum Beispiel dem Externalisieren der Gedanken zur Fallanalyse gegenüber dem Lernpartner, dem gegenseitigen Erklären und Begründen und dem kritischen Nachfragen zugeschrieben (vgl. King 2007, S. 18ff.). Inwiefern kooperatives fallbasiertes Lernen erfolgreich ist, hängt also davon ab, inwieweit solche Lernaktivitäten auftreten bzw. durch instruktionale Unterstützung angeregt werden.

Damit rücken die Lernprozesse in den Fokus – vor allem, wenn die Evaluation keinen eindeutig positiven Befund bezüglich der Lernergebnisse liefert. Dann liegt häufig die Schlussfolgerung nahe, dass die instruktionale Unterstützung gänzlich wirkungslos war. Die Analyse von Lern- und Kooperationsprozessen bietet hingegen neben der Möglichkeit der Erklärung positiver Effekte von instruktionalen Maßnahmen auf den Lernerfolg auf der Grundlage von Lernmechanismen (vgl. Stegmann u.a. 2012, S. 320; Wecker/Fischer 2011, S. 754) auch die Möglichkeit, Schwachstellen der instruktionalen Unterstützung zu identifizieren. In Bezug auf die *Qualität der Argumentation* tendieren Lernende zum Beispiel dazu, Behauptungen aufzustellen, ohne diese zu begründen (vgl. Stegmann u.a. 2012, S. 316f.). Ein weiteres Problem ist die häufig geringe *Transaktivität*: Lernende nehmen kaum Bezug auf Beiträge anderer Lernender und bauen nicht auf den Argumenten ihrer Lernpartner auf (vgl. Fischer u.a. 2011, S. 328).

Die Annahme, dass eine instruktionale Unterstützung etwa der Qualität der Argumentation oder der Transaktivität einen positiven Effekt auf den Wissenserwerb hat, lässt sich als „Hypothesendreieck“ beschreiben (vgl. Abb. 1). Der Effekt der instruktionalen Unterstützung auf den Lernerfolg (Pfeil c in Abb. 1) wird mit mindestens zwei weiteren Annahmen auf einen Lernmechanismus zurückgeführt: Die erste dieser beiden Annahmen besagt, dass bestimmte (Lern-)Aktivitäten positiv bzw. negativ mit dem Lernerfolg zusammenhängen (Pfeil b in Abb. 1). Die zweite Annahme besagt, dass die instruktionale Unterstützung einen positiven oder negativen Effekt auf diese (Lern-)Aktivitäten hat (Pfeil a in Abb. 1).

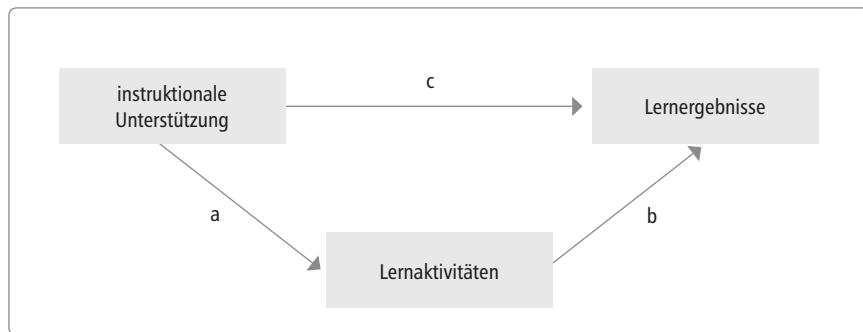


Abbildung 1: Das „Hypothesendreieck“ bei der Analyse von Lernprozessen

In diesem Beitrag wird eine Vorgehensweise beschrieben, mit der sich ein derartiges Gefüge von Annahmen überprüfen lässt. Es handelt sich dabei um eine Methodik, mit der qualitative Analysen im Sinne eines Mixed-methods-Ansatzes einer statistischen Analyse zugänglich gemacht werden können. Als Ausgangspunkt eignen sich dabei prinzipiell beliebige qualitative Auswertungsmethoden, sofern die eingesetzten Codierschemata strukturell mit der hier beschriebenen Vorgehensweise übereinstimmen. Zur Illustration greifen wir dabei auf ein Forschungsprojekt zum fallbasierten Lernen in Online-Diskussion in der universitären Pädagogikausbildung zurück. In diesem Projekt wurde instruktionale Unterstützung u.a. zur Förderung der Qualität der Argumentation und der Transaktivität untersucht (vgl. Fischer u.a. 2011, S. 330). Die Aufgabe der Studierenden bestand darin, mit Hilfe einer Theorie zur Leistungsmotivation drei Probleme zu analysieren, d.h. Diagnosen zu formulieren, Prognosen zu erstellen und Maßnahmenempfehlungen abzugeben.

