

TALIS-Videostudie Deutschland - Erläuterung der nationalen Datensätze

Stand

September 2021

Autor*innen

Benjamin Herbert¹

Julia Käfer¹

Patrick Schreyer¹

Jessica Fischer²

Juliane Grünkorn¹

Eckhard Klieme¹

Anna Praetorius³

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

²Deutsches Institut für Erwachsenenbildung - Leibniz-Zentrum für Lebenslanges Lernen

³Universität Zürich

Kontakt

videostudie@dipf.de

Inhaltsverzeichnis

1	Besonderheiten der nationalen Datensätze	1
2	Matching-Variablen in nationalen und internationalen Datensätzen	3
3	Übersicht nationaler Datensätze und Dokumente.....	4
4	Felddaten.....	5
5	Schüler*innendaten.....	5
6	Lehrer*innendaten.....	6
6.1	Befragungsdaten	6
6.2	Teacher Log Daten.....	6
7	Rating Daten	7
7.1	Artefactratings.....	7
7.2	Videoratings	8
7.2.1	Benennung der Variablen.....	9
7.2.2	Vorgehen beim Aggregieren	9
7.2.3	Sortierung der Datensätze	11
7.2.4	Begleitdokumente	12

1 Besonderheiten der nationalen Datensätze

Die international durchgeführten Erhebungen der *TALIS-Video Study* bzw. *Global Teaching InSights* wurden in Deutschland erweitert. Dadurch entstanden zusätzliche nationale Datensätze. Dieses Dokument enthält wichtige Hinweise zum Aufbau, Inhalt und dem korrekten Umgang mit den verschiedenen nationalen Datensätzen der *TALIS-Videostudie Deutschland*.

Internationale Datensätze

Die internationalen Datensätze der *TALIS-Video Study* werden auf der Seite der OECD zum Download bereitgestellt (<http://www.oecd.org/education/school/global-teaching-insights-technical-documents.htm>)¹. Da die Studie im Zuge der Berichtlegung in *Global Teaching InSights* umbenannt wurde, sind alle Informationen und Veröffentlichungen unter diesem neuen Namen gelistet. Die dort verfügbaren Datensätze enthalten die Daten aller acht Länder, die an der Studie teilgenommen haben, und umfassen keine nationalen Ergänzungen. Neben dem *Technical Report* finden sich auf der Seite der OECD auch Informationen zu den eingesetzten Messinstrumenten sowie das Dokument *User Guide and Codebook*, in dem umfangreiche Informationen zu den einzelnen Datensätzen enthalten sind. Es wird unbedingt empfohlen dieses Dokument vor der Nutzung nationaler oder internationaler Datensätze zu sichten.

Unterschiede zwischen nationalen und internationalen Datensätzen

Obgleich die nationalen Datensätze zu Teilen aus den internationalen reproduziert werden könnten, bestehen einige weitreichende Unterschiede. Zunächst enthalten die internationalen Datensätze Skalenwerte und abgeleitete Variablen, die im Zuge der internationalen Berichtlegung gebildet wurden. Die nationalen Datensätze bestehen weitestgehend aus Rohdaten – Ausnahmen bilden die Ratingdaten der Unterrichtsstunden zum Thema *Quadratische Gleichungen*, die in den nachfolgenden Kapiteln erläutert werden. Einige nationale Datensätze wurden um nationale Ergänzungen erweitert. Ergänzungen fanden dabei auf zwei Arten statt: Zum einen wurden die international einheitlichen Befragungen, die vor und unmittelbar nach der Unterrichtseinheit zum Thema Quadratische Gleichungen stattfanden (Messzeitpunkte A und B bzw. Pre und Post), um zusätzliche Items erweitert. Diese wurden an das Ende der jeweiligen Fragebögen eingefügt und von den Schüler*innen und Lehrpersonen unmittelbar im Anschluss an die anderen Items bearbeitet. Während der Unterrichtseinheit wurden, wie im internationalen Design der Studie vorgesehen, zwei von der Lehrperson ausgewählte Unterrichtsstunden videografiert, jeweils eine Stunde aus der ersten (TVA) und der zweiten (TVB) Hälfte der Unterrichtseinheit. Dieses Design wurde in Deutschland ergänzt. Etwa sechs Wochen nach der untersuchten Unterrichtseinheit wurde eine weitere Befragung der

¹ Stand: 15.09.2021

Schüler*innen und Lehrpersonen durchgeführt (Messzeitpunkt C bzw. Follow-Up). Zudem fand eine zusätzliche Videografie zu einem beliebigen Thema statt (TVC), die zeitlich vor oder nach der Unterrichtseinheit zu Quadratischen Gleichungen liegen konnte. Diese Ergänzungen spiegeln sich in den nationalen Datensätzen wider.

