

Thema

Evaluation der Qualität des Lehrens und Lernens im Fach Sport – Validierung des QUALLIS-Instruments

Christian Herrmann, Pädagogische Hochschule Zürich
Benjamin Niederkofler, Freie Universität Bozen-Bolzano
Harald Seelig, Universität Basel

*Ein wichtiges Teilgebiet der Bildungsforschung ist die Beschreibung von Unterrichtsqualität. Dabei stellen sich Fragen zu Qualitätsmerkmalen im Fach Sport und wie die Qualität des Lehrens und Lernens im Sportunterricht (QUALLIS) systematisch evaluiert werden kann. Der Beitrag arbeitet hierzu zwölf Merkmale guten Sportunterrichts heraus und operationalisiert die Merkmale mehrperspektivisch. Eine Validierungsstudie mit 40 Klassen der Sekundarstufe (634 Schüler*innen und deren Lehrer*innen) bestätigte über konfirmatorische Faktorenanalysen die erwartete zwölf-faktorielle Struktur und die Zusammenhänge mit der Schüler*innenmotivation. Das QUALLIS-Instrument ermöglicht einen empirischen Zugang zur Unterrichtsqualität im Fach Sport und bietet damit einen Mehrwert für die Lehrer*innenbildung.*

1. Problemstellung

Als qualitativ hochwertiger Fachunterricht („quality teaching“) wird ein lernwirksamer Fachunterricht („effective teaching“) verstanden, der gleichzeitig der spezifischen Fachlichkeit des Schulfaches gerecht wird („good teaching“; Berliner, 2005). Entsprechend zielt ein qualitativvoller Unterricht neben fachlichen Kompetenzen auch auf fachübergreifende Kompetenzen und erzieherische Wirkungen ab (E-EDK, 2017). Im Sinne des Angebots-Nutzungs-Modells (Helmke, 2017) wird Unterricht dabei als Angebot an die Lernenden begriffen, das sich einerseits durch die eingesetzten Unterrichtsmethoden und andererseits durch die Unterrichtsqualität beschreiben lässt. Die Unterrichtsmethoden gelten als Sichtstrukturen des Unterrichts und beinhalten Organisationsmerkmale, Methoden und Sozialformen. Unterrichtsqualität beschreibt die Tiefenstruktur des Unterrichts und ist geprägt von den Handlungen, der Wahrnehmung der handelnden Akteure (Lehrperson, der Schüler*innen, Beobachter*innen) sowie deren Interaktionen zueinander (Decristan et al., 2020).

Die Bedeutsamkeit von Unterrichtsqualität für Lernergebnisse wird in der Unterrichtsforschung explizit hervorgehoben (Kunter & Ewald, 2016). So zeigten empirische Befunde, dass weniger die Verwendung spezifischer Unterrichtsmethoden die Lernleistungen der Schüler*innen beeinflusst, als vielmehr die Qualität des Unterrichts. Unterrichtsqualität hat einen bedeutsamen Einfluss auf den Kompetenzerwerb (Hattie, 2013; Seiz et al., 2016). Um derartige Prozesse des Lehrens und Lernens „sichtbar“ (Hattie, 2013) zu machen, ergibt sich die Notwendigkeit, Unterricht systematisch zu evaluieren. Mit Blick auf die gesellschaftliche Legitimation des Faches Sport konstatieren Kurz und Gogoll (2010, S. 238), dass „sich eine mit öffentlichen Mitteln finanzierte Schule in keinem ihrer Fächer der Forderung entziehen kann, an der Qualität ihrer Massnahmen zu arbeiten und über ihre Ergebnisse Rechenschaft abzulegen“.

Im vorliegenden Beitrag werden zentrale Merkmale von Unterrichtsqualität für das Fach Sport benannt, welche für die Erreichung zentraler Erziehungs- und Lernziele des Sportunterrichts von Bedeutung sind. Dies erfolgt im engen Bezug zur Unterrichtsqualitätsforschung und wird durch sportwissenschaftliche Forschung fachspezifisch ausdifferenziert. Darauf aufbauend wird die Entwicklung und empirische Validierung des QUALLIS-Instruments zur Erfassung der Qualität des Lehrens und Lernens im Sport dargestellt.

2. Fachübergreifende und fachspezifische Aspekte der Unterrichtsqualität

Die Unterrichtsqualitätsforschung entwickelte verschiedene Konzeptualisierungen zur Beschreibung von Unterrichtsqualität, welche weitgehend stufen- und fachübergreifend begriffen wurden und werden (Begrich et al., 2023). Da sich die Lernziele, -gegenstände und -prozesse zwischen den Fächern teils stark unterscheiden, stellt

sich die Frage, ob diese fachübergreifenden Modelle für die verschiedenen Schulfächer und für das Fach Sport Gültigkeit besitzen. Welche Merkmale muss also ein guter Sportunterricht aufweisen, wenn dessen fachliche Lernziele und Lerngegenstände spezifisch adressiert werden sollen?

Aktuelle fachübergreifende (Praetorius et al., 2020) und sportspezifische Übersichtsarbeiten (Herrmann & Gerlach, 2020) machen deutlich, dass die verschiedenen Merkmale der Unterrichtsqualität insbesondere auf Ebene der beobachtbaren Indikatoren fachspezifisch ausdifferenziert werden müssen, um die Unterrichtsprozesse in den verschiedenen Fächern und im Fach Sport adäquat beschreiben und operationalisieren zu können

Ausgehend von der fachübergreifenden Unterrichtsqualitätsforschung (2.1) und dem aktuellen Diskurs zur Unterrichtsqualität im Fach Sport (2.2) wird versucht zentrale Merkmale qualitätsvollen Sportunterrichts (2.3) herauszuarbeiten.

2.1 Fachübergreifende Unterrichtsqualitätsforschung

In der Unterrichtsforschung wurden fachübergreifende Merkmalslisten guten Unterrichts entwickelt, die zentrale Aspekte von Unterrichtsqualität zusammenfassen. Im deutschsprachigen Raum waren die Merkmale guten Unterrichts nach Helmke (2017) und Meyer (2005) gebräuchlich, welche für die Unterrichtspraxis eine gute Orientierungshilfe bei der Planung und Reflexion von Unterricht bieten und daher in der Lehrpersonenbildung immer noch Verwendung finden. Inhaltlich existieren zwischen mehreren Merkmalen Überschneidungen, welche aus einer pädagogischen Perspektive durchaus intendiert sind (Meyer, 2005). Mittlerweile werden diese Merkmalslisten in der empirischen Unterrichtsforschung zunehmend kritisiert, da die theoretische Fundierung lückenhaft ist und unklar bleibt, inwieweit die genannten Merkmale sich überlagern. Entsprechend wird eine theoretische und empirische Fundierung der Unterrichtsmerkmale gefordert, in der übergeordnete Konstrukte identifiziert werden, die hinreichend unabhängige und sparsame Kernprinzipien der Unterrichtsqualität darstellen (Klieme & Rakoczy, 2008).

In verschiedenen Studien (u. a. TIMSS, Klieme & Baumert, 2001; COACTIV, Kunter et al., 2013) wurden drei Basisdimensionen von Unterrichtsqualität identifiziert und zu einem fundierten Modell verknüpft (im Überblick Praetorius et al., 2018). Demnach zeichnet sich guter Unterricht aus durch (a) eine strukturierte, klare und störungspräventive Unterrichts- und *Klassenführung*, (b) ein unterstützendes, schüler*innenorientiertes *Unterrichtsklima* und (c) das Potenzial zur kognitiven *Aktivierung*. Diese Basisdimensionen bilden eine empirisch verankerte Systematik didaktischer Prinzipien, die geeignet ist, sowohl die theoretischen Konzepte wie auch die Wirkungsannahmen der Unterrichtsforschung zu vereinen (Klieme, 2008). Diese drei fachübergreifenden Dimensionen konnten sich im deutschsprachigen Raum in den letzten Jahren in der Bildungs- und Unterrichtsforschung weiter etablieren (Praetorius et al., 2018).

Im internationalen Vergleich unterschiedlicher fachübergreifender und mathematikspezifischer Konzeptionen von Unterrichtsqualität wurden weiterführende Syntheseframeworks erarbeitet (u. a. Praetorius & Charalambous, 2018), welche die drei Basisdimensionen erweitern und ausdifferenzieren. Diese Syntheseframeworks bieten sich für fachspezifische Ausdifferenzierungen an, so dass Unterschiede und Gemeinsamkeiten zwischen der fachübergreifenden und fachspezifischen Unterrichtsqualitätsforschung herausgearbeitet werden konnten (Praetorius et al., 2020; für das Fach Sport: Herrmann & Gerlach, 2020). Jedoch handelt es sich dabei wiederum lediglich um Merkmalslisten, weshalb aktuell versucht wird, die verschiedenen Unterrichtsmerkmale in Dimensionen und Subdimensionen in komplexere Unterrichtsqualitätsmodelle zu übertragen, um die Beziehungen zwischen den (Sub-)Dimensionen zu klären (vgl. MAIN-TEACH-Modell, Charalambous & Praetorius, 2020). Dabei bleibt (noch) ungeklärt, wie fachübergreifend oder fachspezifisch diese (Sub-) Dimensionen der Unterrichtsqualität sind. Entsprechend wird in der Verknüpfung der bislang vorrangig fachübergreifend ausgerichteten Unterrichtsqualitätsforschung mit den fachdidaktischen Diskursen in den Schulfächern ein zentrales zukünftiges Handlungsfeld der Unterrichtsqualitätsforschung gesehen. Diese vielversprechenden theoretisch-konzeptionellen Arbeiten konnten jedoch empirisch noch nicht umgesetzt werden, so dass im deutschsprachigen Raum immer noch die drei Basisdimensionen den meisten Operationalisierungen zu Grunde liegen (Begrich et al., 2023).

2.2 Unterrichtsqualitätsforschung im Fach Sport

Die Frage, was als qualitativ hochwertiger Fachunterricht bezeichnet werden kann, hängt in hohem Masse davon ab, welche Erziehungs- und Lernziele erreicht werden sollen. In der deutschsprachigen Sportpädagogik und -didaktik konnte sich das Konzept der *Handlungsfähigkeit im Schulsport* (Kurz, 2013) als pädagogische Leitidee etablieren. Darin sind eine qualifikationsbezogene und eine reflexive Handlungsfähigkeit der Schüler*innen als zentrale Erziehungsziele verankert (Gogoll, 2013). Die *qualifikationsbezogene Handlungs-*

fähigkeit wird durch Kompetenzen begründet, die sich auf Anforderungen der Sport- und Bewegungskultur beziehen. Diese sollen gewährleisten, dass die Schüler*innen sportliche Bewegungshandlungen ausführen und qualifiziert an der Sport- und Bewegungskultur partizipieren können. Die *reflexive Handlungsfähigkeit* wird durch Kompetenzen begründet, die im Zusammenhang mit der reflexiven Auseinandersetzung eigener sportbezogener Welt- und Selbstsichten stehen. Diese versetzen die Schüler*innen in die Lage, Bewusstheit über ihr Tun zu erlangen, eigenständig vernünftige Entscheidungen zu treffen und selbstbestimmt handeln zu können. Da dieses Konzept ebenfalls dem Deutschschweizer Lehrplan 21 (E-EDK, 2017) zu Grunde liegt und im sportdidaktischen Diskurs in der Schweiz zunehmend ausdifferenziert wird, werden die Merkmale des QUALLIS-Instruments zur Evaluation der Unterrichtsqualität im Fach Sport vor dem normativen Hintergrund dieses fachdidaktischen Konzeptes diskutiert. In unserem Verständnis bildet die qualifikationsbezogene und reflexive Handlungsfähigkeit ein Kontinuum an erzieherischen Ansprüchen ab, deren Thematisierung im Sportunterricht unterschiedliche (fachspezifische) Lernprozesse bei den Schüler*innen auslöst.