Mit Hilfe der Methode des lauten Denkens wurden die kognitiven Aktivitäten der Lernenden während der Fallanalyse zugänglich gemacht (vgl. Stegmann u.a. 2012, S. 308, S. 309f.). Dabei wurden die Lernenden dazu aufgefordert, ihre Gedanken unkommentiert laut auszusprechen (vgl. Ericsson/Simon 1984, S. 80–82). Die Beiträge zur Online-Diskussion über die Fallanalysen wurden u.a. bezüglich der Qualität der Argumentation und ihrer Transaktivität analysiert. Obwohl der Schwerpunkt mit diesem Beispiel auf kognitive Lernaktivitäten gelegt wird, sind beispielsweise gruppendynamische bzw. sozialpsychologische Mechanismen strukturell analog analysierbar.

Im Folgenden wird nach einem kurzen Überblick über geeignete Erhebungs-techniken sowie einer Charakterisierung der typischen Datenstruktur bei derartigen Analysen die Vorgehensweise bei der Analyse beschrieben. Dabei wird zunächst die Codierung der Lernaktivitäten thematisiert. Die automatische Analyse sprachlicher Daten wird als Alternative dazu diskutiert. Anschließend werden Varianten der Aufbereitung der Codierdaten erörtert. Zuletzt wird auf verschiedene Möglichkeiten der statistischen Analyse eingegangen.

2. Erhebungstechniken und Datenstruktur

Als Erhebungstechniken steht die gesamte Palette üblicher Verfahren zur Verfügung: von der Nutzung natürlicher *Produkte von Lernaktivitäten* (Artefakte, zumeist Textdaten wie Notizen, Analysen von Fällen oder Lerntagebücher) über *Audio- und Videoaufzeichnungen* der Interaktionen von Personen in Face-to-face-Situationen oder von lautem Denken zur Erfassung kognitiver Lernaktivitäten bis hin zur *Blickbewegungsaufzeichnung* (Eye-Tracking) zur Erfassung der kognitiven Verarbeitung visuell dargebotener Informationen. Bei technologieunterstütztem Lernen können darüber hinaus *Logfiles* sowie *Bildschirmvideos* aufgezeichnet und analysiert werden.

Die dabei anfallenden Daten bilden zusammen mit weiteren Daten über die Lernenden eine Mehrebenenstruktur (vgl. Abb. 2): Jeder Person ist in der Regel eine variable Anzahl einzelner erfasster Lernaktivitäten zuzuordnen. Analog sind häufig mehrere einzelne Lernende einer übergeordneten sozialen Einheit zuzuordnen, etwa einer Lerngruppe oder einem Kurs.

In der beschriebenen Beispielstudie zum fallbasierten Lernen in Online-Diskussionen steuert eine Person als Mitglied einer bestimmten Dreiergruppe im Verlauf des Kurses normalerweise mehrere Beiträge zu Diskussionen über Fallanalysen bei und äußert beim lauten Denken eine Vielzahl einzelner Gedanken.

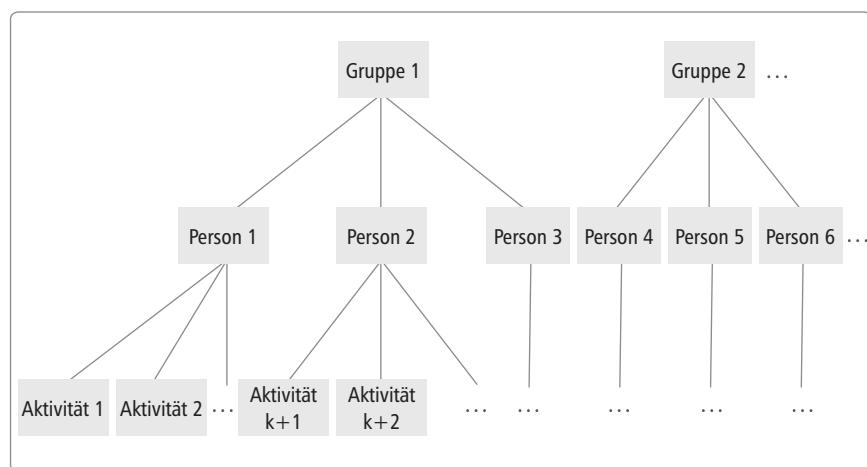


Abbildung 2: Hierarchische Datenstruktur bei der Analyse von Lernprozessen

3. Analyse

3.1 Codierung

3.1.1 Entwicklung eines Codierschemas

Für die Analyse von Einflüssen auf Lernaktivitäten sowie die Rolle von Lernaktivitäten für den Lernerfolg müssen die einzelnen Lernaktivitäten codiert werden. Bei der Entwicklung eines Codierschemas ist zunächst zu klären, welche Analyseeinheiten codiert werden sollen. Dabei ist von den Ereignissen auszugehen, die den Hypothesen zufolge für den Lernerfolg entscheidend sind. Auf dieser Grundlage wird die Art der Einheiten bestimmt, unter denen diese kritischen Ereignisse zu suchen sind. Jeder einzelnen auszuwertenden Analyseeinheit entspricht üblicherweise eine Zeile in einer Datentabelle für die Lernaktivitäten.