Besonderheiten der deutschen Befragungsdaten: Eine ausführliche Erläuterung der nationalen Ergänzungen der Befragungsinstrumente findet sich im Dokument *Skalenhandbuch zu nationalen Erweiterungen der Fragebogenerhebungen*. Dieses enthält unter anderem eine Übersicht der erfassten Konstrukte, welche die nationalen Ergänzungen einschließt. Ein weiterer Unterschied betrifft den Datensatz zu den TeacherLogs. In diesem sind auf nationaler Ebene alle Lehrpersonen als Fälle enthalten, die zugesagt haben, an der Studie teilzunehmen, unabhängig davon, ob die Teilnahme stattfand. Auf internationaler Ebene wurden entsprechende Fälle aus dem Datensatz entfernt. Darüber hinaus enthält der nationale Datensatz Hilfsvariablen, um die einzelnen Unterrichtsstunden einer Lehrperson in den Verlauf der Unterrichtseinheit einzuordnen.

Besonderheiten der deutschen Ratingdaten: Eine weitere Besonderheit der nationalen Datensätze besteht darin, dass die Ratingdaten der Unterrichtsvideos (TVA und TVB) auf verschiedenen Aggregationsebenen vorliegen. Neben den Rohdaten gibt es einen zusätzlich Datensatz, in dem die verschiedenen Ratings einer Unterrichtsstunde bereits auf die Ebene von Unterrichtsstunden gemittelt wurden und einen weiteren Datensatz, in dem die Daten bereits auf die Ebene von Lehrpersonen gemittelt wurden. Zudem liegt ein separater Rohdatensatz mit den Ratings der zusätzlichen Videografie (TVC) und ausgewählter Videos der TIMSS-Videostudie 1995 vor und der Datensatz der Artefaktratings enthält national ergänzte Codes.

ID-Korrekturen in den nationalen Datensätzen: In den hier vorliegenden nationalen Datensätzen wurden Korrekturen einzelner IDs vorgenommen. ACHTUNG: Im internationalen Datensatz sind diese Korrekturen nicht durchgeführt worden, sodass dort noch die ursprünglichen IDs enthalten sind. Im Ordner *Syntax_ID_Korrektur* finden sich die zur Korrektur angewandten Syntaxen.

Konkret handelt es sich in einem Fall um eine Vertauschung zweier Lehrkräfte (00276-0021-0001 und 00276-0021-0003) bei den papierbasierten Erhebungen (Fragebogen, Tests, Teacher-Log). Die IDs wurden in den entsprechenden Datensätzen (*TVS_Teacher_Data.sav*, *TVS_Teacher_Log.sav*; *TVS_Student_Data.sav*) getauscht. Die Ratingdatensätze waren von dieser Vertauschung nicht betroffen. Hier wurden die IDs korrekt erhoben und sind auch im internationalen Datensatz korrekt zugeordnet.

Des Weiteren wurde die ID der Lehrkraft 00276-0059-0001 neu zugeordnet, da die Lehrkraft der Schule 0006 zugehörig war. Die neue ID lautet demzufolge 00276-0006-0005.

Alte IDs (internationaler Datensatz)	Neue IDs (nationale Datensätze)
00276-0021-0001	00276-0021-0003
00276-0021-0003	00276-0021-0001
00276-0059-0001	00276-0006-0005

Historischer Vergleich der Videodaten

Der historische Vergleich der Unterrichtsvideografien aus der TALIS-Videostudie Deutschland mit denen der TIMSS-Videostudie aus dem Jahr 1995 war ein übergeordnetes Ziel der TALIS-Videostudie Deutschland. Die Videodaten der dritten Videografie zu einem beliebigen Thema (TVC) wurden im Wesentlichen für diesen historischen Vergleich angefertigt. Die beliebige Themenauswahl bei diesen Videografien entspricht dem Vorgehen der TIMSS-Videostudie, bei der damals ebenfalls beliebige Themen im Fach Mathematik aufgezeichnet wurden. Für den Vergleich wurden ausgewählte Videos der TIMSS-Videostudie hinzugezogen und ebenfalls mit dem Kodiersystem der TALIS-Video Study ausgewertet. Für beide Videodatensätze (TVC und TIMSS) liegen entsprechend zwei separate Datensätze vor (*TVS_Video_Data_Third_Vid.sav* und *TIMSS_Video_Data.sav*).