Arbeiten Schüler*innen im Sportunterricht an der qualifikationsbezogenen Handlungsfähigkeit, so steht vor allem das **Bewegungslernen** (u.a. Künzell et al., 2023) im Fokus, um (implizites) Bewegungskönnen zu erwerben (z.B. Werfen, Prellen, Rollen können). Sobald sich Schüler*innen mit ihrer reflexiven Handlungsfähigkeit auseinandersetzen, spielt das **verständnisbezogene Lernen** (u.a. Hapke & Waigel, 2019) eine zentrale Rolle, um (explizites) Handlungswissen aufzubauen (z.B. Wissen über Trainingsprozesse und deren Folgen). Dazwischen steht eine Verschränkung von Bewegungslernen und verständnisbezogenen Lernen, welches sich bspw. in einem **spieltaktischen Lernen** (u.a. Raab et al., 2009) äussert und vor allem den Aufbau von (prozeduralem) sport- und bewegungsbezogenen Wissens dient (z.B. spieltaktisches Verständnis anwenden) (vgl. auch Richartz & Kohake, 2021). Entsprechend erscheint es uns zweckmässig die Unterrichtsdimension der kognitiven Aktivierung im Fach Sport entsprechend der Lernziele und Lerngegenstände in eine kognitiv-motorische Aktivierung auszudifferenzieren und zu erweitern.

Neben fachspezifischen Erziehungs- und Lernzielen sind bei der Ausdifferenzierung der Unterrichtsmerkmale im Fach Sport u.a. auch die räumlich-materiellen Bedingungen des Unterrichtsgeschehen (Unterricht in der Sporthalle) als Spezifika zu berücksichtigen (Heemsoth & Krieger, 2018; Herrmann et al., 2016). In bisherigen sportbezogenen Übersichtsarbeiten (Herrmann et al., 2020; Wibowo et al., 2021) konnten die Basisdimensionen *Klassenführung* und *Unterrichtsklima* bereits sportunterrichtsspezifisch angepasst und ausdifferenziert werden (Baumgartner et al., 2020; Heemsoth, 2014). So zeigte sich insgesamt, dass die Unterrichtsdimensionen weitestgehend Gültigkeit besitzen, sie jedoch fachspezifisch ausdifferenziert und operationalisiert werden müssen, um den Besonderheiten des Faches Sport gerecht zu werden. Zentral ist hierbei die konzeptionelle Erweiterung der *Aktivierung* (Herrmann & Gerlach, 2020, S. 380).

2.3 Merkmale qualitätvollen Sportunterrichts

Im Folgendem werden ausgehend von den drei Basisdimensionen zentrale Unterrichtsmerkmale für das Fach Sport benannt, welche für die Entwicklung des QUALLIS-Instruments fachspezifisch ausdifferenziert und ergänzt wurden. Dabei ist hervorzuheben, dass diese Auflistung an Unterrichtsmerkmalen keinen Anspruch auf Vollständigkeit erheben kann, die Merkmale mit unterschiedlicher theoretischer Tiefe in den sportdidaktischen Diskurs eingeführt worden sind und die Interaktionen zwischen den Merkmalen theoretisch und empirisch noch nicht vollständig geklärt sind. Da die Unterrichtsdimensionen *Klassenführung* (u.a. Baumgartner et al., 2020, S. 513–515) und *Unterrichtsklima* (u.a. Heemsoth & Miethling, 2012, S. 228–230) in vorliegenden Studien theoretisch wie empirisch für das Fach Sport bereits ausdifferenziert wurden, werden die im QUALLIS-Instrument genutzten Merkmale nur knapp erläutert. Für die *kognitiv-motorische Aktivierung* liegt dagegen noch keine theoretisch fundierte Auswahl von Merkmalen und deren Operationalisierung vor, weshalb diese Unterrichtsdimension umfassender dargestellt wird.

Die *Klassenführung* fasst Merkmale der klaren, strukturierten und gut organisierten Instruktion zusammen. Guter Unterricht ist durch eine Klassenführung gekennzeichnet, die den Schüler*innen eine intensive Nutzung von Lernzeit mit vielfältigen Lerngelegenheiten ermöglicht und bei dem möglichst wenig Disziplinprobleme auftreten (Haag & Streber, 2012). Lehrpersonen gewährleisten dies durch eine proaktive Steuerung des Unterrichts, indem sie präventiv oder intervenierend auf Unterbrechungen und Disziplinproblemen eingehen. Dadurch können Unterrichtsstörungen effektiv minimiert und die Aufmerksamkeit der Schülerinnen und Schüler gesichert werden (Borich, 2007; Kounin, 1976). Merkmale einer guten Klassenführung im Fach Sport sind eine gute Organisation und *Zielklarheit* des Sportunterrichts (ZIEL), klar formulierte und verbindliche Regeln, eine optimale *Zeitnutzung* (ZEIT) ohne *Disziplinprobleme* (DISZ) sowie eine Gewährleistung der

Sicherheit und Minimierung der Verletzungsgefahr (ausführlich Baumgartner et al., 2020; Herrmann, 2021; vgl. Tabelle 1).

Das **Unterrichtsklima** umschreibt Merkmale, die das Eingehen auf individuelle Lernpotenziale und Bedürfnisse der einzelnen Schüler*innen kennzeichnen. Hierbei stehen die sozial-emotionalen Unterrichtsmerkmale im Fokus, welche für das Unterrichtsklima, die Lern- und Interessensentwicklung sowie für die intrinsische Lernmotivation bedeutsam sind (Gräsel & Göbel, 2011). Das schüler*innenorientierte Unterrichtsklima im Fach Sport ist geprägt von einer positiven Lehrer*innen-Schüler*innen-Beziehung, in der die Lehrpersonen *fürsorglich* (FSLP) handeln sowie individuelles und wertschätzendes *Feedback* (FEED) geben. Sie *unterstützen* (UNT) die Schüler*innen emotional im Lernprozess, nehmen Differenzierungen nach Vorerfahrung und Lernvoraussetzungen vor und orientieren ihr Handeln an den Interessen der Schüler*innen (ausführlich Heemsoth & Miethling, 2012; Herrmann, 2021; vgl. Tabelle 1).

Unter **kognitiv-motorischer Aktivierung** werden Merkmale beschrieben, die die Schüler*innen zur elaborierten Auseinandersetzung mit dem Lerngegenstand und damit zum Lernen anregen. Mit dem Ziel der Handlungsfähigkeit wären dies insbesondere bewegungs- und verständnisbezogene Lernprozesse. Aktuell besteht darüber ein breiter sportdidaktischer Diskurs mit unterschiedlichen theoretischen und methodischen Zugängen (im Überblick Wibowo et al., 2021). Der vorliegende Beitrag verfolgt den Theorieansatz von Niederkofler und Amesberger (2016), welcher die Aktivierung im Sportunterricht über eine handlungstheoretische Perspektive konzipiert.

Handlungstheorien legen nahe, dass Bewegungshandlungen kognitiv, motorisch und emotional reguliert sind (u.a. Munzert, 1995). Eine Kernannahme bezieht sich darauf, dass Menschen aufgrund von subjektiv bedeutsamen Zielen im situativen Kontext handeln und sich über diese Handlungen weiterentwickeln. Der Bewegungshandlung wird zudem eine triadische Phasenstruktur zugeschrieben, mit der alle Regulationsebenen angesprochen werden: in der *Antizipation* die Handlungsorientierung, in der *Realisation* die Handlungsdurchführung und in der *Interpretation* die Handlungsbewertung. Das würde bedeuten, dass sich alle regulativen Ebenen aktivieren und weiterentwickeln, sobald Schüler*innen im Sportunterricht zu Bewegungshandlungen aufgefordert werden (vgl. Niederkofler & Amesberger, 2016, 2023).

Für die Auseinandersetzung mit der Handlungsfähigkeit im Sportunterricht muss ein komplexes Handlungsgefüge angenommen werden, in dem die Situationswahrnehmung von Schüler*innen eine handlungsleitende Rolle einnimmt: von der situativen Wahrnehmung der Aufgabe, der Umwelt (z.B. der Lehrperson) sowie der eigenen Person hängt ab, was Schüler*innen denken und fühlen sowie welche (Bewegungs-)Handlungen zur Weiterentwicklung der Handlungsfähigkeit gesetzt werden (vgl. Nitsch, 2004). Entsprechend wird für die kognitiv-motorische Aktivierung gefolgert, dass die Situationswahrnehmung die Unterrichtsqualität mitbestimmt. Folgend werden die im QUALLIS-Instrument genutzten Merkmale begründet und eingeordnet (siehe auch Niederkofler & Amesberger, 2016; Niederkofler & Herrmann, 2022; vgl. Tabelle 1):

- Schüler*innen planen Bewegungshandlungen bevor sie durchgeführt und ausgewertet werden. Dabei kann es für die Zielerreichung zweckmässig sein, die Aufgabenanforderungen (im Sinne der für die Bewältigung notwendigen Kompetenzen) in der *Antizipation* zu analysieren und darauf aufbauend einen Handlungsplan zu entwickeln. Dies erleichtert in weiterer Folge auch die Handlungsbewertung. Derartige **Verarbeitungsprozesse** (VERA) sind für die Unterrichtsqualität mitzudenken.
- Aus einer handlungstheoretischen Perspektive bleibt die Situationsanalyse allerdings nicht bei der Betrachtung von Aufgabenanforderungen stehen. Die Anforderungen werden mit den eigenen Handlungskompetenzen abgeglichen, bevor versucht wird, die Aufgabe mit aktualisierbaren Kompetenzen zu bewältigen. Dies verweist auf das qualitätsbestimmende Merkmal des **Anspruchsniveaus** (ANSP).
- Es ist von entscheidender Bedeutung, die Aufmerksamkeit im Handlungsprozess gezielt auf handlungsrelevante Aspekte zu **fokussieren** (FOKU). Dazu kann es zweckdienlich sein, sich relevante Übungsmerkmale bewusst zu machen und diese in die Handlungsplanung zu integrieren. Ein derartiger Fokus kann die Orientierung auf die Handlung erhöhen und die wahrgenommene Bewegungskontrolle positiv beeinflussen.
- Häufig werden Aufgaben im Sportunterricht von der Lehrperson formuliert. Dies löst bei Schüler*innen unterschiedliche Prozesse aus, u.a. auch die Frage, was man selbst bei dieser Aufgabe erreichen möchte. Eine derartige Intentionbildung hängt auch davon ab, wie **verständlich** (VERS) die Erklärungen der Lehrperson sind und welche Konsequenzen die Schüler*innen daraus ableiten.
- Die Situationswahrnehmung hängt weiter davon ab, welchen Anregungsgehalt eine Aufgabe hat und welche Umstände Schüler*innen mit der Bewegungshandlung verbinden. Wird der Übungskontext so **strukturiert**

riert (STRU), dass Schüler*innen im Rahmen ihrer realisierbaren Kompetenzen ausreichend Gestaltungsspielraum erleben, kann sich dies förderlich auf die Umsetzung von Bewegungshandlungen auswirken.