In der Studie zum fallbasierten Lernen in Online-Diskussionen besagte eine Hypothese, dass eine hohe Transaktivität der Diskussion den Wissenserwerb fördert. Eine Art von für den Lernerfolg kritischen Ereignissen dürften demnach transaktive Beiträge zur Diskussion sein. Somit können die einzelnen Beiträge als Analyseeinheiten herangezogen werden, die bezüglich des Vorkommens transaktiver Äußerungen codiert werden sollen.

Bezüglich der Abgrenzung der einzelnen Analyseeinheiten voneinander gibt es verschiedene Möglichkeiten. Wenn sie bereits natürlich abgegrenzt vorliegen, wie im Fall der Diskussionsbeiträge, können diese direkt als zu codierende Segmente verwendet werden. Wenn keine derartige natürliche Abgrenzung vorliegt, muss das erhobene Datenmaterial zunächst segmentiert werden – etwa bei den Protokollen lauten Denkmens, die mit Hilfe von Audio- und Bildschirmaufzeichnungen erfasst wurden. Dabei können entweder linguistische Einheiten wie Sätze, aber auch psychologische, wie etwa „Gedankenstränge“, mit Hilfe von Segmentierregeln voneinander abgegrenzt werden, oder es kann ein Zeitraster mit feststehender Segmentdauer angelegt werden. Bei der Nutzung eines Zeitrasters hat sich in vielen Fällen eine Segmentlänge von zehn Sekunden bewährt (vgl. Seidel/Prenzel 2010, S. 146). Bei der Entscheidung für eine dieser beiden Segmentierungstechniken sind der mögliche Aufwand – einschließlich einer ggf. erforderlichen Transkription im Vergleich zur rein computergestützten Codierung mit geeigneten Tools wie z.B. Videograph (vgl. ebd., S. 150) – sowie die erforderliche Genauigkeit zu berücksichtigen.

An die Entscheidung über das Segmentierungsverfahren ist die über die Anfertigung von Transkriptionen eng gekoppelt. Dabei sind ebenfalls der Aufwand sowie der mit der Transkription einhergehende Verlust von bedeutungsrelevanten Informationen (z.B. Intonation oder Gestik) zu berücksichtigen. Beträchtliche Vorteile von Transkriptionen sind dagegen der bessere Überblick bei der späteren Codierung sowie die Entkopplung der Analysedauer von der Dauer der Audio- oder Videoaufzeichnungen. Während bei der Codierung ohne Transkription mindestens die Dauer der Audio- bzw.

Videoaufzeichnung benötigt wird, können Transkriptionen ggf. auch deutlich schneller codiert werden. In der Beispielstudie wurden die kognitiven Aktivitäten zur Bearbeitung des Falls anhand der Protokolle lautend Denkens auf der Grundlage von Segmenten mit einer feststehenden Länge von zehn Sekunden ohne Transkription analysiert.

Auch die Aufstellung von Codierdimensionen ist eng mit der Bestimmung der Analyseeinheiten verknüpft (vgl. Abb. 3). Eine Dimension ist eine Variable, d.h. eine Menge möglicher Werte (hier: Codes). Diese müssen sich gegenseitig ausschließen und insgesamt erschöpfend sein – d.h. die Codes müssen so gewählt werden, dass sie nicht gleichzeitig auf eine Analyseeinheit zutreffen können, und jeder Analyseeinheit muss ein Code zugeordnet werden können. Jeder Codierdimension entspricht üblicherweise eine Spalte in der Datentabelle für die Lernaktivitäten. Jeder Analyseeinheit wird bei der Auswertung auf jeder auf sie anwendbaren Codierdimension ein Wert zugewiesen.

Die Codierdimensionen können verschiedenen Begriffstypen entsprechen und somit unterschiedliche Skalenniveaus aufweisen. Die beiden häufigsten Typen sind klassifikatorische (Nominalskalenniveau) und komparative bzw. metrische Dimensionen (Ordinal- oder Intervallskalenniveau; vgl. Seidel/Prenzel 2010, S. 146f.). Letztere werden oft auch als „Rating-Dimensionen“ bezeichnet. Sowohl klassifikatorische Codierdimensionen als auch Rating-Dimensionen können „niedriginferent“ oder „hochinferent“ sein, d.h. an unmittelbar feststellbare Merkmale gekoppelt sein oder stärker interpretative Schlussfolgerungen erfordern (vgl. ebd., S. 147).

Die verschiedenen Dimensionen eines Codierschemas müssen jedoch nicht demselben Typ angehören, sondern können gemischt auftreten. Zum Beispiel kann eine Äußerung in einer Diskussion über einen Fall sowohl auf der Dimension der Qualität der Argumentation (z.B. ob die Hauptthese begründet wird) klassifikatorisch codiert als auch hinsichtlich ihrer Transaktivität auf einer Skala von 1 bis 5 bewertet werden.