In der TALIS-Videostudie Deutschland konnten bei 44 der 50 teilnehmenden Lehrkräfte eine dritte Videografie zu einem beliebigen Thema (TVC) durchgeführt werden. Für den historischen Vergleich wurde dementsprechend eine gleich große Substichprobe aus den Videos der TIMSS-Videostudie gebildet. Die Auswahl der Videos aus der TIMSS-Videostudie unterlag verschiedenen Kriterien. Zum einen wurden die soziodemographischen Daten der TALIS-Videostudie herangezogen. Der Überhang an Gymnasien und das Fehlen von Hauptschulen in der Stichprobe wurde auf die Substichprobe der TIMSS-Videos übertragen. Ebenfalls wurde die Verteilung in den einzelnen Bundesländern (fast ausschließlich alte Bundesländer) berücksichtigt. In einem letzten Schritt wurden die in den Unterrichtsstunden behandelten Themen verglichen, sodass in beiden Stichproben möglichst ähnliche Themen vertreten waren. Dies führte zu einer geeigneten Substichprobe von ebenfalls 44 Unterrichtsvideografien aus der TIMSS-Videostudie. Eine Übersicht der Themen der ausgewählten Unterrichtsstunden findet sich in den Begleitmaterialien der Ratingdaten.

2 Matching-Variablen in nationalen und internationalen Datensätzen

Alle nationalen und internationalen Datensätze enthalten die Variable *T_ID* bzw. *S_ID*, die als Matching-Variable verwendet werden kann. Die erste der beiden genannten Variablen stellt eine individuelle Kennung für jede Lehrperson dar und besteht aus drei Zahlenblöcken, die durch Bindestriche getrennt sind (z.B. 00276-0001-0001). Der erste Block, bestehend aus fünf Zeichen, gibt den Ländercode an, der zweite vierstellige Block den Schulcode und der dritte ebenfalls vierstellige

Block den Code der Lehrperson an der jeweiligen Schule. Die zweite Variable (*S_ID*) stellt die individuelle Kennung für jede*n Schüler*in dar und ergänzt die Lehrpersonen-Kennung und einen weiteren vierstelligen Block, der für die Schülerin oder den Schüler in der Klasse steht (z.B. 00276-0001-0001-0001).

3 Übersicht nationaler Datensätze und Dokumente

Die Daten der TALIS-Videostudie Deutschland, sind in vier Kategorien/Ordner gliedert:

- Fielding Data
- Rating Data
- Student Data
- Teacher Data

In jeder Kategorie befinden sich die jeweiligen Datensätze sowie Dokumente mit Informationen zu den Instrumenten, die für die Interpretation der Daten relevant sind, z.B. Erhebungsinstrumente oder Kodier-Manuale:

Fielding Data	Datenblatt mit Informationen zum Erhebungsprozess
Rating Data	Datensätze zu den Video-Ratings, Informationen zum Artefakt-Kodiersystem Datensätze zu den Artefact-Ratings, Informationen zum Video-Kodiersystem
Student Data	Schüler*innendatensatz, Schüler*innenfragebögen, Tests, Lösungsschlüssel
Teacher Data	Lehrer*innendatensatz, Lehrer*innenfragebögen, TeacherLog-Datensatz, TeacherLog (Erhebungsbogen)

Neben den Dokumenten innerhalb der vier Kategorien gibt es drei allgemeine Übersichtsdokumente:

(1) Dieses Dokument liefert eine Erläuterung der nationalen Datensätze (*TVS_Erläuterung_nationale_Datensätze.pdf*).

(2) Im Skalenhandbuch zu nationalen Erweiterungen der Fragebogenerhebungen (*TVS_Skalenhandbuch_Ergänzungen_Befragungsdaten.pdf*) sind alle nationalen Ergänzungen der Befragungsinstrumente ausführlich dokumentiert. Es enthält als Anhang eine Konstruktübersicht, in der alle Items der Schüler*innen- und Lehrer*innenfragebögen zu allen Erhebungen (Pre, Post, Follow-up) den intendierten Konstrukten zugeordnet werden.