- Die angenommene triadische Phasenstruktur von Bewegungshandlungen verweist darauf, dass Schüler*innen nach der *Realisation* einschätzen, wie die Durchführung gelungen ist und was es zu merken gilt, um für künftige Bewegungshandlungen die Pläne entsprechend anzupassen. Eine **Rückmeldung (RUCK)** der Lehrperson kann in diesem Rahmen als qualitativvoll erlebt werden, sobald sich Schüler*innen in der Erreichung ihrer subjektiven Ziele konstruktiv unterstützt wahrnehmen.

Ausgehend von diesen theoretischen Überlegungen zu Merkmalen qualitätsvollen Sportunterrichts verfolgt der vorliegende Beitrag das Ziel, ein ökonomisch einsetzbares Fragebogeninstrument zur Evaluation des Sportunterrichts zu entwickeln, um Lehr- und Lernprozesse kriteriengeleitet beurteilen und (weiter-)entwickeln zu können. Hierfür werden im Folgendem bereits durchgeführte Entwicklungs- und Validierungsstudien dargestellt und das QUALLIS-Instrument abgeleitet, welches es ermöglicht den Sportunterricht aus der Sicht der Schüler*innen sowie deren Lehrpersonen zu beurteilen.

Die psychometrische Prüfung des Instruments (angelehnt an Bühner, 2011) fokussiert folgende Forschungsfragen:

1. Sind die Unterrichtsmerkmale faktoriell valide zu erfassen? (Konstruktvalidität)
2. Existieren Zusammenhänge von Unterrichtsmerkmalen mit Merkmalen der Schüler*innen? (Kriteriumsvalidität)
3. Ist die Erfassung der Unterrichtsmerkmale auf Schüler*innen-, Klassen- und Lehrpersonenebene reliabel?
4. Welche Zusammenhänge zwischen Lehrpersoneneinschätzungen und Schüler*inneneinschätzungen von Unterrichtsmerkmalen sind vorhanden?

3. Methode

Unterrichtsqualität wird durch die Einschätzung von Lehr- und Lernprozessen entlang von Evaluationskriterien erfasst und spiegelt damit unterschiedliche Perspektiven einzelner Personen (Lehrpersonen, Schüler*innen, Beobachter*innen) auf bestimmte Unterrichtsmerkmale. Die Einschätzungen durch **Schüler*innen** mittels Fragebögen stellt eine ökonomische Erhebungsweise dar, mit der auch grössere Stichproben erfasst werden können. Gleichzeitig spiegeln die entsprechenden Daten – im Vergleich zu punktuellen Beobachtungen einzelner Unterrichtseinheiten – einen relativ langen Beurteilungszeitraum wider. Durch das Aggregieren der Schüler*innendaten auf Klassenebene gewinnen die Schüler*innenratings an Reliabilität (Lüdtke et al., 2006). Die Einschätzungen durch die **Lehrpersonen** bieten insbesondere für die Unterrichtsbeurteilung und Unterrichtsentwicklung die Möglichkeit, die subjektiven Sichtweisen der Lehrpersonen auf den Unterricht – auch im Vergleich zu den Schüler*inneneinschätzungen – zu erfassen. Einschränkend dabei ist, dass die Lehrpersonen als Akteure im Unterricht ihre eigenen Handlungen nur schwer neutral beurteilen können, weshalb Verzerrungen in der Selbstwahrnehmung des Unterrichtsgeschehens eintreten können. Einschätzungen durch geschulte externe **Beobachter*innen** wird dagegen eine höhere Validität zugesprochen, gehen jedoch mit einem sehr hohen Zeit- und Kostenaufwand einher und beziehen sich nur auf einen punktuellen Beobachtungszeitraum (Praetorius, 2014). Im vorliegenden Beitrag wurden die Einschätzungen der Unterrichtsmerkmale von Schüler*innen sowie deren Lehrpersonen mittels Fragebögen vorgenommen.

3.1 Entwicklung des Instruments

In den letzten Jahren wurden mehrere aufeinander aufbauende Studien durchgeführt, um ein Erhebungsinstrument zur Erfassung der Unterrichtsqualität im Fach Sport zu entwickeln. Hierbei sind vornehmlich Einschätzungen von Schüler*innen genutzt worden, um die Unterrichtsqualität zu erfassen. In der ersten Studie wurden zunächst bestehende Erhebungsinstrumente aus der Unterrichtsforschung inhaltlich und sprachlich für das Fach Sport adaptiert. Mit einer Stichprobe von $N = 884$ Schüler*innen ($M = 13.2$ Jahre, $SD = 0.65$) wurde ein Instrument zur Erfassung der beiden Unterrichtsdimensionen *Klassenführung* und *Unterrichtsklima* entwickelt und validiert, welches auf Basis von 34 Items insgesamt acht Unterrichtsmerkmale erfasst. Die Dimension *Aktivierung* konnte in dieser Studie nicht berücksichtigt werden, da eine fachspezifische Bestimmung und theoretische Fundierung für *Aktivierung* im Fach Sport noch nicht vorlag (ausführlich Herrmann et al., 2015). Nach einer theoretischen Ausdifferenzierung der Unterrichtsdimension *Aktivierung* (Herrmann, 2021; Nieder-

kofler & Amesberger, 2016) wurden in weiteren Studien sukzessive Merkmale wie *Anspruchsniveau*, *Verständlichkeit* und *Rückmeldung* ergänzt. Zusätzlich wurde die Beschreibung der Merkmale zunehmend präzisiert, die Zuordnung der Merkmale zu den Dimensionen neu sortiert sowie die Anzahl der Items auf ein Minimum von je drei Items pro Merkmal reduziert (u. a. Herrmann & Gerlach, 2017; Herrmann et al., 2018).

Auf Basis dieser Vorarbeiten wird im vorliegenden Beitrag die Validierung des QUALLIS-Instruments zur Evaluation der Qualität des Lehrens und Lernens im Sport vorgestellt. Dieses Instrument soll zukünftig im schulischen sowie in einer angepassten Version auch im außerschulischen Kontext Verwendung finden. Die folgende Prüfung der psychometrischen Eigenschaften des QUALLIS-Instruments wird jedoch zunächst ausschliesslich den Kontext Schule berücksichtigen. Die Adaption und Validierung des QUALLIS-Instruments für den außerschulischen Sport in der Schweiz findet sich bei Herrmann et al. (under review).

3.2 QUALLIS-Instrument

Das QUALLIS-Instrument für den Schulsport bezieht sich auf die drei Unterrichtsdimensionen *Klassenführung*, *Unterrichtsklima* und kognitiv-motorische *Aktivierung* und umfasst insgesamt zwölf Unterrichtsmerkmale, welche jeweils mit drei Items erfasst werden (Tabelle 1). Das QUALLIS-Instrument ist konzipiert, um die Unterrichtsqualität im Fach Sport aus drei Perspektiven zu beschreiben. Erfragt werden kann die Einschätzung von Sportunterricht durch (a) *Schüler*innen*, (b) *Lehrpersonen* und (falls vorhanden) (c) «nicht-teilnehmende» *Beobachter*innen*. Durch Vergleiche von Einschätzungen der Schüler*innen (Beispielitem: „Unser*e Sportlehrer*in lobt uns, wenn wir unsere Leistung verbessern.“), Einschätzungen der unterrichtenden Lehrperson (Beispielitem: „Ich als Sportlehrer*in lobe die Schüler*innen, wenn sie ihre Leistung verbessern.“) sowie Einschätzung eines potenziell vorhandenen «externen» Beobachters (Beispielitem: „Der/Die Sportlehrer*in lobt die Schüler*innen, wenn sie ihre Leistung verbessern.“) wird eine mehrperspektivische Gesamtbewertung des Unterrichts ermöglicht (Helmke & Lenske, 2013). Der vorliegende Beitrag fokussiert zunächst eine Validierung des Schüler*innen- und Lehrpersonenfragebogens.

Die Items besitzen für jedes Unterrichtsmerkmal einheitliche Aussageeinleitungen (Schüler*innenfragebogen: „Im Sportunterricht ...“/„Unser*e Sportlehrer*in ...“; Lehrpersonenfragebogen: „In meinem Sportunterricht...“/„Ich als Sportlehrer*in...“) und sollen anhand einer 5-stufigen Antwortskala (1 = „trifft nicht zu“ bis 5 = „trifft zu“) eingeschätzt werden, wie sehr sie die beschriebene Aussage als zutreffend empfinden.

Tabelle 1

Übersicht der QUALLIS-Merkmale mit Beispielitems aus Schüler*innen- und Lehrpersonenperspektive (vollständige Itembeschreibung mit deskriptiven Werten im Anhang).