Die Ausprägungen der Codierdimensionen stellen die Codes dar, die den einzelnen Analyseeinheiten auf den Codierdimensionen zugewiesen werden (vgl. Abb. 3). In der Beschreibung des Codierschemas sollten unter jeder Codierdimension alle einzelnen Codes aufgelistet werden. Die genauen Anwendungskriterien für die einzelnen Codes sollten so klar und einfach wie möglich beschrieben werden (vgl. Abb. 3). Gibt es in den Daten viele Analyseeinheiten, die sich nur schwer zuordnen lassen und die für die Fragestellung von keinem besonderen Interesse sind, kann ein Code mit der Bezeichnung „Sonstiges“ aufgenommen werden. Hilfreich für die Erläuterung der Anwendungskriterien ist außerdem die Aufnahme von authentischen Beispielen (insbesondere von Grenzfällen) aus den zu codierenden Daten unter Verweis auf ihre Herkunft (vgl. Abb. 3).

Bei Rating-Dimensionen sollten jeweils entweder alle möglichen Rating-Werte oder die äußersten möglichen Werte (im Beispiel oben zu Transaktivität: 1 und 5) aufgeführt werden. Es ist zwar möglich, allen einzelnen Rating-Werten Kriterien zuzuordnen, doch besteht der Vorteil einer Rating-Dimension häufig gerade in einer einfachen Einschätzung, die bereits anhand von Kriterien oder Beispielen für einzelne Werte ver-

ankert werden kann. Ein Item für das Rating des Ausmaßes an Transaktivität könnte z.B. den Wortlaut „Die Gruppe diskutierte intensiv unterschiedliche Sichtweisen“ aufweisen, wobei die möglichen Werte von 1 (trifft überhaupt nicht zu) bis 5 (trifft voll und ganz zu) reichen könnten.

Dimension: Qualität der Argumentation

Anwendungsbedingung: Auf dieser Dimension wird codiert, wenn es sich um einen inhaltlich relevanten Beitrag handelt (d. h. kein „Off-topic“-Beitrag).

Codes:

- Code: Behauptung (1)

Kriterium: In der Äußerung wird eine bestimmte Position in Form einer These zum Ausdruck gebracht.

Beispiele:

- „Maria Steiner attribuiert Erfolg external variabel.“ (VPn 256, Seg. 32)
- „Maria wird es in Ihrem Studium nicht weit bringen.“ (VPn 172, Seg. 12)

- Code: Beleg (2)

Kriterium: In der Äußerung wird Information zur Untermauerung der These angeführt.

Beispiele:

- „Sie führt ihre gute Prüfungsnote darauf zurück, dass die Lehrerin so kurz vor den Ferien gut gelaunt war.“ (VPn 256, Seg. 33)
- „Ich habe auch einmal ein Mädchen gekannt, bei dem war das genauso.“ (VPn 172, Seg. 13)

- ...

Dimension: Transaktivität

...

Abbildung 3: Auszug aus einem Codierschema für die Analyse von Lernprozessen

Eine Besonderheit mancher Codierdimensionen besteht darin, dass sie nur auf Analyseeinheiten angewandt werden, die bestimmte Bedingungen erfüllen, d.h. bei denen zuvor auf einer anderen Codierdimension ein bestimmter Code vergeben wurde. Diese Anwendungsbedingungen sollten im Codierschema eindeutig unter Verweis auf diese andere Codierdimension und den (oder die) dort vorausgesetzten Code(s) angegeben werden. So wird im Beispiel die Codierung der Qualität der Argumentation auf Diskussionsbeiträge mit inhaltlichem Bezug zum Fall beschränkt (vgl. Abb. 3).

3.1.2 Training der Codierenden und Bestimmung der Objektivität

Der Einsatz eines Codierschemas erfordert stets ein hinreichend intensives Training der Codierenden, das neben einer *Einführung in das Codierschema* mehrere Zyklen, bestehend aus *unabhängiger Codierung einer Trainingsstichprobe, Ermittlung der Objektivität und Besprechung von Abweichungen*, umfasst. Diese werden so lange wiederholt, bis auf allen in der späteren Analyse zu verwendenden

Codierdimensionen zufriedenstellende Kennwerte für die Objektivität erzielt werden. Dabei wird das Codierschema nach Bedarf modifiziert oder ergänzt.

Die Durchführung des Trainings wird möglichst genau dokumentiert und in Publikationen im Methodenteil beschrieben. Wichtige Informationen sind dabei die Größe der einzelnen Trainingsstichproben, die Anzahl der Durchgänge und der Zeitaufwand (getrennt für Trainingscodierungen und Besprechungen) sowie die am Codierschema vorgenommenen Änderungen (vgl. De Wever u.a. 2006, S. 11).

Nach Abschluss des Trainings wird die endgültig erzielte Objektivität bestimmt. Zu diesem Zweck wird eine weitere Probecodierung durchgeführt, zu der keine Besprechungen mehr stattfinden. Die Größe der Stichprobe von Analyseeinheiten bei dieser Probecodierung (Anzahl bzw. Anteil der Segmente) wird wiederum dokumentiert und in Publikationen grundsätzlich berichtet (vgl. ebd.). Üblich sind 10 bis 20 Prozent der Daten, die in die statistische Analyse eingehen.