(3) In einem dritten Dokument (*TVS_VideoID_Recodings_FDZ.pdf*) ist aufgelistet, wie Videos und Transkripte benannt wurden, bevor sie vom Forschungsdatenzentrum zur Verfügung gestellt wurden

(<http://dx.doi.org/10.7477/352:1:0>)². In der Auflistung ist der Originalname der Videos (z.B. 276-0001-01-TVA-16042018) angegeben - dieser setzt sich zusammen aus der ID der Unterrichtsstunde (*IDLESSON*, zu finden in einigen nationalen Datensätzen) und dem Datum der Videografie. Die ersten 15 Zeichen, die der Variablen *IDLESSON* entsprechen, setzen sich ähnlich zusammen wie die im vorherigen Abschnitt beschriebene Matching Variable *T_ID*: Ländercode, Schulcode und Lehrpersonencode. Anschließend findet sich die Information, um welche Videografie der Lehrperson es sich handelt (TVA, TVB oder TVC). Die Auflistung enthält darüber hinaus die neuen Namen der Videos, unter denen sie beim Forschungsdatenzentrum hinterlegt sind sowie den Namen des Moduls, dem die Videos zugeordnet wurden. Jedes Modul beinhaltet alle beim Forschungsdatenzentrum hinterlegten Daten zu einer Unterrichtsstunde. In der Regel handelt es sich dabei um ein Video aus Schüler*innenperspektive, ein Video aus Lehrpersonenperspektive, ein Transkript und die eingesetzten Unterrichtsmaterialien.

4 Felddaten

Im Ordner *Fielding Data* befindet sich die Excel-Datei *TVS_Fielding_Data.xlsx*, die nähere Information zum Erhebungsprozess der TALIS-Videostudie liefert. In dieser ist dargestellt, an welchen Tagen die verschiedenen Erhebungen, d.h. Vorbefragung/Vortest, erste Videografie (Video A), zweite Videografie (Video B), Nachbefragung/Nachttest und Follow-up-Befragung/Follow-up-Test stattfanden. Zudem ist vermerkt, ob Schüler*innen an den Erhebungen teilnahmen, und wenn nicht, was der Grund für den Ausfall war. Das Datum der dritten Videografie ist nicht in dem Dokument enthalten; dieses lässt sich aus der Variable *DATE_VIDEO* im Datensatz *TVS_Video_Data_Third_Vid.sav* auslesen.

5 Schüler*innendaten

Im Ordner *Student Data* befindet sich der Schüler*innendatensatz *TVS_Student_Data.sav*. Dieser enthält die Angaben der Schüler*innen aus den Fragebögen und Tests zu allen Erhebungszeitpunkten. Dabei ist der Datensatz so aufgebaut, dass vorne die Fragebogendaten und hinten die Testdaten stehen. Um die Items und Skalen den einzelnen Erhebungen zuordnen zu können, gilt folgendes System:

Vortest	STA	Nachttest	STB	Follow-up-Test	STC
Vorbefragung	SQA	Nachbefragung	SQB	Follow-up-Befragung	SQC

Entsprechend ist SQA01 das erste Item im Schüler*innenfragebogen der Vorerhebung und STA01 das erste Item im Schüler*innentest der Vorerhebung. Die intendierten Zusammensetzung der Items zu

² Stand: 15.09.2021

Skalen ist in der Konstruktübersicht in Anhang A des Skalenhandbuchs zu den nationalen Ergänzungen der Befragungsinstrumente dokumentiert (*TVS_Skalenhandbuch_Ergänzungen_Befragungsdaten.pdf*).

Innerhalb der sich anschließenden Testdaten befinden sich die mit 0 (=falsch) und 1 (=richtig) codierten Angaben zu den einzelnen Testitems vorne, die skalierten Testwerte (WLE aus nationalen Rasch-Skalierungen, getrennt für STA, STB und STB) hinten. Die WLE-Reliabilitäten sind: Pretest: .76, Posttest: .66, Follow-Up: .74. Der Itemfit der Rasch-Modelle ist zu allen drei Messzeitpunkten zufriedenstellend mit Infit-Werten zwischen 0.8 und 1.2.

Als Begleitdokumente wurden die eingesetzten Schüler*innenfragebögen und -tests mit Lösungsschlüsseln für alle drei Erhebungszeitpunkte bereitgestellt.

6 Lehrer*innendaten

6.1 Befragungsdaten

Im Ordner *Teacher Data* befindet sich der Lehrer*innendatensatz *TVS_Teacher_Data.sav*. Dieser enthält die Angaben der Lehrpersonen aus den Fragebögen zu allen Erhebungszeitpunkten. Um die Items den einzelnen Erhebungen zuordnen zu können, gilt folgendes System:

Vorbefragung	TQA	Nachbefragung	TQB	Follow-up-Befragung	TQC
--------------	-----	---------------	-----	---------------------	-----

Entsprechend ist TQA01 das erste Item im Lehrer*innenfragebogen der Vorerhebung. Die intendierte Zusammensetzung der Items zu Skalen ist in der Konstruktübersicht in Anhang A des Skalenhandbuchs zu den nationalen Ergänzungen der Befragungsinstrumente dokumentiert (*TVS_Skalenhandbuch_Ergänzungen_Befragungsdaten.pdf*).