	Merkmale	Einschätzungen der Schüler*innen	Einschätzungen der Lehrpersonen
Klassenführung	Zeitnutzung (ZEIT)	Im Sportunterricht wird viel Zeit vertrödelt. (-)	In meinem Sportunterricht wird viel Zeit vertrödelt. (-)
	Disziplin (DISZ)	Im Sportunterricht gibt es viele Störungen durch die Schüler*innen. (-)	In meinem Sportunterricht gibt es viele Störungen durch die Schüler*innen. (-)
	Zielklarheit (ZIEL)	Unser/e Sportlehrer/in macht uns die Ziele der Unterrichtsstunde klar. (+)	Ich als Sportlehrer/in mache den Schülern/innen die Ziele der Unterrichtsstunde klar. (+)
Unterrichtsklima	Feedback (FEED)	Unser*e Sportlehrer*in lobt uns, wenn wir unsere Leistung verbessern. (+)	Ich als Sportlehrer*in lobe die Schüler*innen, wenn sie ihre Leistung verbessern. (+)
	Fürsorglichkeit (FSLP)	Unser*e Sportlehrer*in schenkt allen Schüler*innen Beachtung. (+)	Ich als Sportlehrer*in schenke allen Schüler*innen Beachtung. (+)
	Unterstützung (UNT)	Unser*e Sportlehrer*in unterstützt uns, wenn wir Fehler machen. (+)	Ich als Sportlehrer*in unterstütze die Schüler*innen, wenn sie Fehler machen. (+)
Kognitiv-motorische Aktivierung	Fokussierung (FOKU)	Im Sportunterricht überlege ich vorher, wie ich die Übungen ausführe. (+)	In meinem Sportunterricht rege ich die Schüler*innen an, sich vorher zu überlegen, wie sie die Übungen ausführen. (+)
	Verarbeitung (VERA)	Im Sportunterricht gehe ich die Übungen im Kopf durch. (+)	In meinem Sportunterricht rege ich die Schüler*innen dazu an, die Übungen im Kopf durchzugehen. (+)
	Verständlichkeit (VERS)	Im Sportunterricht verstehe ich die Erklärungen unseres*r Sportlehrers*in. (+)	Ich als Sportlehrer*in achte darauf, die Übungen verständlich zu erklären. (+)
	Anspruchsniveau (ANSP)	Unser*e Sportlehrer*in lässt uns Übungen machen, bei denen ich mich richtig anstrengen muss. (+)	Ich als Sportlehrer*in lasse die Schüler*innen Übungen machen, bei denen sie sich richtig anstrengen müssen. (+)
	Strukturierung (STRU)	Unser*e Sportlehrer*in baut die Übungen Schritt für Schritt auf. (+)	Ich als Sportlehrer*in baue die Übungen Schritt für Schritt auf. (+)
	Rückmeldung (RUCK)	Unser*e Sportlehrer*in gibt uns Hinweise, um die Übungsausführung zu verbessern. (+)	Ich als Sportlehrer*in gebe den Schülern*innen Hinweise, um die Übungsausführung zu verbessern. (+)

Kodierung: 1 = trifft nicht zu, 2 = trifft eher nicht zu, 3 = trifft teils teils zu, 4 = trifft eher zu, 5 = trifft zu

3.3 Ergänzende Instrumente

Im Schüler*innen-Fragebogen wurde zu Beginn das *Geschlecht* (Jungen = 0, Mädchen = 1) und das *Alter* (in Jahren) der Schüler*innen erfasst. Dies aufgrund vorliegender Studien (im Überblick Herrmann et al., 2020), bei denen sich geschlechts- und altersspezifische Unterschiede in den Einschätzungen von Sportunterricht zeigten: Zum einen fielen die Werturteile über den Sportunterricht bei Mädchen negativer aus als bei Jungen. Zum anderen wurde der Sportunterricht von den Schüler*innen mit zunehmendem Alter schlechter beurteilt (Deutscher Sportbund, 2006; Lamprecht et al., 2014).

Zur Einschätzung der *Kriteriumsvalidität* wurden die auf das Fach Sport bezogenen motivationalen Aspekte *Interesse am Sportunterricht* (3 Items, Beispielitem: „Der Sportunterricht ist für mich wichtig.“ $M = 3.73$, $SD = 1.18$, $C's \alpha = .89$, $ICC = .189$) und *Anstrengungsbereitschaft im Sportunterricht* (4 Items, Beispielitem: „Im Sportunterricht versuche ich mich immer anzustrengen.“ $M = 3.80$, $SD = 0.95$, $C's \alpha = .91$, $ICC = .114$) auf Seiten der Schüler*innen anhand einer 5-stufigen Antwortskala (1 = „trifft nicht zu“ bis 5 = „trifft zu“) erfasst. Diese motivationalen Aspekte haben sich in verschiedenen Studien als wichtige Kriteriumsvariablen für die Unterrichtsmerkmale erwiesen (z.B. Heemsoth, 2014; Klieme & Rakoczy, 2008).

3.4 Stichprobe

An der Befragung nahmen auf freiwilliger Basis sportunterrichtende Lehrpersonen mit ihren Schüler*innen aus den Kantonen Luzern, Basel-Land und Basel-Stadt teil. Die Schüler*innen hatten verpflichtend drei Sport-

stunden (jeweils 45 Minuten) pro Woche. Die gedruckten Fragebögen waren vollständig anonymisiert (ohne Vor- und Nachname, ohne Geburtsdatum) und wurden zu Beginn des Sportunterrichts unter der Anleitung einer wissenschaftlichen Hilfskraft ausgefüllt. Die Schüler*innen wurden explizit darauf hingewiesen, dass die Teilnahme freiwillig ist und die Lehrpersonen keine Einsicht in die einzelnen Schüler*innen-Fragebögen erhalten. Die Bearbeitungsdauer lag zwischen zehn und 15 Minuten.

Es wurden querschnittliche Daten zu zwölf Unterrichtsmerkmalen in 40 Klassen der siebten bis neunten Jahrgangsstufe mit $N = 634$ Schüler*innen (davon 42.8 % Jungen, Alter: $M = 13.9$ Jahre, $SD = 1.13$) an einem Messzeitpunkt erfasst. Die mittlere Klassengröße lag bei $M = 15.8$ Schüler*innen. Die unterrichtenden Lehrpersonen waren im Mittel $M = 46.0$ Jahre ($SD = 10.5$) alt, 42.5 % von ihnen waren männlich.

3.5 Datenauswertung

Nach dem Rücklauf der Fragebögen, wurden diese eingescannt und digitalisiert. Dies erfolgte mit dem Programm EvaSys®, welches eine (semi-) automatisierte und standardisierte Weiterverarbeitung der Daten ermöglichte. Häufigkeits- und Korrelationsanalysen wurden mit SPSS 25, die latenten Strukturgleichungsmodelle mit Mplus 8.3 und die Mehrebenenanalysen mit HLM 7.03 durchgeführt. Bei allen Strukturgleichungsmodellen wurden potenzielle Abhängigkeiten innerhalb der vorhandenen Mehrebenenstruktur (Schüler*innen in Klassen) berücksichtigt, indem die relevanten Standardfehler mit der Mplus-Funktion *type = complex* für geschachtelte Datensätze angepasst wurden. Dabei wurde der Maximum Likelihood Robust (MLR)-Schätzer verwendet. Der Anteil an fehlenden Werten pro Item lag zwischen 0.0% und 2.2 %. Einzelne fehlende Werte wurden anhand des *Full-Information-Maximum-Likelihood* (FIML)-Algorithmus geschätzt. Die Beurteilung der Anpassungsgüte der verschiedenen Modelle erfolgte auf Basis der in der Literatur vorgeschlagenen Fit-Indices (Weiber & Mühlhaus, 2014).

Die Datenauswertung teilte sich in die psychometrische Überprüfung (angelehnt an Bühner, 2011; Geiser, 2011) des QUALLIS-Fragebogens der Schüler*innen und der Lehrpersonen sowie des Vergleichs zwischen den Perspektiven. Vor dem Hintergrund der theoretischen Annahmen wurde das QUALLIS-Instrument hinsichtlich der faktoriellen Validität, als zentraler Teil der Konstruktvalidität, sowie der Kriteriumsvalidität empirisch geprüft.

Von faktorieller Validität kann dann gesprochen werden, wenn die Zusammenhänge der Testitems auf gemeinsame, nicht direkt beobachtbare Konstrukte (hier: Unterrichtsmerkmale) zurückgeführt werden können und damit die angenommene Dimensionsstruktur gestützt werden kann (Hartig et al., 2012). Bei der Prüfung der Kriteriumsvalidität handelt sich hier um den Zusammenhang der Einschätzung der Unterrichtsmerkmale mit Kriterien (u.a. Motivation), mit denen die Konstrukte aufgrund ihres Messanspruches korrelieren sollten (Bühner, 2011).

Während die Überprüfung des Schüler*innenfragebogens über latente Strukturgleichungsmodelle erfolgen konnte, musste die Überprüfung des Lehrpersonenfragebogens aufgrund der geringen Stichprobengröße auf manifester Ebene erfolgen. Die Analyseschritte orientieren sich an den eingangs formulierten vier Fragestellungen:

- (1) *Konstruktvalidität und Faktorreliabilität*: Mit den 36 Items des Schüler*innenfragebogens wurde eine konfirmatorische Faktorenanalyse mit zwölf Faktoren berechnet. Die Zuordnung der jeweils drei Items zu einem Faktor erfolgte auf Basis der konzeptionellen Vorarbeiten (Tabelle 1). Die Faktorladungen der Items auf die latenten Faktoren wurden frei geschätzt. Aufgrund der umfassenden Vorstudien wurden keine explorativen Strukturgleichungsmodelle (ESEM, vgl. Herrmann et al., 2015) zur Exploration der Datenstruktur vorgenommen. Zur Überprüfung der Reliabilität (auf latenter Ebene) wurden die Faktorreliabilitäten (FR) der zwölf Merkmale berechnet. Die Faktorreliabilität gilt als Mass der Reliabilität über die Gesamtsumme aller Items, die ein Konstrukt bilden und sollten Werte grösser als .60 annehmen. Zur weiteren Einschätzung der Reliabilität wurden die entsprechenden Cronbach's α -Werte (auf manifester Ebene) berechnet, welche Werte grösser als .70 annehmen sollten (Weiber & Mühlhaus, 2014).
- (2) *Kriteriumsvalidität*: Die konfirmatorische Faktorenanalyse wurde um die manifesten Korrelate *Geschlecht* und *Alter* sowie um die latent modellierten Korrelate *Anstrengungsbereitschaft im Sportunterricht* und *Interesse am Sportunterricht* ergänzt.
- (3) *Reliabilität auf Klassen- und Lehrpersonenebene*: Hierfür wurden zunächst die Intraklassen-Korrelationen ICC (1) und ICC (2) für die Unterrichtsmerkmale berechnet. Die ICC (1) gibt den prozentualen Anteil der Gesamtvarianz an, der auf Unterschiede zwischen den Klassen zurückzuführen ist. Hohe Werte in der

ICC (1) bedeuten, dass es bei dem entsprechenden Merkmal grosse Unterschiede zwischen Klassen gibt. Die ICC (2) gibt die Genauigkeit des über alle Schüler*innen einer Klasse gemittelten Urteils an und gilt als Mass der Reliabilität des Klassenmittelwertes der Schülerinnenurteile (Lüdtke et al., 2006). Zur Bewertung der Lehrpersoneneinschätzungen wurden Cronbach's α -Werte sowie die Pearson-Korrelationen zwischen den von den Lehrpersonen eingeschätzten Unterrichtsmerkmalen berechnet.

(4) *Zusammenhänge zwischen Schüler*innen- und Lehrpersoneneinschätzungen:* Mittels Mehrebenenanalysen wurde geprüft, ob die Lehrpersoneneinschätzungen der Unterrichtsmerkmale (Klassenebene) die Schüler*inneneinschätzungen der Unterrichtsmerkmale (Individualebene) präzisieren können. Die Mehrebenenanalysen wurden als Random-Intercept-Modelle berechnet. Dabei wurden alle Variablen vorgängig z-standardisiert, um die Zusammenhänge als standardisierte β -Werte interpretieren zu können (Raudenbush & Bryk, 2002). Für jedes Unterrichtsmerkmal wurde ein eigenes Modell berechnet.