Für die Daten, die für das Training und die Bestimmung der Objektivität verwendet werden, gilt:

1. Sofern verfügbar, werden Daten verwendet, die später nicht in die eigentliche Analyse einbezogen werden (z.B. aus Pilotstudien bzw. fehlgeschlagenen Erhebungen).
2. Die Stichprobe von Analyseeinheiten umfasst gleich große Anteile aus allen experimentellen Bedingungen und deckt die gesamte Breite an möglichen Merkmalen ab.
3. Die Codierenden erfahren nicht, aus welcher die Daten stammen.

Die anwendbaren Kennwerte für die Objektivität richten sich nach dem Typ bzw. Skalenniveau der einzelnen Codierdimensionen. Im Folgenden geben wir einen Überblick über die gängigsten jeweils einschlägigen Kennwerte; für die Berechnung verweisen wir auf die umfassende Darstellung von Wirtz und Caspar (2002).

Bei klassifikatorischen Codierdimensionen ist derzeit die Angabe der *prozentualen Übereinstimmung* sowie des zufallskorrigierten Maßes *Cohens Kappa Standard* (vgl. ebd., S. 56–59). Unter bestimmten Umständen – etwa wenn die Übereinstimmung von mehr als zwei Codierenden in einem einzigen Kennwert ausgedrückt werden soll – sind *Fleiss' Kappa* (vgl. ebd., S. 75f.) sowie *Krippendorffs Alpha* geeignet. Bei komparativen Rating-Dimensionen eignet sich das *gewichtete Kappa* (vgl. ebd., S. 78–82) sowie bei metrischen Rating-Dimensionen die *Intraklassenkorrelation* in ihren verschiedenen Varianten (ebd., Kap. 6.1).

Für jede einzelne klassifikatorische Codierdimension wird jeweils ein eigener Wert für das verwendete Objektivitätsmaß berechnet. Idealerweise werden die einzelnen Objektivitätswerte für alle Codierdimensionen im Methodenteil eines Forschungsberichts angegeben (vgl. De Wever u.a. 2006, S. 11), zumindest jedoch ihr Minimum und Maximum sowie Mittelwert oder Median. Ab 0,4 gelten Objektivitätswerte als „akzeptabel“, ab 0,6 als „gut“ und ab 0,75 als „exzellent“ (vgl. Wirtz/Caspar 2002, S. 59). Hauptfolgen einer geringen Objektivität sind zugunsten der Nullhypothese verzerzte

inferenzstatistische Tests, da die Fehlervarianz in den Daten vergleichsweise hoch ist. Dies bedeutet, dass eine geringe Objektivität letztlich weniger problematisch ist, wenn die statistischen Hypothesentests dennoch Signifikanz anzeigen.

3.1.3 Durchführung der Codierung

Bei der eigentlichen Durchführung der Codierung sind verschiedene Formen der Arbeitsaufteilung möglich. Beispielsweise kann das gesamte Datenmaterial von einer einzigen Person ausgewertet werden, sobald deren Analyseergebnisse ausreichend objektiv sind. Ein entscheidender Vorteil besteht dabei darin, dass die Fehlervarianz in den Codierungen gering ausfällt, da die Auswertungsergebnisse frei von systematischen Unterschieden zwischen Codierenden sind. Leider ist diese eigentlich optimale Variante häufig aufgrund der zu codierenden Datenmenge nicht praktikabel. Bei mehreren Codierenden werten dagegen alle auf jeden Fall aus jeder Bedingung denselben Anteil der Daten aus, um zumindest den internen Aspekt der Validität der Untersuchung nicht zu beeinträchtigen.

3.2 Automatische Codierung sprachlicher Daten als Alternative

Wenn die zu analysierenden Daten in digitaler Textform vorliegen, bietet sich ab einem gewissen Datenumfang auch der Einsatz von Technologien der automatischen Codierung sprachlicher Daten an. Auf der Grundlage segmentierter und bereits codierter Daten werden mit Hilfe von Ansätzen des *machine learning* Algorithmen „trainiert“. Zur Bestimmung der Übereinstimmung mit den menschlichen Codierenden werden üblicherweise 90 Prozent der von Menschen codierten Daten als Trainingsmaterial verwendet und die Objektivität anhand der restlichen zehn Prozent bestimmt. Die trainierten Algorithmen können dann für die automatische Codierung beliebiger Mengen von Daten mit hoher Objektivität eingesetzt werden.