Begleitdokumente sind die eingesetzten Lehrer*innenfragebögen der drei Erhebungszeitpunkte.

6.2 Teacher Log Daten

Zusätzlich zu den Befragungen wurden Lehrpersonen gebeten, die inhaltlichen Schwerpunkte aller Stunden aus der Unterrichtseinheit zu Quadratischen Gleichungen in einem TeacherLog zu dokumentieren. Der Datensatz *TVS_Teacher_Log.sav* enthält für jede einzelne Stunde eine Einschätzung der Lehrperson zur Behandlung von neun Inhalten aus dem Themengebiet (siehe *TVS_Teacher_Log.pdf*). Jeder Fall stellt die Angaben zu einer Stunde dar. Das Behandeln der verschiedenen Inhalte wurde von den Lehrpersonen wie folgend eingeschätzt: „0 = gar nicht behandelt“, „1 = mit behandelt“ und „2 = als Schwerpunkt der Stunde behandelt“.

Der Datensatz enthält zwei Hilfsvariablen: Die Variable *lessonnumber* gibt für jeden Fall an, um die wievielte Unterrichtsstunde es sich im Verlauf der Unterrichtseinheit handelt. Die relative Position einer

Stunde kann zudem über die Variable *relative_lessonnumber* ausgelesen werden, sie nimmt den Wert 0 für die erste Stunde der Unterrichtseinheit und 1 für die letzte Stunde an.

Als Begleitdokument ist der Erhebungsbogen (*TVS_Teacher_Log.pdf*) bereitgestellt, der den Lehrpersonen zu Beginn der Unterrichtseinheit zu Quadratischen Gleichungen ausgeteilt, und am Ende der Einheit eingesammelt wurde.

7 Rating Daten

Im Ordner *Rating Data* befinden sich ein Datensatz mit den Artefaktratings (*TVS_Artefact_Data.sav*) sowie verschiedene Datensätze zu den Videoratings. Für die Videoratings zur Unterrichtseinheit Quadratische Gleichungen, die auch im internationalen Datensatz zu finden sind, wurden zusätzlich zu dem Rohdatensatz (*TVS_Video_Data.sav*) aggregierte Datensätze auf Stunden- und Lehrpersonenebene bereitgestellt (*TVS_Video_Data_Aggregated_Lesson.sav* und *TVS_Video_Data_Aggregated_Teacher.sav*). Die Ratings der zusätzlichen Videografie (TVC) und der TIMSS-Videos liegen als Rohdatensätze vor (*TVS_Video_Data_Third_Vid.sav*; *TIMSS_Video_Data.sav*).

7.1 Artefactratings

Im Ordner *Rating Data* befindet sich ein Datensatz mit den Artefaktratings: *TVS_Artefact_Data.sav*. Dieser beinhaltet die Ratings der Unterrichtsmaterialien der beiden videografierten Stunden zum Thema Quadratische Gleichungen (TAA und TAC) sowie deren jeweiliger Folgestunde (TAB und TAD). Darüber hinaus sind Ratings für den Leistungstest enthalten, der von der jeweiligen Lehrperson am Ende der Unterrichtseinheit zu Quadratischen Gleichungen durchgeführt wurde (TAT). Im Datensatz ist jedes erhobene Artefakt-Set mindestens zweimal vertreten, weil mindestens zwei (bei den sogenannten Validierungs-Sets auch mehr) Rater*innen unabhängig voneinander die Materialien eingeschätzt haben. Jeder Fall im Datensatz entspricht einem Rating.

Die Artefaktratings wurden um die national entwickelten Codes *Sprachlogische Komplexität* und *Aufgabenklassen - Typen mathematischen Arbeitens* ergänzt, die gemeinsam mit den anderen Codes ausgewertet wurden. Beide Codes wurden in Anlehnung an Codes der COACTIV Studie entwickelt.

Die Variablen sind im Datensatz wie folgend sortiert:

1. Identifikationsvariablen (IDLESSON, IDTEACHER, IDCNTRY, ...)
2. Scorability Variablen (SCO_01, SCO_02, SCO_03)
3. Component Variablen (COM_01, COM_02, ...)
4. Component Variablen – nationale Ergänzungen (COM_12 bis COM_15)
5. Subtopic Variablen (SUB01, SUB02, ...)