4. Ergebnisse

Entlang der vier Hauptfragestellungen und der statistischen Analyseschritte werden folgend die Ergebnisse der Prüfung der faktoriellen Validität und der Reliabilitäten der Schüler*inneneinschätzungen (4.1) sowie die Korrelationen zwischen den Unterrichtsmerkmalen und Kriteriumsvariablen (4.2) dargestellt. Darauf aufbauend werden die Reliabilitäten der klassenweise aggregierten Schüler*inneneinschätzungen und der Lehrpersoneneinschätzungen berichtet (4.3). Abschliessend erfolgt der Vergleich zwischen den individuellen Schüler*inneneinschätzungen mit den Lehrpersoneneinschätzungen (4.4).

4.1 Konstruktvalidität und Faktorreliabilität

Zur Prüfung der faktoriellen Validität des QUALLIS-Schüler*innenfragebogens wurden in einer konfirmatorischen Faktorenanalyse die 36 Items den zwölf Merkmalen zugeordnet (Tabelle 1). Hierbei zeigte sich eine sehr gute Anpassung der zwölf-faktoriellen Struktur auf die vorliegenden Daten ($\text{Chi}^2 = 817.47$, $\text{df} = 528$, $p < .001$; $\text{CFI} = .966$, $\text{TLI} = .959$, $\text{RMSEA} = .029$, $\text{SRMR} = .043$).

Die resultierenden Faktorwerte lagen bei $.56 \leq \beta \leq .90$ (Tabelle 2, ausführlich Anhang 1). Die errechneten Faktorreliabilitäten waren mit $.68 \leq \text{FR} \leq .87$ als gut bis sehr gut zu beurteilen (Tabelle 2). Die interne Konsistenz der zwölf Merkmale – als Mass der Reliabilität auf manifester Ebene – war mit $.72 \leq \alpha \leq .87$ ebenfalls durchgängig gut bis sehr gut (Tabelle 2). Auf Basis der gegebenen Reliabilität und faktoriellen Validität der Messung von Merkmalen konnte gefolgert werden, dass die Bildung von (manifesten) Skalenmittelwerten zulässig ist. Die Skalenmittelwerte lagen zwischen $2.99 \leq M \leq 4.12$. Die Merkmale *Feedback*, *Verständlichkeit* und *Rückmeldung* wurden mit Mittelwerten von $4.08 \leq M \leq 4.12$ relativ hoch eingeschätzt, während *Disziplin* mit $M \leq 2.99$ am niedrigsten bewertet wurde (Tabelle 2).

4.2 Interkorrelationen der Unterrichtsmerkmale und Kriteriumsvalidität

In der konfirmatorischen Faktorenanalyse mit den Kovariaten Alter, Geschlecht, Interesse und Anstrengungsbereitschaft ($\text{Chi}^2 = 1369.51$, $\text{df} = 827$, $p < .001$; $\text{CFI} = .954$, $\text{TLI} = .945$, $\text{RMSEA} = .032$, $\text{SRMR} = .044$) wurden die Interkorrelationen der durch die Schüler*innen eingeschätzten Merkmale berechnet sowie Zusammenhänge mit potenziellen Kriteriumsvariablen geprüft, um die Kriteriumsvalidität abschätzen zu können.

Die Interkorrelationen der Unterrichtsmerkmale zeigten grösstenteils mittlere bis hohe Zusammenhänge (Tabelle 3). Entsprechend konnte gezeigt werden, dass die Merkmale zwar verbunden sind, jedoch ausreichend trennscharf abgebildet werden können. Die Merkmale, welche aus den theoretischen Überlegungen heraus einer Basisdimension zugeordnet wurden (Tabelle 1), verfügten zwar über teils hohe Zusammenhänge. Ein klares Bild, dass diese Merkmale prinzipiell höher miteinander korrelierten als mit Merkmalen anderer Basisdimensionen, konnte jedoch nicht durchgängig erkannt werden. Beispielsweise stand die *Zielklarheit* in keinem oder nur kleinen Zusammenhang mit den zwei weiteren Merkmalen der Klassenführung *Disziplin* und *Zeitnutzung*. Dagegen zeigten die drei Merkmale des Unterrichtsklimas durchgängig hohe Interkorrelationen. Die Merkmale der kognitiv-motorischen Aktivierung fielen mittel bis hoch aus, wobei die hohen Korrelationen zwischen *Verarbeitung* und *Fokussierung* ($r = .86$, $p < .01$), *Rückmeldung* und *Strukturierung* ($r = .80$, $p < .01$) sowie *Strukturierung* und *Verständlichkeit* ($r = .73$, $p < .01$) auffallend hoch waren.

Tabelle 2

*Faktorwerte (FW), Faktorreliabilitäten (FR), interne Konsistenzen (C's α) und deskriptive Werte (M, SD) und Intraklassenkorrelationen (ICC) des Schüler*innenfragebogens; interne Konsistenzen (C's α) und deskriptive Werte (M, SD) des Lehrpersonenfragebogens und Perspektivvergleich*

Merkmale	Schüler*inneneinschätzung					Lehrpersoneneinschätzung					Vergleich	
	FW	FR	C's α	M	SD	ICC (1)	ICC (2)	C's α	M	SD	β	p
Disziplin ¹	.78 – .87	.87	.86	2.99	1.15	.46	.93	.90	3.68	1.02	.45	<.01
Zeitnutzung ¹	.67 – .90	.76	.83	3.58	1.11	.28	.86	.86	4.33	0.62	.30	<.01
Zielklarheit	.66 – .87	.81	.81	3.90	0.93	.12	.68	.71	4.34	0.63	.07	.33
Feedback	.80 – .87	.79	.87	4.11	0.89	.26	.85	.76	4.59	0.46	.02	.80
Fürsorglichkeit	.70 – .77	.77	.78	3.95	0.89	.26	.85	.26	4.34	0.42	-.05	.58
Unterstützung	.71 – .83	.80	.82	3.91	0.87	.27	.85	.62	4.64	0.37	.02	.86
Fokussierung	.56 – .76	.68	.72	3.45	0.91	.09	.61	.70	3.45	0.75	.02	.80
Verarbeitung	.66 – .83	.74	.77	3.60	0.95	.05	.45	.65	3.87	0.78	.04	.52
Verständlichkeit	.74 – .76	.81	.80	4.08	0.82	.21	.81	.90	4.80	0.36	-.08	.37
Anspruchsniveau	.59 – .80	.74	.75	3.78	0.91	.15	.74	.69	4.29	0.47	.03	.64
Strukturierung	.73 – .79	.79	.80	3.84	0.91	.16	.75	.85	4.43	0.60	.00	.99
Rückmeldung	.74 – .86	.85	.84	4.12	0.81	.16	.75	.70	4.63	0.38	-.04	.59

Kodierung: (1) trifft nicht zu, (2) trifft eher nicht zu, (3) trifft teils zu, (4) trifft eher zu, (5) trifft zu

¹Die invers kodierten Items wurden für die Berechnung der deskriptiven Werte rekodiert.

Tabelle 3

*Unterhalb der Diagonalen: Latente Korrelationen der Unterrichtsmerkmale aus Schüler*innensicht (N = 634) mit Kriteriumsvariablen; Oberhalb der Diagonalen: Pearson Korrelationen der Unterrichtsmerkmale aus Lehrpersonensicht (N = 40)*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Disziplin		.49*	.17	.31	.20	.29	-.12	.15	.19	.03	.36*	.10
2. Zeitznutzung	.51*		.35*	.05	.35*	.30	.37*	.44*	.27	.43**	.57**	.42**
3. Zielklarheit	.04	.12*		.09	.57*	.36*	.44**	.62**	.31	.27	.68**	.11
4. Feedback	.13	.29*	.34*		.43*	.50*	-.06	-.05	.05	.45**	.10	.21
5. Fürsorglichkeit	.12	.41*	.39*	.66*		.59**	.24	.38*	.36*	.59**	.62**	.39*
6. Unterstützung	.01	.32*	.36*	.77*	.79*		.17	.25	.30	.53**	.57**	.46**
7. Fokussierung	.09	.20*	.40*	.36*	.35*	.40*		.51*	-.12	.38*	.35*	.22
8. Verarbeitung	.13*	.26*	.38*	.42*	.38*	.42*	.86*		.35*	.37*	.75**	.12
9. Verständlichkeit	.10	.44*	.42*	.62*	.75*	.73*	.46*	.50*		.13	.35*	.23
10. Anspruchsniveau	.03	.31*	.31*	.49*	.48*	.49*	.38*	.41*	.46*		.41**	.44**
11. Strukturierung	.05	.28*	.65*	.64*	.68*	.75*	.49*	.54*	.73*	.49*		.37*
12. Rückmeldung	.05	.32*	.50*	.55*	.65*	.70*	.39*	.45*	.66*	.45*	.80*	
13. Anstrengung	.13*	.26*	.28*	.41*	.48*	.49*	.57*	.57*	.59*	.41*	.49*	.39*
14. Interesse	.06	.24*	.25*	.39*	.48*	.47*	.44*	.44*	.61*	.26*	.49*	.39*
15. Geschlecht ¹	.20*	.02	-.04	-.11	-.18*	-.11	-.11	-.10	-.22*	-.17*	-.13	-.10
16. Alter	.00	.06	.07	.05	.01	.10	.07	-.01	.07	.17*	.10	.15*

* p < .05; ** p < .01; ¹ Jungen = 0, Mädchen = 1

Während die Mehrheit der Merkmale mit den Kriteriumsvariablen *Anstrengungsbereitschaft im Sportunterricht* und *Interesse am Sportunterricht* in einem mittleren bis hohen Zusammenhang standen, wiesen die drei Merkmale der Klassenführung *Disziplin*, *Zeitnutzung*, *Zielklarheit* vergleichsweise geringe Korrelationen auf. Die höchsten Zusammenhänge mit den Kriteriumsvariablen erreichten die drei Merkmale *Fokussierung*, *Verarbeitung* und *Verständlichkeit* (Tabelle 3). An dieser Stelle ist darauf hinzuweisen, dass aufgrund des querschnittlichen Designs keine kausalen Wirkungen der Unterrichtsmerkmale auf die Motivation der Schüler*innen gefolgert werden können. Vielmehr könnte auch angenommen werden, dass bspw. interessierte Schüler*innen die Unterrichtsmerkmale anders einschätzten als nicht-interessierte Schüler*innen.

Die Korrelationen zwischen den Merkmalen und den Kovariaten *Geschlecht* und *Alter* fielen gering und überwiegend nicht signifikant aus. Die *Disziplin* wurde von Mädchen positiver eingeschätzt als von Jungen, während die *Fürsorglichkeit* und *Verständlichkeit* von den Mädchen negativer bewertet wurde. Das *Alter* wies lediglich signifikante positive Korrelationen in geringer Höhe mit den Merkmalen *Anspruchsniveau* und *Rückmeldung* auf. Ältere Schüler*innen nehmen ein höheres Anspruchsniveau und positivere Rückmeldungen wahr als jüngere Schüler*innen. Diese geringen Zusammenhänge können als erster Hinweis darauf dienen, dass die Erfassung der Unterrichtsmerkmale über Schüler*inneneinschätzungen weitestgehend geschlechts- und altersübergreifend erfolgen kann.