Diese Verfahren beruhen nicht auf definierten Regeln des Analyseschemas, sondern sie bestimmen mit Hilfe zuvor geschätzter probabilistischer Parameter auf der Grundlage bestimmter Texteigenschaften (z.B. Satzlänge, Häufigkeit von Wörtern und Wortkombinationen) den wahrscheinlichsten Code für ein Segment. Daher kann es vorkommen, dass Codes auf der Basis von Oberflächenmerkmalen falsch zugeordnet werden. So könnte zum Beispiel ein sehr langes Segment als kritische Äußerung codiert werden, oder das Wort „weil“ zum Code „Begründung“ führen, wenn eindeutige Merkmale fehlen, die gegen diese Codes sprechen. Üblicherweise ist der Anteil so entstehender Fehler jedoch nicht größer als der Fehleranteil aufgrund der schwanken- den Leistung menschlicher Codierender. Eine detaillierte Anleitung zum Training von Algorithmen zur Analyse natürlich-sprachlicher Daten findet sich bei Mu u.a. (2012).

Für Daten aus der hier verwendeten Beispielstudie konnten die Qualität der Argumentation, die inhaltliche Fallbearbeitung und auch die Transaktivität von Beiträgen in Gruppendiskussionen mit zufriedenstellender Objektivität teil-automatisch codiert werden (vgl. Rosé u.a. 2008, S. 262f.). Zuvor war jedoch die Segmentierung der Daten

durch menschliche Codierende erforderlich. Inzwischen wurden allerdings Ansätze entwickelt, bei denen auch die Segmentierung automatisch erfolgt (z.B. das ACODEA-Framework, vgl. Mu u.a. 2012, S. 299f.).

3.3 Aufbereitung der Codierdaten

Auf der Grundlage der codierten Daten können die zu prüfenden Hypothesen getestet werden. Dafür ist zu entscheiden, ob und gegebenenfalls wie diese Daten weiter verrechnet werden sollen. Prinzipiell können entweder die Daten auf der Ebene der Lernaktivitäten direkt verwendet werden, oder es kann eine Aggregation von der Ebene der Lernaktivitäten auf die Ebene der Personen oder sogar Gruppen durchgeführt werden. Bei Letzterer ist jedoch in der Regel der Verlust möglicherweise wichtiger Informationen über die Details des zu untersuchenden Mechanismus in Kauf zu nehmen.

				Skalenbildung			
Person	Beitrag	Trans-aktivität	Qualität d. Arg.	Person	mittlere Trans-aktivität	Anteil Belege (1)	Argument-Trans-aktivität
1	1	4	0	1	3,4	0,4	0,5
1	2	3	2	2
1	3	5 Aggregation 1	
1	4	1	0				
1	5	4	1				
2	6	2	0				
2	7	1	0				
2				
...				

Skalenbildung im Beispiel:

$$\text{Argument-Transaktivität} = ((\text{mittlere Transaktivität} - 1) / 4 + \text{Anteil Belege}) / 2$$
|
mögliche Spanne der Ratings
mögliches Minimum der Ratings

$$\text{Argument-Transaktivität} = ((3,4 - 1) / 4 + 0,4) / 2 = 0,5$$

Abbildung 4: Aggregation und Skalenbildung

Bei der Aggregation werden die Datensätze zu den verschiedenen Lernaktivitäten einer Person zu einem einzigen Datensatz zusammengefasst (vgl. Abb. 4). Zu diesem Zweck müssen die Ausprägungen der Variablen bei den einzelnen Lernaktivitäten jeweils durch einen einzigen statistischen Kennwert ausgedrückt werden.

Bei klassifikatorischen Codierdimensionen werden häufig recht einfache Kennwerte gebildet: Beispielsweise kann eine aggregierte Variable schlicht die Information enthalten, ob ein bestimmtes Ereignis (z.B. ein „Aha-Erlebnis“) im Verlauf der Lernphase überhaupt vorgekommen ist (0 vs. 1). Weitere Möglichkeiten sind die Bildung der absoluten oder relativen Häufigkeit einzelner Ausprägungen der Codierdimensionen (vgl. Abb. 4). Welche Variante verwendet wird, hängt vom angenommenen Lernmechanismus ab.

Bei Rating-Dimensionen besteht die nächstliegende Aggregierungsmethode in der Ermittlung eines Maßes der zentralen Tendenz der Ratings der Lernaktivitäten bei jeder einzelnen Person.

Unabhängig von der Aggregation, d.h. auf der nicht aggregierten Ebene der Lernaktivitäten wie der aggregierten der Personen, können verschiedene einzelne Variablen auch zu einer Skala zusammengefasst werden (vgl. Abb. 4). Dies geschieht zumeist per Summen- oder Mittelwertsbildung. Beispielsweise könnte etwa aus verschiedenen aggregierten Ratings ein Mittelwert für die Transaktivität einer Person gebildet werden. Im Fall einer derartigen Skalenbildung wird ein Kennwert für deren Reliabilität – gewöhnlich Cronbachs Alpha – angegeben.

3.4 Statistische Analyse

In Bezug auf die statistische Analyse können hier nur einige Hinweise auf grundsätzliche Möglichkeiten gegeben werden. Bezüglich der Durchführung verweisen wir auf die einschlägige Literatur zu den verschiedenen Verfahren.