Als Begleitdokumente sind das Codebook und das Manual hinterlegt, mit denen die Artefakt-Sets eingeschätzt wurden. Das Codebook beinhaltet alle Codes, das Manual enthält zusätzliche Informationen zu den Codes und dazu, wie sie anzuwenden sind. Beide Dokumente umfassen die nationalen Ergänzungen. Zur Auswertung der Artefakt-Sets lagen den Rater*innen beide Dokumente vor.

7.2 Videoratings

Die Datensätze der Videoratings enthalten jeweils alle Variablen des Component- und des Indicator-Kodiersystems. Der Dokumentenname gibt dabei die jeweilige Aggregationsebene der Daten an:

TVS_Video_Data.sav;

TVS_Video_Data_Third_Vid.sav;

TIMSS_Video_Data.sav

Jeder Fall entspricht den Ratings einer Raterin oder eines Raters zu einer Unterrichtsstunde. Die Datensätze enthalten alle Rohdaten (also Ratings für einzelne Segmente). Der Datensatz der Hauptstudienvideos (*TVS_Video_Data.sav*) umfasst zusätzlich Variablen auf Rating-Ebene, bei denen die Ausprägungen einer Variablen über die einzelnen Segmente hinweg gemittelt wurden. In diesen Datensätzen ist jede gefilmte Stunde mindestens zweimal vertreten, weil sowohl zu den Components als auch zu den Indicators mindestens zwei (bei den sogenannten Validierungs-Videos auch mehr) Rater*innen unabhängig voneinander die Videos eingeschätzt haben.

Wichtig: Das erste Components-Rating und das erste Indicators-Rating bilden zusammen einen Fall der Datei, obwohl beide von unterschiedlichen Rater*innen eingeschätzt wurden. Die Rater*innen lassen sich anhand der Rater-IDs identifizieren (Variablen *IDRATER_Components* und *IDRATER_Indicators*). Die Zusammensetzung der Fälle aus unabhängigen Indicator- und Component-Ratings zieht sich durch den gesamten Datensatz.

TVS_Video_Data_

Aggregated_Lesson.sav

Jeder Fall entspricht den Ratings für eine Unterrichtsstunde. Der Datensatz enthält alle Variablen über Ratings hinweg aggregiert auf Stunden-Ebene. Hierzu wurden die (mindestens) zwei Ratings pro Unterrichtsstunde gemittelt.

TVS_Video_Data_

Aggregated_Teacher.sav

Jeder Fall entspricht den Ratings für eine Lehrperson. Der Datensatz enthält alle Variablen über Stunden hinweg aggregiert auf Lehrpersonen-Ebene. Hierzu wurden die Ratings der beiden

Videos aus der Unterrichtseinheit zu Quadratischen Gleichungen gemittelt.

7.2.1 Benennung der Variablen

ICE4CT_S1

Kodiersystem: Indicators

Konstrukt: Student Cognitive Engagement

Variable innerhalb des Kodiersystems und Konstrukts: 4

Name des Items: Classroom Technology

Nummer des Segments: Segment 1

Auf Aggr.-Ebene der Ratings fällt der Zusatz zum Segment weg (ICE4CT).

Auf Aggr.-Ebene der Stunde ist der Zusatz „_mean“ angehängt (ICE4CT_mean).

Auf Aggr.-Ebene der Lehrperson ist der Zusatz „_mean_mean“ angehängt (ICE4CT_mean_mean).

7.2.2 Vorgehen beim Aggregieren

Aggregation von metrischen Ratingdaten

Für die Aggregation auf Rating-Ebene (*TVS_Video_Data.sav*) wurden die Ratings der Segmente (separat für alle Rater*innen) gemittelt. Da es keine Ausfälle bei den Ratings gibt, mussten keine Regelungen zur Mindestanzahl gültiger Ratings getroffen werden. Die Variablen *Segment_lesson_Components* und *Segment_lesson_Components* geben Auskunft darüber, aus wie vielen Segmenten eine Unterrichtsstunde besteht und folglich aus wie vielen Einzelwerten sich das Rating-Mittel zusammensetzt.

Für die Aggregation auf Stunden-Ebene (*TVS_Video_Data_Aggregated_Lesson.sav*) wurden die im vorherigen Absatz beschriebenen aggregierten Ratings gemittelt. Jedes Rating einer Stunde erhält damit das gleiche Gewicht. Alle Stunden-Werte setzen sich aus mindestens zwei Ratings zusammen (Validierungsvideos wurden von bis zu 5 Rater*innen ausgewertet). Die Variablen *Ratings_lesson_Components* und *Ratings_lesson_Indicators* geben Auskunft darüber, wie viele Rater*innen eine Stunde ausgewertet haben und folglich aus wie vielen Einzelwerten sich das Stundenmittel zusammensetzt.