4.3 Reliabilität auf Klassen- und Lehrpersonenebene

Die Unterrichtsqualität in einer Klasse gilt als ein Mass für eine geteilte schulische Umwelt, welche in Beziehung zu anderen Konstrukten (z. B. Lernleistung, Schulform) gesetzt werden kann. Hierfür wurden die individuellen Einschätzungen der Schüler*innen einer Klasse gemittelt (Lüdtke et al., 2006). Die Unterschiede in den Schüler*inneneinschätzungen zwischen den Schulklassen lassen sich durch den ICC (1) charakterisieren. Tabelle 2 zeigt, mit Ausnahme der Merkmale *Fokussierung* und *Verarbeitung*, hohe ICC (1)-Werte, welche aufzeigten, dass die Klassenzugehörigkeit einen hohen Anteil in der Varianz der Merkmale aufklären kann. Beispielsweise erklärte die Klassenzugehörigkeit 46 % der Gesamtvarianz im Merkmal *Disziplin*. Die Bewertungen der Merkmale *Fokussierung* und *Verarbeitung* waren dagegen eher auf Unterschiede zwischen den einzelnen Schüler*innen zurückzuführen und weniger von der Klasse abhängig. Entsprechend fielen die Werte zur ICC (2), verstanden als Reliabilität des Klassenmittelwertes der Schüler*inneneinschätzungen, ebenfalls hoch aus (Tabelle 2). Ausnahme bilden wiederum die Merkmale *Fokussierung* und *Verarbeitung*, deren Reliabilität mit $.45 \leq \text{ICC} (2) \leq .61$ niedrig ausfällt und somit für eine Aggregation auf Klassenebene nur bedingt geeignet sind.

Die Einschätzungen der $N = 40$ Lehrpersonen lagen auf Klassenebene vor. Hier wiesen die Cronbach's α -Werte, mit Ausnahme des Merkmals *Fürsorglichkeit* ($\alpha = .26$), auf eine ausreichende interne Konsistenz hin (Tabelle 2). Die Interkorrelationen der durch die Lehrpersonen eingeschätzten Merkmale waren gering bis mittel (Tabelle 3, oberhalb der Diagonalen) und fielen damit insgesamt etwas niedriger aus als die Schüler*inneneinschätzungen (Tabelle 3, oberhalb der Diagonalen).

4.4 Zusammenhänge der Schüler*innen- und Lehrpersoneneinschätzungen

Um die Einschätzungen der Lehrpersonen (auf Klassenebene) mit den Einschätzungen der Schüler*innen (auf Individualebene) zu vergleichen, wurde mittels Mehrebenenanalysen geprüft, inwiefern die Lehrpersoneneinschätzungen die Einschätzungen der Schüler*innen präzisieren können. Hier zeigte sich mit Ausnahme der Merkmale *Disziplin* ($\beta = .45$) und *Zeitnutzung* ($\beta = .30$) keine signifikanten Zusammenhänge (Tabelle 2). Entsprechend konnten die Lehrpersoneneinschätzungen nur sehr eingeschränkt die Schüler*inneneinschätzungen erklären.

5. Diskussion und Ausblick

Nachdem in den letzten Jahrzehnten vor allem der adäquate Einsatz verschiedener Unterrichtsmethoden und sportdidaktischer Konzepte im Sportunterricht diskutiert wurde (im Überblick Balz, 2009), wird im Fach Sport der Fokus zunehmend auf die Unterrichtsqualität gerichtet. Neben einer normativ-theoretischen Fundierung dieses zentralen Themas wird gefordert, die Qualität des Sportunterrichts auch empirisch zu erfassen, um den Unterricht systematisch weiterentwickeln zu können. Hierbei stellt sich das Problem, dass der Frage „Was ist guter Sportunterricht?“ zunächst die Frage „Welche Erziehungs- und Lernziele sollen im Sportunterricht überhaupt erreicht werden?“ vorausgehen muss.

Entsprechend bedarf es einer klaren Bestimmung der Erziehungs- und Lernziele, der Lerngegenstände und der damit in Verbindung stehenden Lehr- und Lernprozesse. Mit dem Konzept der *Handlungsfähigkeit im Schulsport* (Kurz, 2013) als (normative) Leitidee einer pragmatischen Fachdidaktik liegt zwar innerhalb der Sportpädagogik weitestgehend ein Konsens hinsichtlich der Erziehungs- und Lernziele vor. Jedoch sind diese im Rahmen der Handlungsfähigkeit derart vielfältig, dass die theoretische Fundierung der Lehr- und Lernprozesse zukünftig ausdifferenziert werden muss.

Im vorliegenden Beitrag wurde unter einer handlungstheoretischen Perspektive (Niederkofler & Amesberger, 2016) die Unterrichtsdimension *kognitiv-motorische Aktivierung* für das Fach Sport näher bestimmt und in ausgewählte Merkmale ausdifferenziert. Bei den Unterrichtsdimensionen *Klassenführung* und *Unterrichtsklima* konnte auf theoretische und empirische Vorarbeiten zurückgegriffen werden, wobei auch deutlich wurde, dass diese Dimensionen einen stärker fachübergreifenden Charakter besitzen (im Überblick Herrmann & Gerlach, 2020).

Mit dem QUALLIS-Instrument wurde eine Operationalisierung der Basisdimensionen und den zugehörigen Merkmalen guten Sportunterrichts realisiert. Ziel war es, ein ökonomisch einsetzbares Instrument zur Evaluation der Qualität des Lehrens und Lernens im Sport zu entwickeln. Im Rahmen dieser Studie konnte die faktorielle Validität des QUALLIS-Instruments nachgewiesen werden. Die Unterrichtswahrnehmung von Schüler*innen wurde anhand von 36 Items erfasst, welche sich den zwölf theoretisch begründeten Merkmalen zuordnen liessen. Die Merkmale waren gering mit dem Alter und Geschlecht der Schüler*innen korreliert, so dass eine Verwendung des QUALLIS-Instruments in der Sekundarstufe I unabhängig vom Geschlecht und Jahrgangsstufe möglich ist. Auf Basis dieser querschnittlichen Ergebnisse zeigten sich deutliche Zusammenhänge zwischen den Merkmalen und den motivationalen Aspekten *Interesse* und *Anstrengungsbereitschaft*. Diese wiesen auf die Bedeutsamkeit der Unterrichtsqualität für motivationale Aspekte im Fach Sport hin, welche in vergleichbarer Weise für das Fach Sport (z.B. Heemsoth & Krieger, 2018) sowie für Schulfächer in sprachlichen und mathematisch-naturwissenschaftlichen Fächern nachgewiesen wurden (Cornelius-White, 2016; Kunter & Baumert, 2006). Aufgrund des querschnittlichen Designs der Studie können jedoch nur Zusammenhänge mit den erfassten Unterrichtsmerkmalen des QUALLIS-Instruments angenommen werden. Zwar ist davon auszugehen, dass eine hohe Unterrichtsqualität die Motivation der Schüler*innen positiv beeinflusst, jedoch konnte auch die entgegengesetzte Wirkrichtung nachgewiesen werden. Schüler*innen mit hohem Fachinteresse nehmen die Unterrichtsqualität positiver wahr (Anderman, 2003).

Erste längsschnittlich angelegte Studien im Fach Sport geben Hinweise darauf, dass die Schüler*inneneinschätzungen der *Klassenführung* (Seiler, 2019) und des *Unterrichtsklimas* (Niederkofler et al., 2015) einen positiven Einfluss auf motivationale Aspekte besitzen. Bei den vorliegenden Studien aus dem Fach Sport ist zu berücksichtigen, dass Unterrichtsqualität und motivationale Variablen anhand der Schüler*inneneinschätzungen erhoben wurden. Damit wäre es auch denkbar, dass die dargestellten Zusammenhänge (anteilig) auf Methodeneffekte zurückzuführen sind (Schermelleh-Engel & Schweizer, 2012). Entsprechend müssen sich die dargestellten Zusammenhänge dieser Validierungsstudie in längsschnittlichen Studien noch erhärten. Für die *kognitiv-motorische Aktivierung* stellt sich darüber hinaus die Frage, ob sich der postulierte Zusammenhang zwischen den Merkmalen und den (Bewegungs-)Handlungen von Schüler*innen auch empirisch bestätigt werden kann. Inwieweit auch die Merkmale der *Klassenführung* und des *Unterrichtsklimas* einen Einfluss auf konkrete fachliche Lernleistungen besitzen, muss ebenfalls in zukünftigen Studien geklärt werden.

Mit ausreichend grossen Stichprobengrössen würde dann auch die Frage empirisch bearbeitet werden können, ob sich die faktorielle Validität des QUALLIS-Instruments auch bei den Lehrpersoneneinschätzungen zeigen lässt. Die vorliegenden manifesten Berechnungen zur internen Konsistenz waren bei einzelnen Merkmalen (z.B. Fürsorglichkeit: $\alpha = .26$) nicht zufriedenstellend, so dass hier eine erneute Prüfung vorgenommen werden sollte. Weiterhin stellt sich die Frage, ob sich auf Ebene der Lehrpersonen die gleiche faktorielle Struktur zeigt, wie auf Ebene der Schüler*innen (Heemsoth & Krieger, 2018).

Im Abgleich der Einschätzungen der Schüler*innen mit denen der Lehrpersonen deutete sich in der vorliegenden Studie an, dass es sich um zwei eigenständige Perspektiven auf Sportunterricht handelt und diese kaum im Zusammenhang stehen. Zur weiterführenden Einschätzung der Kriteriumsvalidität wären ergänzende, systematische Fremdbeobachtungen des Sportunterrichts mittels validierter Beobachtungsinstrumente hilfreich und gewinnbringend (z.B. CLASS, Kohake et al., 2023). Diese sind jedoch nur mit vergleichsweise hohem personellen und zeitlichem Aufwand umzusetzen und es stellen sich Fragen der Äquivalenz der erfassten Konstrukte. Wird also mit der CLASS-Domain *Instructional Support* das gleiche Konstrukt erfasst, wie mit der QUALLIS-Dimension *kognitiv-motorische Aktivierung*?

Das QUALLIS-Instrument liegt neben der Version für Lehrpersonen und Schüler*innen auch bereits in einer Beobachter*innenversion vor (Herrmann, 2019), welche eine kriteriengeleitete Erfassung des Unterrichtsgeschehens aus einer dritten Perspektive zulässt. Ob diese niederschwellige Möglichkeit der Fremdbeobachtung valide und reliable Ergebnisse liefert und ob eine umfassende Schulung der Beobachter*innen zwingend nötig ist, muss sich in zukünftigen Studien zeigen.

Vor dem Hintergrund des aktuellen Diskurses in der Sportwissenschaft, was unter der Basisdimension *Aktivierung* im Fach Sport zu verstehen ist (Wibowo et al., 2021; Künzell et al., 2023), muss auch die Auswahl und Operationalisierung der QUALLIS-Merkmale hinterfragt werden. Daher sollten die im QUALLIS-Instrument verwendeten Evaluationsmerkmale zur Erfassung der *kognitiv-motorischen Aktivierung* als eine (erste) Auswahl begriffen werden, welche lediglich Teilaspekte des fachspezifischen Lehr-Lernprozesses abbilden.