Neben der Prüfung des Effekts von instruktionalen Rahmenbedingungen auf den Lernerfolg (vgl. Abb. 1, Pfeil c) kommt bei der Analyse von Lernmechanismen dem Test des Effekts der instruktionalen Rahmenbedingungen auf die Lernaktivitäten (Pfeil a) sowie des Zusammenhangs von Lernaktivitäten und Lernerfolg (Pfeil b) eine besondere Bedeutung zu. Der Zusammenhang von Lernaktivitäten und Lernerfolg kann mittels Regressions- bzw. Korrelationsanalysen mit dem Lernerfolg als Kriterium und aggregierten Indikatoren für Lernaktivitäten als Prädiktoren analysiert werden. Der Effekt von instruktionalen Rahmenbedingungen auf die Lernaktivitäten (Pfeil a) kann auf zwei Arten analysiert werden: Entweder werden verschiedene experimentelle Bedingungen hinsichtlich aggregierter Indikatoren für Lernaktivitäten verglichen, oder es werden, im Rahmen von Mehrebenenanalysen, unaggregierte Indikatoren für Lernaktivitäten als Kriterien sowie weitere Variablen auf beliebigen Ebenen als Prädiktoren verwendet (Raudenbush/Bryk 2002, Kap. 6 und 10; für ein Anwendungsbeispiel vgl. Wecker/Fischer 2011, S. 752f.).

Im zweiten Fall können wesentlich differenziertere Hypothesen über kausale Bedingungen für das Auftreten bestimmter Lernaktivitäten oder -ereignisse überprüft werden: So könnte beispielsweise das Auftreten von Lernaktivitäten mit bestimmten Merkmalen im Zeitverlauf in Abhängigkeit von Merkmalen der Personen oder Gruppen untersucht werden. Auf diese Weise können spezifische Annahmen über genau umrissene Lernmechanismen auf der Grundlage kausaler Zusammenhänge zwischen Ereignissen und Lernaktivitäten direkt getestet werden.

Im Fall der Analyse auf Personenebene kann dagegen mit Hilfe von Mediatoranalysen überprüft werden, inwiefern das Ausmaß bestimmter Lernaktivitäten (aggregiert auf Personenebene) den Effekt der Intervention auf den Lernerfolg vermittelt (vgl. Baron/Kenny 1986, S. 1176f.). Die simultane Überprüfung des gesamten Hypothesen-gefüges ist mit Hilfe von Pfadanalysen bzw. Strukturgleichungsmodellen möglich (Nachtigall u.a. 2003; für ein Anwendungsbeispiel vgl. Stegmann u.a. 2011, S. 239f.). Im gewöhnlichen Fall liegen dabei ebenfalls alle einbezogenen Variablen auf derselben Ebene, d.h. es werden wiederum aggregierte Merkmale der Lernaktivitäten verwendet.

det. In Mehrebenen-Pfadanalysen bzw. -Strukturgleichungsmodellen können darüber hinaus Zusammenhänge zwischen Variablen auf verschiedenen Ebenen berücksichtigt werden, sofern dabei keine Variable auf einer höheren Ebene als ihre unmittelbaren Prädiktoren angesiedelt ist. So kann z.B. untersucht werden, ob sich die Ähnlichkeit des Vorwissens innerhalb einer Gruppe auf die Transaktivität der Beiträge auswirkt.

4. Fazit

Die differenzierte Analyse von Lern- und Kooperationsprozessen ermöglicht es, theoretische Annahmen zum Lernen von Erwachsenen und zur Wirkung von unterschiedlichen Rahmenbedingungen zu überprüfen. In diesem Beitrag wurde eine Vorgehensweise vorgestellt, in der Lern- und Kooperationsprozesse zunächst qualitativ mit Codierschemata analysiert werden, ehe dann die Ergebnisse dieser Analyse im Rahmen quantitativer Methoden Verwendung finden. Auch Ansätze zur Teilautomatisierung der Analyse wurden vorgestellt. Automatische Analysen von Lern- und Kooperationsprozessen werden derzeit intensiv beforscht und gelten (unter dem Sammelbegriff „Learning Analytics“) als wichtiger Zukunftstrend.

Die hier verwendeten Beispiele zum fallbasierten Lernen Erwachsener sind dem formalen Bildungskontext des Universitätsstudiums entnommen. In diesen Studien zeigte sich an verschiedenen Stellen, dass der Aufwand für die Analysen der Lernaktivitäten mehr als gerechtfertigt war. So zeigte z.B. eine Intervention zur Förderung der Qualität der Argumentation keinen Effekt auf den inhaltlichen Wissenserwerb (vgl. Fischer u.a. 2011, S. 331). Die Analysen der Lernaktivitäten zeigten jedoch, dass deren Förderung erfolgreich war und dass diese substanzell mit dem Wissenserwerb zusammenhingen. Daraus lässt sich die Vermutung ableiten, dass eine längerfristige Förderung der Lernaktivitäten auf den Wissenserwerb durchschlagen könnte. Jüngere Projekte aus der Weiterbildungsforschung belegen ferner, dass die in diesem Beitrag skizzierte Analysestrategie in der Forschung im Weiterbildungskontext zu aufschlussreichen Ergebnissen führen kann. Beispielhaft sei hier das von der Deutschen Forschungsgemeinschaft finanzierte Projekt „Förderung von Lehrexpertise: Das Verhältnis von individuellen Lernvoraussetzungen und Instruktionsprozessen in computerunterstützten fallbasierten Lernumgebungen“ genannt. In diesem Projekt lernten die Teilnehmenden in unterschiedlichen Varianten einer fallbasierten Lernumgebung. Sie sollten darin unterstützt werden, unterschiedliche Perspektiven zu übernehmen und didaktische Ansätze bei der Analyse von Videofällen zu verwenden. Neben Plenarphasen gab es in den mehrtägigen Kursen auch individuelle und Kleingruppenlernphasen (vgl. Goeze u.a. 2010, S. 1100–1102). Es wurde analysiert, ob durch die spezifische Unterstützung verschiedene Perspektiven tatsächlich besser eingenommen werden konnten und ob es den Lernenden mit der Unterstützung besser gelang, die didaktischen Ansätze bei der Analyse der Fälle anzuwenden.