Für die Aggregation auf Lehrpersonen-Ebene (*TVS_Video_Data_Aggregated_Teacher.sav*) wurden die Stundenmittelwerte gemittelt. Dadurch erhält jede Stunde das gleiche Gewicht, unabhängig davon, welchen zeitlichen Umfang sie hat.

Aggregation von nicht-metrischen Ratingdaten

Für nicht-metrische Indicator-Variablen wurden bei der Aggregation auf Rating-Ebene separate Variablen für die einzelnen Ausprägungen erstellt. Die Werte der Variablen entsprechen relativen

Häufigkeiten für das Auftreten der jeweiligen Ausprägung über die Segmente hinweg. In den weiteren Aggregationsschritten wurden nur noch die neu erstellten Variablen verwendet, die die einzelnen Ausprägungen abbilden, und entsprechend der Erläuterungen im vorherigen Abschnitt weiter aggregiert.

Beispielsweise wurde für die Variable *Activity structure & frequency* in der Ausprägung *Whole group* (ICM2AW) für jede der vier Antwortkategorien „not used“, „briefly used“, „more than briefly used“ und „used the entire segment“ eine eigene Variable erstellt. Die Namen der neu entstanden Variablen setzen sich zusammen aus dem ursprünglichen Variablennamen ohne Segmentbezeichnung plus einer Zahl, die die Antwortkategorie angibt (z.B. ICM2AW_01 für „not used“). Alle relevanten Informationen zur Zuordnung sind im Datensatz auch im Variablenlabel enthalten. Das Vorgehen wurde für die folgenden Variablen angewandt:

- ICM2AW_01 = Activity structure & frequency: Whole group 01 Not used
- ICM2AW_02 = Activity structure & frequency: Whole group 02 Briefly used
- ICM2AW_03 = Activity structure & frequency: Whole group 03 more than briefly used
- ICM2AW_04 = Activity structure & frequency: Whole group 04 used the entire segment
- ICM3AS_01 = Activity structure & frequency: Small group 01 Not used
- ICM3AS_02 = Activity structure & frequency: Small group 02 Briefly used
- ICM3AS_03 = Activity structure & frequency: Small group 03 more than briefly used
- ICM3AS_04 = Activity structure & frequency: Small group 04 used the entire segment
- ICM4AP_01 = Activity structure & frequency: Pairs 01 Not used
- ICM4AP_02 = Activity structure & frequency: Pairs 02 Briefly used
- ICM4AP_03 = Activity structure & frequency: Pairs 03 more than briefly used
- ICM4AP_04 = Activity structure & frequency: Pairs 04 used the entire segment
- ICM5AI_01 = Activity structure & frequency: Individual 01 Not used
- ICM5AI_02 = Activity structure & frequency: Individual 02 Briefly used
- ICM5AI_03 = Activity structure & frequency: Individual 03 more than briefly used
- ICM5AI_04 = Activity structure & frequency: Individual 04 used the entire segment

Für dichotome Variablen wurde nur eine Variable erstellt, die das Vorhanden-Sein des Merkmals angibt. Das Vorgehen dabei war identisch mit der Beschreibung im vorherigen Abschnitt und wurde für folgende Variablen angewandt:

- IDC1DO_02 = Discussion opportunities 02 Present
- IQS6RG_02 = Types of representations: Graphs 02 present
- IQS7RT_02 = Types of representations: Tables 02 present
- IQS8RD_02 = Types of representations: Drawings or Diagrams 02 present
- IQS9RE_02 = Types of representations: Equations and Expressions 02 present

- IQS10RO_02 = Types of representations: Objects 02 present
- ICE10SU_02 = Software use for learning 02 present

Eine Ausnahme bilden die Konstrukte *Classroom Technology* (ICE4CT1, ICE5CT2, ICE6CT3) und *Student Technology* (ICE7ST1, ICE8ST2, ICE9ST3), die jeweils über drei Variablen erfasst wurden. Für den Fall, dass mehrere Technologien in einem Segment verwendet wurden, standen drei Variablen zur Verfügung, um diese einzutragen. Rater*innen waren dazu angehalten, je eine verwendete Technologie in eine beliebige der drei Variablen einzutragen. Bei der Aggregation auf Rating-Ebene wurde für jede Technologie eine eigene Variable erstellt:

- ICE4CT_01 = Classroom technology 01 Overhead projector or visualiser_document camera.
- ICE4CT_02 = Classroom technology 02 Smartboard or projector.
- ICE4CT_03 = Classroom technology 03 Graphing calculator.
- ICE4CT_04 = Classroom technology 04 Non-graphing calculator.
- ICE4CT_05 = Classroom technology 05 Computer/Laptop.
- ICE4CT_06 = Classroom technology 06 Television.
- ICE4CT_07 = Classroom technology 07 Tablet.
- ICE4CT_08 = Classroom technology 08 Cell phone.
- ICE4CT_09 = Classroom technology 09 No technology that requires electricity is used.
- ICE7ST_03 = Student technology 03 Graphing calculator.
- ICE7ST_04 = Student technology 04 Non-graphing calculator.
- ICE7ST_05 = Student technology 05 Computer/Laptop.
- ICE7ST_07 = Student technology 07 Tablet.
- ICE7ST_08 = Student technology 08 Cell phone.
- ICE7ST_09 = Student technology 09 No technology that requires electricity is used.

7.2.3 *Sortierung der Datensätze*

Sortierung der Variablen im Datensatz „TVS_Video_Data.sav“.

1. Identifikationsvariablen Teil 1 (IDLESSON, IDRATER_Components, ...)
2. Component Variablen für alle Segmente (CCM1RT_S1, CCM2MN_S1, ... CCM1RT_S2, ...)
3. Component Variablen Mittelwerte auf Rating-Ebene (CCM1RT, CCM2MN, ...)
4. Identifikationsvariablen Teil 2 (IDRATER_Indicators, ...)
5. Indicator Variablen für alle Segmente (ICM1TT_S1, ICM2AW_S1, ... ICM1TT_S2, ...)
6. Indicator Variablen Mittelwerte auf Rating-Ebene (ICM1TT, ...) bzw. separate Variablen zu relativen Häufigkeit von Ausprägungen von Indicators (ICM2AW_01, ...)

Die Sortierung der Variablen in den Datensätzen „*TVS_Video_Data_Third_Vid.sav*“ und „*TIMSS_Video_Data.sav*“ verhält sich identisch, mit der Ausnahme, dass keine Mittelwerte auf Ratingebene enthalten sind, wodurch die Punkte 3 und 6 wegfallen. Die Variablen „*T_ID*“ und „*DATE_VIDEO*“ sind im Datensatz „*TIMSS_Video_Data.sav*“ nicht enthalten. Die TIMSS-Videos lassen sich über die Variable „*IDLESSON*“ unterscheiden, die der damaligen Benennung der Videos entspricht. Das konkrete Erhebungsdatum der TIMSS-Videos lag nicht vor.

Sortierung der Variablen in den aggregierten Datensätzen „*TVS_Video_Data_Aggregated_Lesson.sav*“ und „*TVS_Video_Data_Aggregated_Teacher.sav*“:

1. Identifikationsvariablen Teil 1 (*IDLESSON* bzw. *IDTEACHER*, ...)
2. Component Variablen Mittelwerte auf Rating-/Lehrpersonen-Ebene (*CCM1RT_mean* bzw. *CCM1RT_mean_mean*, ...)
3. Identifikationsvariablen Teil 2 (*Segment_lesson_indicators*, ...)
4. Indicator Variablen Mittelwerte auf Rating-/Lehrpersonen-Ebene (*ICM1TT_mean* bzw. *ICM1TT_mean_mean*, ...)

7.2.4 Begleitdokumente

Unter den Begleitdokumenten („*TVS_Video_Documents.7z*“) befinden sich für das Components- und das Indicator-Kodiersystem jeweils das Codebook und das Manual. Das Codebook beinhaltet alle Codes, das Manual enthält zusätzliche Informationen zu den Codes und dazu, wie sie anzuwenden sind. Zur Auswertung der Videos lagen den Rater*innen beide Dokumente vor. Bei den Begleitdokumenten „*TVC_TIMSS_Documents.7z*“ finden sich Dokumente für den historischen Vergleich. Zum einen finden sich dort Auflistungen der Themen der videografierten Unterrichtsstunden (/TVC und TIMSS) und weiterhin ist dort eine Liste mit Startzeiten der TIMSS Videos zu finden. Die Startzeiten sind vom Forschungsteam festgelegte Zeiten, ab denen das Rating beginnen konnte, da die TIMSS-Videos oft eine längere Vorlaufzeit hatten (bspw. wurde der Klassenraum ohne Lehrperson gefilmt, anschließend ist ein harter Schnitt erkennbar).