Die teils hohen Interkorrelationen der Merkmale (z.B. Verarbeitung mit Fokussierung, Rückmeldung mit Strukturierung) geben bereits jetzt Hinweise darauf, dass es sich bei der Unterrichtsdimension der *kognitiv-motorischen Aktivierung* um ein mehrdimensionales Konstrukt handeln könnte. Demnach – und um die normativen Überlegungen zur qualifikatorischen und reflexiven Handlungsfähigkeit (Gogoll, 2013) stärker zu berücksichtigen – sollte zukünftig überlegt werden, die Aktivierung theoretisch-konzeptionell sowie empirisch in eine *kognitiv-motorische Aktivierung* und eine *kognitiv-reflexive Aktivierung* zu unterscheiden (Herrmann, 2023). Dies könnte für eine differenzierte Betrachtung spezifischer Lernziele und Lernprozesse gewinnbringend sein (Begrich et al., 2023, S. 88). Beispielsweise bemüht sich jüngst auch die Sportmotorik um eine Konzeption von *kognitiver Aktivierung* vor dem Hintergrund einer „Embodied Cognition“. Hierbei wird Bewegungslernen als kognitiver Prozess verstanden und damit das Erlernen motorischer Kontrolle auch als (verkörperte) kognitive Aktivität begriffen (Künzell et al., 2023).

Damit rückt im Fach Sport – wie auch in anderen Schulfächern – eine fachdidaktische Unterrichtsqualitätsforschung in den Fokus, die ausgehend von den (1) *fachlichen Zugängen* (z. B. problemlösen, bewegen, experimentieren, argumentieren, beurteilen), (2) *Lernziele, Inhalte und Methoden* ableitet (z. B. Balancieren und Springen, Kooperieren im Spiel, Regelverständnis) und zugehörige (3) *Lehr- und Lernprozesse* beschreibt (z.B. motorisches, soziales und verständnisorientiertes Lernen). Aus diesen spezifischen Lehr- und Lernprozessen können notwendige Anpassungen der bestehenden Unterrichtsqualitätsmodelle gefolgert und in konkrete beobachtbare Indikatoren guten Fachunterrichts operationalisiert werden (Begrich et al., 2023).

Als mögliches Anwendungsfeld des QUALLIS-Instruments kann die systematische und mehrperspektivische Unterrichtsevaluation begriffen werden, welche durch eine datenbasierte Einschätzung des Unterrichtsgeschehens eine gezielte Massnahmenplanung ermöglicht (Balzer & Beywl, 2015; Helmke & Lenske, 2013). Es werden Vergleiche der Einschätzungen der Schüler*innen, der unterrichtenden Lehrperson sowie eines Beobachters bzw. einer Beobachterin ermöglicht (Herrmann, 2019). Solche Vergleiche können die Voraussetzung für die Entwicklung einer professionellen Feedbackkultur und darauf aufbauender Unterrichtsentwicklung bilden (Rahn et al., 2016).

Das QUALLIS-Instrument verspricht, neben einen ersten theoretisch-konzeptionellen sowie empirischen Zugang zur Unterrichtsqualität im Fach Sport, einem Mehrwert für die Aus- und Fortbildung von Lehrpersonen im Fach Sport. Der Innovationsgehalt des QUALLIS-Instruments besteht in den (1) sportspezifischen Merkmalen und deren Operationalisierung, (2) der für die Schüler*innen gut verständlichen Itemformulierungen und (3) der Möglichkeit unterschiedliche Perspektiven erfassen und vergleichen zu können.

Literaturverzeichnis

- Anderman, L. H. (2003). Academic and social perceptions as predictors of change in middle school students' sense of school belonging. *The Journal of Experimental Education*, 72(1), 5–22. <https://doi.org/10.1080/00220970309600877>
- Balz, E. (2009). Fachdidaktische Konzepte update oder: Woran soll sich der Schulsport orientieren? *Sportpädagogik*, 33(1), 25–32.
- Balzer, L., & Beywl, W. (2015). *evaluiert: Planungsbuch für Evaluationen im Bildungsbereich*. hep verlag.
- Baumgartner, M., Oesterhelt, V., & Reuker, S. (2020). Konstruktion und Validierung eines multidimensionalen Beobachtungsinstrumentes zur Erfassung der klassenführungsbezogenen Performanzen von sportunterrichtenden Lehrkräften (KlaPe-Sport). *German Journal of Exercise and Sport Research*, 50(4), 511–522. <https://doi.org/10.1007/s12662-020-00675-6>
- Begrich, L., Praetorius, A.K., Decristan, J. et al. (2023). Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft. *Unterrichtswissenschaft*, 51, 63–97 (2023). <https://doi.org/10.1007/s42010-023-00163-4>
- Borich, G. D. (2007). *Effective teaching methods* (6. ed). Pearson Merrill.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). Pearson Studium.
- Charalambous, C. Y., & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*, 67, 100894. <https://doi.org/10.1016/j.stueduc.2020.100894>
- Cornelius-White, J. (2016). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, 77(1), 113–143. <https://doi.org/10.3102/003465430298563>

- Decristan, J., Hess, M., Holzberger, D., & Praetorius, A.-K. (2020). Oberflächen- und Tiefenmerkmale. *Zeitschrift für Pädagogik, Beiheft* (66), 102–116. <https://doi.org/10.3262/ZPB2001102>
- Deutscher Sportbund (Hrsg.). (2006). *DSB-Sprint-Studie: Eine Untersuchung zur Situation des Schulsports in Deutschland*. Meyer & Meyer.
- E-EDK. (2017). *Lehrplan 21*. <https://zh.lehrplan.ch/index.php?code=e|200|3>
- Geiser, C. (2011). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung*. Springer VS.
- Gogoll, A. (2013). Sport- und bewegungskulturelle Kompetenz: Zur Begründung und Modellierung eines Teils handlungsbezogener Bildung im Fach Sport. *Zeitschrift für Sportpädagogische Forschung*, 1(2), 5–24.
- Gräsel, C., & Göbel, K. (2011). Unterrichtsqualität. In H. Reinders, H. Dittton, C. Gräsel, & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung: Gegenstandsbereiche* (1. Aufl., S. 87–98). VS.
- Haag, L., & Streber, D. (2012). *Klassenführung: Erfolgreich unterrichten mit Classroom Management*. Beltz.
- Hapke, J., & Waigel, S. (2019). „Sporttreiben mit Köpfchen“ – Kognitive Aktivierung im Sportunterricht. In A. Gawatz & K. Stürmer (Hrsg.), *Kognitive Aktivierung im Unterricht: Befunde der Bildungsforschung und fachspezifische Zugänge* (S. 148–161). Westermann.
- Hartig, J., Frey, A., & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 143–172). Springer.
- Hattie, J. (2013). *Lernen sichtbar machen: Überarbeitete deutschsprachige Ausgabe von Visible Learning*. Schneider Hohengehren.
- Heemsoth, T. (2014). Unterrichtsklima als Mediator des Zusammenhangs von Klassenführung und Motivation im Sportunterricht. *Psychologie in Erziehung und Unterricht*, 61(3), 203–215.
- Heemsoth, T., & Krieger, C. (2018). Perspektiven auf Sportunterricht. *Unterrichtswissenschaft*, 46(4), 499–522. <https://doi.org/10.1007/s42010-018-0026-8>
- Heemsoth, T., & Miethling, W.-D. (2012). Schülerwahrnehmungen des Unterrichtsklimas: Entwicklung eines Fragebogens und Befund zum Sportunterricht. *Sportwissenschaft*, 42(4), 228–239.
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (7. Aufl.). Klett-Kallmeyer.
- Helmke, A., & Lenske, L. (2013). Unterrichtsdiagnostik als Voraussetzung für Unterrichtsentwicklung. *Beiträge zur Lehrerbildung*, 31(2), 214–233.
- Herrmann, C. (2019). Evaluation der Unterrichtsqualität im Sportunterricht mit dem QUALLIS-Instrument. *Bewegung und Sport*, 73(2), 12–17.
- Herrmann, C. (2021). Unterrichtsdimensionen und -merkmale als Orientierungspunkte eines wirksamen Sportunterrichts. In P. Neumann & E. Balz (Hrsg.), *Unterrichtspraxis: Perspektiven von Expertinnen und Experten: Bd. 17. Wirksamer Sportunterricht* (1. Aufl., S. 49–59). Schneider Hohengehren.
- Herrmann, C. (2023). *Qualität des Lehrens und Lernens im Sportunterricht (QUALLIS) - Ein lerngegenstandsspezifisches Modell der Unterrichtsqualität*. Zenodo. <https://doi.org/10.5281/zenodo.7625336>
- Herrmann, C., & Gerlach, E. (2017). Unterrichtsqualität im Schneesport. In I. Bach (Hrsg.), *Schriftenreihe der ASH: Bd. 24. Skilauf und Snowboard in Lehre und Forschung* (S. 118–129). Feldhaus.
- Herrmann, C., & Gerlach, E. (2020). Unterrichtsqualität im Fach Sport – Ein Überblicksbeitrag zum Forschungsstand in Theorie und Empirie. *Unterrichtswissenschaft*, 48, 361–384. <https://doi.org/10.1007/s42010-020-00080-w>
- Herrmann, C., Gogoll, A., & Gerlach, E. (2020). Unterrichtsqualität im Fach Sport. In E. Balz, C. Krieger, W.-D. Miethling & P. Wolters (Hrsg.), *Edition Schulsport: Bd. 20. Empirie des Schulsports* (3. Aufl., S. 61–81). Meyer & Meyer.
- Herrmann, C., Seiler, S., Gerlach, E., & Sohnmeyer, J. (2018). Unterrichtsqualität im Fach Sport - Dimensionen, Merkmale und Evaluation. In E. Balz & D. Kuhlmann (Hrsg.), *Schriften der Deutschen Vereinigung für Sportwissenschaft: Band 269. Sportwissenschaft in pädagogischem Interesse* (S. 44–46). Feldhaus, Edition Czwalina.
- Herrmann, C., Seiler, S., & Niederkofler, B. (2016). „Was ist guter Sportunterricht?“ – Dimensionen der Unterrichtsqualität. *Sportunterricht*, 65(3), 7–12.
- Herrmann, C., Seiler, S., Pühse, U., & Gerlach, E. (2015). „Wie misst man guten Sportunterricht?“ – Erfassung der Dimensionen von Unterrichtsqualität im Schulfach Sport. *Zeitschrift für Sportpädagogische Forschung*, 3(1), 6–26.
- Herrmann, C., Seiler, S., Siffert, A., Dapp, L., Gashaj, V. & Studer, F. (under revision). Entwicklung und Validierung eines Evaluationsinstruments zur Erfassung der Qualität des Lehrens und Lernens im Schweizer Kinder- und Jugendsport. *Forum Kinder- und Jugendsport*.
- Klieme, E. (Hrsg.). (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Beltz.
- Klieme, E., & Baumert, J. (Hrsg.). (2001). *BMBF publik. TIMSS - Impulse für Schule und Unterricht: Forschungsbefunde, Reforminitiativen Praxisberichte und Video-Dokumente*.
- Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik: Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237.
- Kohake, K., Richartz, A., & Maier, J. (2023). Measuring pedagogical quality in children's sports: validity and reliability of the classroom assessment scoring system K-3 in extracurricular sports training. *German Journal of Exercise and Sport Research*, 53(1), 47–58. <https://doi.org/10.1007/s12662-022-00836-9>
- Kounin, J. S. (1976). *Techniken der Klassenführung*. Huber.
- Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9(3), 231–251.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*. Springer.
- Kunter, M., & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer, & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts*. Waxmann.

- Künzell, S., Maurer, L., Voigt, L., & Zentgraf, K. (2023). Kognitive Aktivität im Sportunterricht. *Sportunterricht*, 72(3), 98–103. <https://doi.org/10.30426/SU-2023-03-1>
- Kurz, D. (2013). Zur Entwicklung einer pragmatischen Fachdidaktik. In P. Neumann & E. Balz (Hrsg.), *Sportdidaktik: Pragmatische Fachdidaktik für die Sekundarstufe I und II* (S. 13–23). Cornelsen.
- Kurz, D., & Gogoll, A. (2010). Standards und Kompetenzen. In N. Fessler, A. Hummel, & G. Stibbe (Hrsg.), *Handbuch Schulsport* (S. 227–244). Hofmann.
- Lamprecht, M., Fischer, A., & Stamm, H. (2014). *Sport Schweiz 2014: Sportaktivität und Sportinteresse der Schweizer Bevölkerung*. Bundesamt für Sport BASPO.
- Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Analyse von Lernumwelten: Ansätze zur Bestimmung der Reliabilität und Übereinstimmung von Schülerwahrnehmungen. *Zeitschrift für pädagogische Psychologie*, 20(1/2), 85–96.
- Meyer, H. (2005). *Was ist guter Unterricht?* (3., Aufl.). Scriptor.
- Munzert, J. (1995). Bewegung als Handlung verstehen. In R. Prohl & J. Seewald (Hrsg.), *Bewegung verstehen: Facetten und Perspektiven einer qualitativen Bewegungslehre* (S. 77–97). Verlag Karl Hofmann.
- Niederkofler, B., & Amesberger, G. (2016). Kognitive Handlungsrepräsentationen als Strukturgrundlage zur Definition von kognitiver Aktivierung im Sportunterricht. *Sportwissenschaft*, 46(3), 188–200. <https://doi.org/10.1007/s12662-016-0414-3>
- Niederkofler, B., & Amesberger, G. (2023). Förderung von motorischen Basiskompetenzen über das Können, Wissen und Wollen der Schulkinder: Effekte einer 8-wöchigen Unterrichtsreihe in der zweiten und vierten Schulstufe. *Unterrichtswissenschaft*. Vorab-Onlinepublikation. <https://doi.org/10.1007/s42010-023-00176-z>
- Niederkofler, B., & Herrmann, C. (2022). Aufmerksamkeit als Modalität der kognitiven Aktivität der Schüler*innen im Sportunterricht. In R. Sygusch, J. Hapke, S. Liebl, & C. Töpfer (Hrsg.), *Beiträge zur Lehre und Forschung im Sport: Bd. 198. Kompetenzorientierung im Sport: Grundlagen, Modellentwurf und Anwendungsbeispiele* (S. 101–116). Hofmann.
- Niederkofler, B., Herrmann, C., Seiler, S., & Gerlach, E. (2015). What influences motivation in physical education? A multi-level approach to identify climate determinants of achievement motivation. *Psychological Test and Assessment Modeling*, 57(1), 70–93.
- Nitsch, J. R. (2004). Die handlungstheoretische Perspektive: Ein Rahmenkonzept für die sportpsychologische Forschung und Intervention. *Zeitschrift für Sportpsychologie*, 11(1), 10–23. <https://doi.org/10.1026/1612-5010.11.1.10>
- Praetorius, A.-K. (2014). *Messung von Unterrichtsqualität durch Ratings*. Waxmann.
- Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM*, 50(3), 535–553. <https://doi.org/10.1007/s11858-018-0946-0>
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B., & Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*, 47(1), 97. <https://doi.org/10.1007/s11858-018-0918-4>
- Raab, M., Masters, R. S. W., Maxwell, J., Arnold, A., Schlapkohl, N., & Poolton, J. (2009). Discovery learning in sports: Implicit or explicit processes? *International Journal of Sport and Exercise Psychology*, 7(4), 413–430. <https://doi.org/10.1080/1612197X.2009.9671917>
- Rahn, S., Gruehn, S., Keune, M., & Fuhrmann, C. (2016). Aus Schüleraussagen lernen?! - Auf dem Weg zu einer professionellen Feedbackkultur an Schulen. *Die Deutsche Schule*, 108(2), 163–175.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publications.
- Richartz, A., & Kohake, K. (2021). Zur (Fach-)Spezifität von Unterrichtsqualität im Fach Sport. *Unterrichtswissenschaft*, 49(2), 243–251. <https://doi.org/10.1007/s42010-021-00112-z>
- Schermelleh-Engel, K., & Schweizer, K. (2012). Multitrait-Multimethod-Analysen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 345–362). Springer.
- Seiler, S. (2019). *Lernleistungen im Sport. Bildung und Sport*. Springer VS.
- Seiz, J., Decristan, J., Kunter, M., & Baumert, J. (2016). Differenzielle Effekte von Klassenführung und Unterstützung für Schülerinnen und Schüler mit Migrationshintergrund. *Zeitschrift für Pädagogische Psychologie*, 30(4), 237–249. <https://doi.org/10.1024/1010-0652/a000186>
- Weiber, R., & Mühlhaus, D. (2014). Güteprüfung reflektiver Messmodelle. In R. Weiber & D. Mühlhaus (Hrsg.), *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS* (2. Aufl., S. 127–172). Springer Gabler.
- Wibowo, J., Krieger, C., Gerlach, E., & Buekers, F. (Hrsg.). (2021). *Aktivierung im Sportunterricht* (2. überarbeitete und erweiterte Auflage). Universität Hamburg. <https://doi.org/10.25592/AktivierungImSU>

Schlagworte: Sportunterricht; Fragebogenentwicklung; Klassenführung; Unterrichtsklima; Aktivierung

Évaluation de la qualité de l'enseignement et de l'apprentissage dans l'éducation physique – Validation de l'instrument QUALLIS

Résumé

La description de la qualité de l'enseignement est un domaine important dans la recherche en éducation. Des questions se posent sur les critères de la qualité de l'éducation physique et sur la manière dont la qualité de l'enseignement et de l'apprentissage peut être évaluée. L'article élabore douze critères d'une bonne éducation physique et les opérationnalise d'une manière multi-perspective. Une étude de validation portant sur 40 classes de l'école secondaire (634 élèves et leurs enseignant·e·s) a confirmé au moyen d'analyses factorielles confirmatoires la structure à douze facteurs attendue ainsi que les corrélations avec la motivation des élèves. L'instrument QUALLIS fournit un accès empirique à la qualité de l'enseignement du sport et offre ainsi une valeur ajoutée à la formation des enseignant·e·s.

Mots-clés: Éducation physique; élaboration d'un questionnaire; gestion de la classe; climat de la classe; activation

Valutazione della qualità dell'insegnamento e dell'apprendimento nell'educazione fisica – Validazione dello strumento QUALLIS

Riassunto

Un importante campo della ricerca educativa è la descrizione della qualità dell'insegnamento. Questo solleva delle domande sulle caratteristiche della qualità nell'ambito dell'educazione fisica e su come la qualità dell'insegnamento e dell'apprendimento nell'educazione fisica (QUALLIS) possa essere valutata sistematicamente. L'articolo elabora 12 caratteristiche di una buona educazione fisica operazionalizzandole in modo multiprospettico. Uno studio di validazione su 40 classi del livello secondario (634 studenti e i loro insegnanti) ha confermato la struttura a dodici fattori prevista e le correlazioni con la motivazione degli studenti attraverso analisi fattoriali di conferma. Lo strumento QUALLIS fornisce un accesso empirico alla qualità dell'insegnamento in materia di educazione fisica e offre quindi un valore aggiunto per la formazione delle e degli insegnanti.

Parole chiave: Educazione fisica; sviluppo del questionario; gestione della classe; clima in classe; attivazione

Assessment of teaching and learning quality in physical education – Validation of the QUALLIS instrument

Summary

An important subfield of educational research is the description of teaching quality. This raises questions about quality characteristics in the subject of physical education and how the quality of teaching and learning in physical education (QUALLIS) can be systematically evaluated. The article elaborates twelve characteristics of good physical education and operationalizes the characteristics in a multi-perspective way. A validation study with 40 secondary school classes (634 students and their teachers) confirmed the expected twelve-factorial structure and the correlations with student motivation via confirmatory factor analyses. The QUALLIS instrument provides an empirical approach to teaching quality in physical education and thus offers added value for teacher education and training.

Keywords: Physical education; questionnaire development; classroom management; classroom climate; activation

Christian Herrmann, Prof. Dr., ist Professor für Didaktik, Bewegung und Sport an der PH Zürich. Er hat an den Universitäten Bayreuth, Mainz und Jena promoviert. Während seiner Postdoc-Phase arbeitete er an der Universität Basel. Seine Habilitation zum Thema “Motorische Basiskompetenzen” schloss er an der Universität Potsdam ab. Seine aktuellen Forschungsschwerpunkte liegen in der Konzeption und Evaluation von motorischen Basiskompetenzen und der Unterrichtsqualität in den Fachdidaktiken.

Pädagogische Hochschule Zürich, Forschungsgruppe Didaktik Bewegung und Sport, Lagerstrasse 2, CH-8090 Zürich

E-Mail: christian.herrmann@phzh.ch

Benjamin Niederkofler, Prof. Dr., ist Professor für Sportpädagogik und -didaktik an der Fakultät für Bildungswissenschaften der Freien Universität Bozen. Studium der Sportwissenschaften an den Universitäten Innsbruck und Basel. Lehrdiplom für Maturitätsschulen an der PH FHNW. Promotion an der Universität Salzburg. Post-Doc-Forschung an der PH Salzburg im Bereich der Sportdidaktik zu Kognitionen sowie zur Kompetenzentwicklung von Schüler*innen im Sportunterricht.

Freie Universität Bozen-Bolzano, Fakultät für Bildungswissenschaften, Regensburger Allee 16, I-39042 Brixen

E-Mail: benjamin.niederkofler@unibz.it

Harald Seelig, Dr., ist wissenschaftlicher Angestellter am Departement für Sport, Bewegung und Gesundheit der Universität Basel. Seine Forschungsgebiete sind: Methodologie, Evaluation, Qualitätsmanagement, Monitoring motorischer Kompetenzen, Motivations- und Volitionstheorien.

Universität Basel, Departement für Sport, Bewegung und Gesundheit, Grosse Allee 6, CH-4052 Basel

E-Mail: harald.seelig@unibas.ch