Zu den interessantesten methodischen Herausforderungen für die Forschung zu Lern- und Kooperationsprozessen zählt derzeit die Integration unterschiedlicher Ebenen bei der Analyse (z.B. Individuum, Kleingruppe, Plenum), um die komplexen Wirkungen von instruktionalen Interventionen auf das Handeln von Lehrenden und Lernenden besser zu verstehen.

Literatur

- Baron, R.M./Kenny, D.A. (1986): The Moderator-Mediator Variable Distinction in Social Psychology Research: conceptual, strategic, and statistical considerations. In: *Journal of Personality and Social Psychology*, Bd. 51, S. 1173–1182
- De Wever, B./Schellens, T./Valcke, M./Van Keer, H. (2006): Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: a review. In: *Computers & Education*, Bd. 46, S. 6–28
- Ericsson, K.A./Simon, H.A. (1984): *Protocol Analysis: Verbal Reports as Data*. Cambridge, Massachusetts
- Fischer, F./Stegmann, K./Wecker, C./Kollar, I. (2011): Online-Diskussionen in der Hochschullehre: Kooperations-skripts können das fachliche Argumentieren verbessern. In: *Zeitschrift für Pädagogik*, Bd. 57, S. 326–337
- Goeze, A./Zottmann, J./Schradier, J./Fischer, F. (2010): Instructional Support for Case-Based Learning with Digital Videos: Fostering Pre-Service Teachers' Acquisition of the Competency to Diagnose Pedagogical Situations. In: Gibson, D./Dodge, B. (Hg.): *Proceedings of the Society for Information Technology & Teacher Education (SITE) International Conference 2010*. Chesapeake, Virginia, S. 1098–1104
- King, A. (2007): Scripting Collaborative Learning Processes: A cognitive perspective. In: Fischer, F./Kollar, I./Mandl, H./Haake, J. (Hg.): *Scripting Computer-Supported Collaborative Learning: Cognitive, computational, and educational perspectives*. New York, S. 13–37
- Mu, J./Stegmann, K./Mayfield, E./Rosé, C./Fischer, F. (2012). The ACODEA Framework: developing classification schemes for fully automatic classification of online discussions. *International Journal of Computer-Supported Collaborative Learning*, 7(2), 285–304
- Nachtigall, C./Kroehne, U./Funke, F./Steyer, R. (2003): (Why) Should We Use SEM? Pros and Cons of Structural Equation Modeling. *Methods of Psychological Research Online*, Bd. 8, S. 1–22
- Raudenbush, S.W./Bryk, A.S. (2002): *Hierarchical Linear Models: Applications and data analysis methods*. 2. Aufl. Thousand Oaks
- Rosé, C.P./Wang, Y.C./Arguello, J./Cui, Y./Stegmann, K./Weinberger, A./Fischer, F. (2008): Analyzing Collaborative Learning Processes Automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, Bd. 3, S. 237–271
- Seidel, T./Prenzel, M. (2010): Beobachtungsverfahren: Vom Datenmaterial zur Datenanalyse. In: Holling, H./Schmitz, B. (Hg.): *Handbuch Statistik, Methoden und Evaluation*. Göttingen, S. 139–152
- Stegmann, K./Wecker, C./Weinberger, A./Fischer, F. (2012): Collaborative Argumentation and Cognitive Processing – An empirical study in a computer-supported collaborative learning environment. *Instructional Science*, Bd. 40, S. 297–323
- Stegmann, K./Weinberger, A./Fischer, F. (2011): Aktives Lernen durch Argumentieren: Evidenz für das Modell der argumentativen Wissenskonstruktion in Online-Diskussionen. *Unterrichtswissenschaft*, Bd. 39, S. 231–244
- Wecker, C./Fischer, F. (2011): From Guided to Self-Regulated Performance of Domain-General Skills: The role of peer monitoring during the fading of instructional scripts. *Learning and Instruction*, Bd. 21, S. 746–756
- Wirtz, M./Caspar, F. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen