

INDIVIDUELLES INTERESSE AM SPORT – VALIDIERUNG DES FRAGEBOGENS ISPO

von Jan Sohnsmeyer und Rüdiger Heim

ZUSAMMENFASSUNG | Im Rahmen der Validierung des Fragebogens *Interesse am Sport (iSpo)* wurden dessen faktorielle, konvergente, prädiktive und Kriteriumsvalidität sowie die Messinvarianz an einer Stichprobe von $N = 643$ geprüft. Die Ergebnisse bestätigten die bifaktorielle Struktur des Fragebogens: Er umfasst drei Subskalen (Rezeption, Aktivität und Vermittlung) mit jeweils sechs Items, die die Varianz zu gleichen Teilen erklären wie der G-Faktor. Zudem konnten die Reliabilität der Skalen (Rezeption: $\alpha = .89$; Aktivität: $\alpha = .92$; Vermittlung: $\alpha = .94$; G-Faktor: $\alpha = .91$) und deren Messinvarianz über das Geschlecht und die Vereinsmitgliedschaft ebenso bestätigt werden wie die Annahmen zur konvergenz und Kriteriumsvalidität. Mit dem iSpo kann daher das individuelle Interesse am Sport ökonomisch und psychometrisch überzeugend erfasst werden.

Schlüsselwörter: Interesse, Fragebogen, Validierung, Jugendforschung

INDIVIDUAL INTEREST IN SPORT AND EXERCISE – VALIDATION OF THE ISPO-QUESTIONNAIRE

ABSTRACT | The iSpo-Questionnaire, an instrument measuring individual interest in sport and exercise, was validated using data of $N = 643$ subjects. The study examined the factorial, convergent, predictive, and criterion validity. Our research confirmed the earlier findings showing that iSpo measures three dimensions with six items each (reception, activity and instruction) and a g-factor (bifactor structure). The g-factor explains half of the common variance extracted with the other half spread across the group factors. Results indicate good reliability (reception: $\alpha = .89$; activity: $\alpha = .92$; instruction: $\alpha = .94$; g-factor: $\alpha = .91$). The construct was invariant across sex and sports club membership and showed good convergent and criterion validity. iSpo can be regarded as an economic instrument with good psychometric properties.

Key Words: Interest, Questionnaire, Validation, Youth Research

INDIVIDUELLES INTERESSE AM SPORT – VALIDIERUNG DES FRAGEBOGENS ISPO

1 | EINLEITUNG

Aus unserer Sicht dürfte das Interesse am Sport als theoretisch fundiertes Konstrukt sowohl in der Jugend- als auch in der empirischen Schulsportforschung für viele Fragestellungen hohe Relevanz besitzen. Denn weil das Konstrukt des individuellen Interesses, verstanden als relativ stabile dispositionelle oder habitualisierte motivationale Struktur eines Individuums, vor allem emotionale Aspekte betont, ergänzt es kognitiv akzentuierte Theoriezugänge der Lern- und Leistungsmotivation (z. B. Erwartungs-Wert-Modelle). Während für das situationale Interesse, das als spezifischer motivationaler Zustand in einem aktuellen Handlungsablauf verstanden wird (Krapp, 2010), einige englischsprachige Inventare im Zusammenhang mit Sport zur Verfügung stehen (z. B. Chen, Darst & Pangrazi, 1999), liegen zum individuellen Interesse bisher keine Erhebungsinstrumente vor, die das Konstrukt mehrdimensional erfassen. Daher haben wir in dieser Zeitschrift vor Kurzem die Konstruktion eines Fragebogens (Heim & Sohnsmeier, 2016) sowie dessen erste testtheoretische Prüfung vorgestellt. Nachdem dort die konzeptuellen Überlegungen, der Forschungsstand und die dimensionale Struktur individuellen Interesses am Sport fokussiert wurden, prüft der vorliegende Beitrag die Validität des (etwas verfeinerten) Fragebogens. Im Mittelpunkt steht dabei die Prüfung der Konstruktvalidität¹ an einer weiteren Stichprobe im Hinblick auf die faktorielle und konvergente bzw. diskriminante Validität. Ergänzt werden diese Untersuchungen durch eine Überprüfung der Kriteriumsvalidität, mit der eine differenzierte Analyse der Messinvarianz für das Geschlecht und die Vereinsmitgliedschaft verbunden ist, sowie eine Kontrolle der Test-Retest-Reliabilität des Fragebogens. Darüber hinaus stellen wir einige methodische Anregungen für die zukünftige empirische Forschung in der Sportpädagogik vor, indem einerseits bifaktorielle Strukturgleichungsmodelle (Morin, Arens & Marsh, 2016) bei der Prüfung der psychometrischen Multidimensionalität zur Anwendung kommen und andererseits, ebenso auf Strukturgleichungsmodellen basierende, systematische Analysen der Messinvarianz.

2 | THEORETISCHE UND METHODOLOGISCHE ÜBERLEGUNGEN

Nach der zugrunde gelegten Person-Gegenstands-Theorie entfaltet sich Interesse in der Interaktion eines Individuums mit seiner gesellschaftlich gerahmten Umwelt, indem die Person eine herausgehobene Beziehung zu einem Gegenstands- oder Erfahrungsbereich entwickelt (Krapp, 2010). Diese spezifische Relation zwischen Individuum und Gegenstand konstituiert sich in interessenthematischen Handlungen (Prenzel, 1988), die allerdings in Gestalt unterschiedlicher Aktivitätsformen auftreten können. Ferner postuliert die Person-Gegenstands-Theorie eine mehrdi-

1 Wir orientieren uns bei der hier verwendeten Terminologie im Rahmen der Validitätsproblematik an der Systematik von Hartig, Frey und Jude (2012), die die *Konstruktvalidität* als verbindendes Konzept der verschiedenen Validitätsaspekte verstehen.

mensionale Struktur des Interesses (Prenzel, Krapp & Schiefele, 1986), die die gefühlsbezogene und wertbezogene Valenz sowie den intrinsischen Charakter umfasst.²

Vor diesem Hintergrund lässt sich das individuelle Interesse am Sport also einerseits im Hinblick auf die beiden Valenzen sowie den intrinsischen Charakter und andererseits im Hinblick auf die Handlungsmodalitäten der unmittelbaren sportlichen Aktivität, rezeptiver und anleitend-vermittelnder Aktivitäten strukturieren (Heim & Sohnsmeier, 2016). Auf der Ebene des Messzugriffs ließen sich die Dimensionen der Handlungsmodalitäten gut identifizieren, denn die Überprüfung mit explorativen Faktorenanalysen (EFA) ergab eine dreifaktorielle Lösung. Demgegenüber konnte – wie auch in der, in dieser Hinsicht theoretisch ähnlich angelegten, Studie von Schiefele, Krapp, Wild und Winteler (1993) – eine Differenzierung der Komponenten der gefühls- und wertbezogenen Valenz sowie des intrinsischen Charakters in einer explorativen Faktorenanalyse nicht bestätigt werden. Die Ergebnisse konfirmatorischer Faktorenanalysen (CFA) zeigten eine bifaktorielle Struktur mit dem G-Faktor Interesse am Sport und den gut differenzierten Subskalen Rezeption, Aktivität und Vermittlung mit jeweils sechs Items. Die Subskalen beinhalteten auf Itemebene die wesentlichen Interessendimensionen wertbezogene und gefühlsbezogene Valenz und intrinsische Qualität, deren Eigenständigkeit sich, wie bereits betont, nicht faktoranalytisch nachweisen ließ. Die internen Konsistenzen der Skalen lagen an der untersuchten Stichprobe von $N = 667$ Schülerinnen und Schüler an weiterführenden allgemeinbildenden Schule in der Metropolregion Rhein-Neckar zwischen .82 und .91 (Rezeption: $\alpha = .82$; Aktivität: $\alpha = .91$; Vermittlung: $\alpha = .89$; G-Faktor: $\alpha = .91$). Dies lässt auf eine gute Reliabilität der Skalen schließen, wenngleich die Itemkennwerte (Itemschwierigkeiten & Trennschärfen) einiger weniger Fragen im Test eine weitere Verfeinerung des Fragebogens nahelegten (Heim & Sohnsmeier, 2016).

Da sich bereits in unserer ersten Studie die Frage stellte, inwieweit beim Interesse am Sport von einem eher globalen (eindimensionalen) oder differenzierten (mehrdimensionalen) Konstrukt hinsichtlich der Handlungsmodalitäten auszugehen ist (Heim & Sohnsmeier, 2016), soll die dimensionale Struktur des Konstrukts im Rahmen eines bifaktoriellen Strukturgleichungsmodells erneut analysiert werden. Damit ist die faktorielle Validität angesprochen, die neben der konvergenten bzw. diskriminanten Validität geprüft werden soll. Und schließlich ist die Kriteriumsvalidität des iSpo-Fragebogens von Interesse, also inwieweit vom individuellen Interesse am Sport auf (Verhaltens-)Merkmale jenseits dieses Interesses geschlossen werden kann. In diesem Zusammenhang gehen wir davon aus, dass das individuelle Interesse am Sport vor allem für die Aufnahme, Aufrechterhaltung und Veränderung des (aktiven) Sportengagements Relevanz besitzt. Daher sollten sich aus der sportbezogenen (Jugend-)Forschung bekannte Muster (zuletzt Burrmann, Seyda, Heim & Konowalczyk, 2016) auch im Zusammenhang mit dem Interesse zeigen, sodass wir die Häufigkeit und Geschlechtsunterschiede sportlicher Aktivitäten sowie die Vereinsmitgliedschaft

2 Neben diesen emotionalen Dimensionen wird zuweilen auch die *kognitive Dimension* genannt, die sich auf die gegenstandsspezifische Komplexität bezieht und mit der epistemischen Tendenz verknüpft ist, das Wissen und Können im Interessengebiet zu erweitern und zu vertiefen (Krapp, 1998).

als Kriterien in den Blick nehmen. Bevor allerdings potenzielle Unterschiede im Interesse untersucht werden können, ist eine Prüfung der Messinvarianz des Erhebungsinstruments unabdingbar (Horn & McArdle, 1992), um sicherzustellen, dass die dimensionale Struktur des Interessenkonstrukts bei allen in den Unterschiedsprüfungen fokussierten Gruppen identisch ist.

Über diese Aspekte der Übereinstimmungsvalidität hinaus soll die Kriteriumsvalidität zudem im Hinblick auf die prognostische Validität untersucht werden: Da wir davon ausgegangen sind, dass individuelles Interesse am Sport durchaus Relevanz für die Schulsportforschung hat, ziehen wir im Rahmen der hier vorgestellten Querschnittstudie die Wahl von Sport als vierstündiges Fach in der Oberstufe als Kriterium für die prädiktive Validität heran. Wenngleich auch andere Gründe eine solche Wahl beeinflussen können – z. B. Aufwand-Nutzen-Kalkül, bisherige Leistungen oder das Fähigkeitsselbstkonzept (Möller & Trautwein, 2015; Schiefele & Schaffner, 2015) –, spielt das Fachinteresse nach Lage der Forschung nicht nur für Mathematik (Köller, Daniels, Schnabel & Baumert, 2000), sondern auch in anderen Fächern (Abel, 2002) eine wichtige Rolle.

Um die konvergente Validität des Fragebogens zu prüfen, sind Konstrukte (bzw. Erhebungsinstrumente) in den Blick zu nehmen, die eine konzeptuelle Nähe zum individuellen Interesse am Sport aufweisen. Daher kommen vor allem die Konzepte „Subjektive Werte“ (Steinmayr & Spinath, 2010), „Wichtigkeit von Sport“ (Dickhäuser & Schrahe, 2006), „Physisches Selbstkonzept“ (Gerlach, Wilschmann, Kehne, Oesterreich & Stucke, 2005) und „Freude am Sport“ (Kendzierski & DeCarlo, 1991) in Betracht. Sie sollen in gebotener Kürze vorgestellt werden.

Vor dem Hintergrund des Erwartungs-Wert-Modells von Eccles und Kollegen (Eccles et al., 1983) haben Steinmayr und Spinath (2010) subjektive schulische Werte im Hinblick auf die drei Komponenten *persönliche Wichtigkeit*, *Nutzen* und *intrinsische Wertzuschreibung* fächerspezifisch (Mathematik, Deutsch, Physik und Chemie) und allgemein schulisch konzipiert. Bei der *Wichtigkeitsdimension* handelt es sich um die subjektive Bedeutung, in einem Fach (bzw. in der Schule) gute Leistungen zu erzielen, während die *Nutzenkomponente* dessen (instrumentelle) Bedeutung für die Pläne und Ziele der Person erfasst und der *intrinsische Wert* aus einer positiven emotionalen und kognitiven Bewertung resultiert. Erwartbar wäre also, dass insbesondere die subjektive Bedeutung und die intrinsischen Werte hoch mit Skalen des iSpo korrelieren, weil sie mit gefühls- und wertbezogenen Valenzen und dem intrinsischen Charakter korrespondieren. Im Kontext etwas anders gerichteter Überlegungen fokussiert der Ansatz von Dickhäuser und Schrahe (2006) gleichfalls die subjektive *Wichtigkeit des Sports*. Denn konzeptuell verwandt mit wertbezogenen Aspekten des Interesses ist die Wichtigkeit, die im Sinne des „attainment value“ (Wigfield & Eccles, 1992) auf die persönliche Bedeutung bezogen ist, in einem bestimmten Gegenstandsfeld gute Leistungen zu erzielen. Der Person-Gegenstands-Theorie des Interesses folgend, ist insbesondere die wertbezogene Valenz darüber hinaus eng mit dem *Selbstbild* verknüpft. Da „der Gegenstand des Interesses vorübergehend oder dauerhaft in die zentralen Regionen des Selbstkonzepts einer Person integriert wird“ (Krapp, 1999, S. 400), kann im Horizont eines mehrdimensionalen Selbstkonzeptmodells (sensu Shavelson, Hubner & Stanton, 1976) ein enger Zusammenhang mit dem

Selbstkonzept der sportlichen Fähigkeiten plausibel angenommen werden. Das Konzept der Freude am Sport(treiben) (*Physical Activity Enjoyment*, Kendzierski & DeCarlo, 1991) akzentuiert positive Gefühle im Zusammenhang mit sportlichen Aktivitäten und kann daher ebenfalls plausibel mit dem Interesse am Sport, insbesondere im Hinblick auf die gefühlsbezogene Wertkomponente, in Verbindung gebracht werden. Für die Zusammenhänge zwischen den genannten Konstrukten und dem individuellen Interesse am Sport ist zu erwarten, dass sie jeweils am stärksten mit der Handlungsmodalität unmittelbarer sportlicher Aktivität korrelieren, während vermittelnde und vor allem rezeptive Aktivitäten weniger eng verknüpft sind. Diese Ergebnisse deuten wir als (erste) Hinweise auf diskriminante Validität, auch wenn hierfür im strengen Sinne Konstrukte heranzuziehen wären, die inhaltlich zwar Nähe zu individuellem Interesse am Sport aufweisen, aber konzeptionell eigenständig sind.³

3 | METHODISCHE HINWEISE

Stichprobe

Es wurden insgesamt 643 Schülerinnen und Schüler (54,5 % weiblich) an verschiedenen Gymnasien in der Metropolregion Rhein-Neckar mit dem standardisierten Fragebogen iSpo untersucht.⁴ Der Altersrange reichte von 10 bis zu 20 Jahren und betrug im Mittel 14,6 Jahre (SD = 2.6). Von den Schülerinnen und Schülern waren 59 % in einem Sportverein aktiv und der Anteil von Heranwachsenden mit Migrationshintergrund lag bei 27 %.

Messinstrumente und Kriterien

Die Items der iSpo-Skalen wurden aufgrund der Untersuchungsergebnisse von Heim und Sohnsmeier (2016) im Hinblick auf zwei Items optimiert. Die Skalen zur Erfassung subjektiver schulischer Werte (SESSW, Steinmayr & Spinath, 2010) wurden fachspezifisch adaptiert und erfassen die drei Komponenten *persönliche Wichtigkeit*, *Nutzen* und *intrinsische Wertzuschreibung* für den Sportunterricht. Die jeweils drei Items pro Wertkomponente wurden auf einer Skala von „Trifft gar nicht zu“ [1] bis „Trifft genau zu“ [4] beantwortet. Zusätzlich wurde auf die von Dickhäuser und Schrahe (2006) entwickelte Skala „Wichtigkeit des Sportunterrichts“ mit fünfstufigem Antwortformat zurückgegriffen.⁵

Zur Erfassung des physischen Selbstkonzepts diente eine bewährte Skala, die bereits in der SPRINT-Studie (Gerlach et al., 2005) erfolgreich und mit guten psychometrischen Kennwerten eingesetzt wurde. Die Beantwortung der vier Items erfolgte gleichfalls auf einer vierstufigen Skala von „Stimmt nicht“ [1] bis „Stimmt genau“ [4].

3 Entsprechend geeignete Konstrukte im Horizont emotional-motivationaler Ansätze liegen nach unseren Recherchen nicht vor.

4 Das jeweilige N für die einzelnen Untersuchungsschritte unterscheidet sich: So erfolgte die Prüfung der konvergenen bzw. diskriminanten Validität an einer Teilstichprobe von N = 408, die Analyse der Test-Retest-Reliabilität mit N = 114 und die Überprüfung der prädiktiven Validität mit N = 82.

5 Auf das Item „Bei Aufgaben im Sport ist es mir persönlich egal, wenn ich mich ungeschickt anstelle“ wurde aufgrund geringer Trennschärfe in bisherigen Untersuchungen (Sohnsmeier & Heim, 2015) verzichtet.

Die Messung der Sportfreude erfolgte mit der von Jekauc, Voelkle, Wagner, Mewes und Woll (2013) für den deutschen Sprachraum adaptierten *Physical Activity Enjoyment Scale (PACES)*. Die PACES umfasst 16 Items, die auf einer Skala von „Stimmt gar nicht“ [1] bis „Stimmt genau“ [5] beantwortet wurden. Die internen Konsistenzen der verwendeten Skalen sind als gut zu bezeichnen (vgl. Tab. 1).

Im Hinblick auf die Kriteriumsvalidität wurde zusätzlich das Geschlecht, die Vereinsmitgliedschaft und die Häufigkeit des Sportengagements erfasst. Zur Analyse der prädiktiven Validität haben wir bei den Schülerinnen und Schülern in Klasse 10 zusätzlich die Sportnote erfragt und ob sie Sport als vierstündiges Fach in der Kursstufe wählen würden. Der Fragebogen wurde von studentischen Testleiter(innen)n administriert und von den Schülerinnen und Schülern jeweils zu Beginn des Sportunterrichts ausgefüllt.

Tab. 1: Messinstrumente zur Überprüfung der konvergenten Validität (dargestellt sind Beispiel-items), die internen Konsistenzen α sind die Werte der vorliegenden Stichprobe

Indikatoren	Operationalisierung	α
Subjektive schulische Werte	Skala zur Erfassung subjektiver schulischer Werte (Steinmayr & Spinath, 2010)	
	<i>Persönliche Wichtigkeit</i> <i>Es ist wichtig für mich, in Sport gut zu sein.</i>	.95
	<i>Nutzen</i> <i>Dinge, die ich in Sport lerne, finde ich für mein Leben im Allgemeinen nützlich.</i>	.88
	<i>Intrinsische Wertzuschreibung</i> <i>Ich mache Sport gerne.</i>	.89
Physisches Selbstkonzept	PSK (Gerlach et al., 2005) <i>Ich bin sehr gut im Sport.</i>	.88
Wichtigkeit von Sport	Wichtigkeit (Dickhäuser & Schrahe, 2006) <i>Für mich persönlich sind Fähigkeiten in Sport „nicht wichtig“ [1] – „sehr wichtig“ [5].</i>	.82
Sportenjoyment	Physical Activity Enjoyment Scale (Jekauc et al., 2013) <i>Mich zu bewegen ... macht mir Freude.</i>	.93
Kriterien	Geschlecht Vereinsmitgliedschaft Sporthäufigkeit Sportnote im letzten Zeugnis Wahlentscheidung „Sport als 4-stündiges Fach in der Kursstufe“	

Statistische Auswertung⁶

Im ersten Schritt wurden konfirmatorische Faktorenanalysen (CFA) durchgeführt, um die postulierte Faktorstruktur zu prüfen. Dazu wurde dem klassischen CFA-Modell eine hierarchische (G-Faktor CFA) und eine bifaktorielle Lösung (Bi-Faktor CFA) gegenübergestellt (Abb. 1).

6 Da die hier verwendeten Verfahren gegenwärtig nicht zum Standardrepertoire der empirischen Sportpädagogik gehören, stellen wir unsere Vorgehensweise für methodisch interessierte Leserinnen und Leser (sehr) detailliert vor. Anderen mag eine kursorische Lektüre genügen.

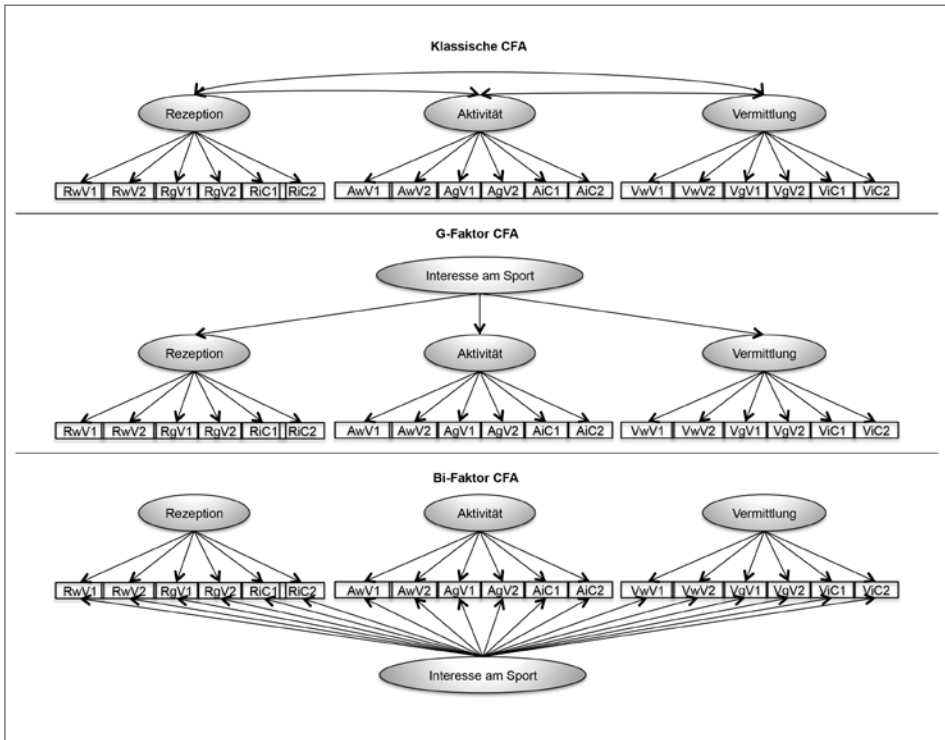


Abb. 1: Grafische Darstellung der drei Modelle (klassische CFA, hierarchische CFA (G-Faktor) und bifaktorielle CFA) nach Morin et al. (2016)

Das bifaktorielle Modell ist durch einen Generalfaktor, durch Gruppenfaktoren und durch eine entsprechende Mustermatrix mit einer bifaktoriellen Struktur charakterisiert. Generalfaktor und Gruppenfaktoren entsprechen dem G-Faktor und den Subdomänen im hierarchischen CFA-Modell zweiter Ordnung (G-Faktor CFA). Im Gegensatz zu dieser hierarchischen Konstruktion werden im bifaktoriellen Modell jedoch für jedes Item die Ladungen sowohl auf dem Generalfaktor als auch auf den Gruppenfaktoren frei geschätzt. D. h., jedes Item lädt auf dem Generalfaktor und auf einem Gruppenfaktor (Mustermatrix mit bifaktorieller Struktur). Im hierarchischen CFA-Modell zweiter Ordnung steht der G-Faktor nur in direkter Verbindung zu den Faktoren erster Ordnung (nicht zu den Items). Damit ist die G-Faktor-CFA nicht in der Lage, zu modellieren, dass (wenigstens einige) Items sowohl direkt durch den Generalfaktor als auch durch die spezifischen Gruppenfaktoren beeinflusst werden.

Die bisherige theoretische und empirische Auseinandersetzung mit dem Konstrukt *Interesse am Sport* legt die Operationalisierung durch ein Instrument nahe, das sich durch einen latenten Generalfaktor und drei latente Subdomänen konstituiert (Heim & Sohnsmeier, 2016). Während der Generalfaktor die gesamte Breite des Konstrukts *Interesse am Sport* widerspiegelt, spezifizieren

die Subdomänen bestimmte Bereiche innerhalb dieser konzeptuellen Breite (Interesse an der Rezeption, der Aktivität und der Vermittlung). Die Items stehen damit zum einen in direkter Verbindung zum Generalfaktor und zum anderen in direkter Verbindung zu den Subdomänen, sodass der bifaktorielle Messzugriff eine äußerst tragfähige Modellierung dieser Annahme darstellt (Myers, Martin, Ntoumanis, Celimli & Bartholomew, 2014).

Um den Modell-Fit zu evaluieren, stehen verschiedene Fit-Indizes zur Verfügung: Wie von Marsh, Hau und Wen (2004) vorgeschlagen, wird der Tucker-Lewis-Index (TLI), der Comparative-Fit-Index (CFI), der Root Mean Square Error of Approximation (RMSEA) und der Standardized Root Mean Square Residual (SRMR) berichtet. Für den RMSEA ist zudem das entsprechende Konfidenzintervall (CI) angegeben. Die Modelle wurden anhand von Cut-off-Werten für die Fit-Indizes beurteilt: RMSEA \leq .05 bzw. .08, SRMR \leq .05 bzw. .10 und CFI und TLI \geq .95 bzw. .90 für eine guten bzw. akzeptablen Fit.⁷ Die Berechnungen wurden mit Mplus 7.31 (Muthén & Muthén, 1998-2012) unter Verwendung des robusten Maximum-Likelihood-Schätzers (MLR) durchgeführt. Denn der MLR-Schätzer ist robust gegenüber Verletzungen der Normalverteilungsannahme und zeigt sich effizient in der Schätzung von latenten Variablen, die auf Basis von Items mit mehrstufigen Antwortformaten erhoben wurden (Scalas, Marsh, Nagengast & Morin, 2013).

Die Ergebnisse der bifaktoriellen konfirmatorischen Faktorenanalyse wurden ergänzt um die Berechnung der „Explained common variance (ECV)“ (Rodriguez, Reise & Haviland, 2016) sowohl auf Item- als auch Faktorebene.⁸ Hohe ECV-Werte bedeuten einen starken Generalfaktor. Auf Item-Level zeigen die I-ECV-Werte ebenfalls den Anteil gemeinsamer Varianz an, die dem Generalfaktor zuzuordnen ist. Hohe Werte sprechen dann dafür, dass diese Items „reine“ Marker für das breite Konstrukt *Interesse am Sport* darstellen, während niedrige Werte die Spezifität des Gruppenfaktors wiedergeben.

Die Prüfung der Messinvarianz für Geschlecht und Vereinsmitgliedschaft erfolgte anhand der von Marsh, Morin, Parker und Kaur (2014) vorgeschlagenen Taxonomie (Tab. 2).

Tab. 2: Taxonomie zur Prüfung der Messinvarianz (nach Marsh et al., 2014)

Modell	Invariante Modellparameter
Modell 1	None (konfigurale Invarianz: gleiche Struktur der Faktorladungsmatrix)
Modell 2	Faktorladungen (metrische Invarianz)
Modell 3	Faktorladungen, Konstanten (skalare Invarianz)
Modell 4	Faktorladungen, Messfehlervarianzen, Konstanten (strikte Invarianz)

7 Dabei handelt es sich allerdings nur um Richtwerte (Chen, Curran, Bollen, Kirby & Paxton, 2008; Marsh et al., 2004).

8 „ECV indexes variance specific to a general factor by taking the ratio of variance explained by a general factor and dividing it by the variance explained by a general and group factors where factors are assumed to be uncorrelated“ (Rodriguez et al., 2016, S. 144).

So prüft der Test auf konfigurale Invarianz (gleiche Struktur der Faktorladungsmatrix) in Modell 1, ob die Konstrukte die gleiche Dimensionalität aufweisen und eine identische Zuordnung der Items zu den latenten Faktoren in den spezifischen Gruppen vorliegt. In Modell 2 (metrische Invarianz) werden die Faktorladungen gleichgesetzt, d. h., es wird geprüft, ob der Zusammenhang zwischen den Faktoren und den zugehörigen Indikatoren in allen Gruppen identisch ist. Bei vorliegender metrischer Invarianz wird in Modell 3 die Gleichheitsrestriktion auf die Konstanten (intercepts) der Indikatoren erweitert (skalare Invarianz). Wird die Annahme skalarer Invarianz nicht verworfen, so liegt eine starke faktorielle Invarianz vor. Diese ist eine wesentliche Voraussetzung, um eine nicht verzerrte Schätzung der latenten Faktoren über die verschiedenen Gruppen hinweg zu gewährleisten – und um einen sinnvollen Mittelwertvergleich durchzuführen. Eine Erweiterung der skalare Invarianz ist die zusätzliche Restriktion der Messfehlervarianzen der Indikatoren (item uniqueness), die als strikte faktorielle Varianz bezeichnet wird (Marsh et al., 2014).

Um zu überprüfen, ob sich die Modellannahmen halten lassen, werden die Fit-Indizes des höheren (sparsameren) Modells mit dem jeweils niedrigeren Modell verglichen (Modell 1 vs. Modell 2; Modell 2 vs. Modell 3, Modell 3 vs. Modell 4). Als Cut-off-Wert, bei dem die Modellannahmen verworfen werden sollten, wurde von Chen (2007) eine Verschlechterung der inkrementellen Fit-Indizes von $-.01$ oder eine Veränderung des RMSEA um weniger als $.015$ vorgeschlagen. Wenn sich also z. B. der CFI um weniger als $-.01$ verschlechtert oder der RMSEA um weniger als $.015$ ansteigt, spricht dies für das sparsamere Modell.

Die Invarianzprüfungen bilden die Basis für weitere Untersuchungen der Kriteriumsvalidität. Ob sich bekannte Muster der Vereinsmitgliedschaft und Geschlechtstypik auch im sportsspezifischen Interesse zeigen, haben wir mithilfe von latenten Mittelwertvergleichen analysiert.⁹ Dazu wurden die Mittelwertdifferenzen im Rahmen des skalaren Invarianzmodells geschätzt und auf Signifikanz geprüft. In diesem Modell sind die Faktorladungen sowie die Itemkonstanten in beiden Gruppen äquivalent. Die Faktormittelwerte für Mädchen resp. Vereinsmitglieder dienten als Referenz ($= 0$) und die Faktormittelwerte für Jungen resp. Nichtmitglieder wurden frei geschätzt.¹⁰ Die zugehörige Prüfstatistik ist approximativ z-verteilt und die gefundenen Differenzen entsprechen Cohens d , sodass aus der Ausprägung der Differenz direkt die zugehörige Effektstärke abgelesen werden kann (Steinmetz, 2010; Steinmetz, Schmidt, Tina-Booh, Wieczorek & Schwartz, 2008). Ergänzt wird die Untersuchung der Kriteriumsvalidität durch eine Inspektion der Korrelationen zwischen dem Interesse am Sport und der Häufigkeit des Sporttreibens.

9 Zur Untersuchung von *Mittelwertdifferenzen* werden zumeist Varianzanalysen und t-Tests in Anschlag gebracht. Das typische Vorgehen umfasst das Aggregieren von Items zu Mittelwerten, die dann zwischen den Gruppen verglichen werden. Allerdings stimmen die zu Mittelwerten zusammengefassten beobachteten Indikatoren nicht mit dem latenten Konstrukt überein, das mit den Indikatoren gemessen werden soll (Steinmetz, 2010). Der vorliegende Ansatz erlaubt demgegenüber die Schätzung der latenten Mittelwerte und deren Vergleich über verschiedene Gruppen. Die statistischen Tests haben eine höhere Teststärke als die klassischen Methoden und ermöglichen die nicht verzerrte Schätzung der Mittelwerte.

10 Die Spezifikation eines Strukturgleichungsmodells in Mplus beinhaltet das Freisetzen von zu schätzenden Parametern, das Fixieren von Parametern auf vorgegebene Werte und das Gleichsetzen von Parametern, die dann auf einen gemeinsamen Wert geschätzt werden.

Ebenfalls mithilfe von Korrelationsprüfungen werden Zusammenhänge zwischen den Interessenskalen und den genannten Instrumenten im Hinblick auf die konvergente Validität in den Blick genommen. Abgeschlossen werden die Validierungsschritte mit einer schrittweisen logistischen Regression, um zu prüfen, inwieweit der iSpo-Fragebogen die geplante Kurswahl in der gymnasialen Oberstufe – in querschnittlich statistischem, nicht längsschnittlichem Sinn – „vorherzusagen“ vermag.

Der Anteil fehlender Werte lag bei den analysierten Variablen zwischen 0.2 und 4.3 %. Angenommen wurde, dass dieser Ausfall bedingt zufällig ist (Missing at Random: MAR), d. h., dass er von weiteren, im Datensatz beobachteten Variablen abhängen kann. Im Rahmen der Faktorenanalysen wurden diese fehlenden Werte durch den FIML-Algorithmus geschätzt (Full Information Maximum Likelihood).

4 | ERGEBNISSE

Faktorielle Validität

Die Ergebnisse der konfirmatorischen Faktorenanalysen belegen die bifaktorielle Struktur (Morin et al., 2016) mit dem G-Faktor *Interesse am Sport* und den Subskalen *Rezeption*, *Aktivität* und *Vermittlung* mit jeweils sechs Items.

Tab. 3: Fit-Indizes von CFA, hierarchischer CFA und bifaktorieller CFA

Modell	Chi ² /df	CFI	TLI	RMSEA	RMSEA 90 % CI	SRMR	AIC	BIC adjust.
CFA	622.9/132	.926	.914	.076	.070 - .082	.045	29108.5	29182.1
G-CFA	622.9/132	.926	.914	.076	.070 - .082	.045	29108.5	29182.1
Bi-CFA	458.6/117	.948	.933	.067	.061 - .074	.037	28941.9	29034.8

Die Fit-Indizes des bifaktoriellen Modells liefern zufriedenstellende Werte (χ^2 (117) = 458.6, $p < .001$, CFI = .948, TLI = .933, RMSEA = .067 [.061 -.074], SRMR = .037) und damit eine akzeptable Passung an die Daten. Die bifaktorielle Struktur des iSpo mit drei Subskalen Rezeption, Aktivität und Vermittlung (Heim & Sohnsmeier, 2016) konnte somit repliziert werden.

Anhand der signifikanten Ladungen der 18 Items auf dem G-Faktor wird deutlich, dass dieser klar definiert ist ($|\lambda|$ = .18 bis .88, M = .53) und die Subskalen wohlbestimmt sind. Die Faktorladungen liegen zwischen $|\lambda|$ = .03 und .84, M = .54.¹¹ Damit zeigt sich der Beitrag der Items einerseits zur Spezifität und andererseits zum übergeordneten G-Faktor. Insbesondere wird anhand der „Item Explained Common Variance (I-ECV)“ (Rodriguez et al., 2016)

¹¹ Nebenladungen sind nicht zugelassen und auf 0 fixiert.

Tab. 4: Ergebnis der konfirmatorischen bifaktoriellen Faktorenanalyse, standardisierte Faktorladungen und „Item Explained Common Variance (I-ECV)“ (Rodriguez et al., 2016)

Items	Faktorladung λ				I-ECV
	1	2	3	iSpo	
Es macht mir Spaß, Sport im Fernsehen anzuschauen.	.67			.28	.15
Das Lesen von Sportnachrichten in der Zeitung bereitet mir Freude.	.84			.35	.15
Es ist mir egal, wenn ich eine Sportsendung im Fernsehen verpasse.*	.45			.18	.14
Es ist wichtig für mich, Sportberichte in der Zeitung zu lesen.	.70			.36	.21
Wenn ich genügend Zeit hätte, würde ich häufiger Sport im Fernsehen anschauen.	.70			.32	.17
Wenn ich die Gelegenheit habe, lese ich Sportberichte in der Zeitung.	.80			.33	.15
Ich mache gerne mehrmals in der Woche Sport.		.44.		.76	.75
Sport ist meine liebste Freizeitaktivität.		.03		.86	.99
Sporttreiben ist in meinem Leben wichtig.		.42		.75	.76
Es ist wichtig für mich, mehrmals in der Woche Sport zu machen.		.45		.76	.74
Immer, wenn ich die Gelegenheit habe, gehe ich sportlichen Aktivitäten nach.		-.07		.88	.99
Ich will auf meine Sporttermine in der Woche nicht verzichten.		.43		.63	.68
Es macht mir Spaß, anderen im Sport etwas beizubringen.			.70	.56	.39
Es macht mir Spaß, anderen bei der Verbesserung ihrer Fehler im Sport zu helfen.			.77	.48	.28
Es hat eine große Bedeutung für mich, anderen im Sport etwas beizubringen.			.60	.55	.46
Es ist wichtig für mich, anderen bei der Verbesserung ihrer Fehler im Sport zu helfen.			.66	.48	.35
In meiner Freizeit bringe ich gerne anderen im Sport etwas bei.			.63	.59	.47
Immer, wenn ich Zeit habe, versuche ich, anderen bei der Verbesserung ihrer Fehler im Sport zu helfen.			.65	.52	.39

Anmerkung: Komponente 1 = Rezeption, Komponente 2 = Aktivität, Komponente 3 = Vermittlung, *umcodiert.

deutlich, dass die Aktivitätsitems stärker als die Items zur Vermittlung und diese wiederum stärker als die Items zur Rezeption zum G-Faktor beitragen. Insgesamt zeigt die „Explained Common Variance (ECV)“ (Rodriguez et al., 2016) mit einem Wert von .48, dass der Generalfaktor 48 % der gemeinsamen Varianz erklärt und 52 % auf die Gruppenfaktoren zurückzuführen ist.

Reliabilität

Die an der Gesamtstichprobe (N = 643) berechneten internen Konsistenzen der Skalen (Rezeption: $\alpha = .89$; Aktivität: $\alpha = .92$; Vermittlung: $\alpha = .94$; G-Faktor: $\alpha = .91$) sind als gut zu bezeichnen. Die Test-Retest-Reliabilitäten, die an einer Teilstichprobe von N = 114 über einen Zeitraum von zwei Wochen bestimmt wurden, lagen mit $.60 < r_{tt} < .80$ im akzeptablen Bereich. (Tab. 5).

Tab. 5: Test-Retest-Reliabilität (r_{tt}) und Cronbachs α (N = 114)

	r_{tt}	Cronbachs α (Test)	Cronbachs α (Retest)
Rezeption	.80	.87	.86
Aktivität	.78	.86	.87
Vermittlung	.60	.89	.94
G-Faktor-iSpo	.79	.88	.90

Invarianzprüfung

Ein weiteres Ziel der Studie war der Nachweis der Messinvarianz der iSpo-Skalen für die Variablen *Geschlecht* und *Vereinsmitgliedschaft*. Die folgenden Analysen basieren auf der Taxonomie zur Prüfung der Messinvarianz nach Marsh et al. (2014) (Tab. 2).

Tab. 6: Fit-Statistiken der Geschlechts- und Vereinssport-Invarianz-Modelle

	Chi ² /df	CFI	TLI	RMSEA [CI]	SRMR
Geschlecht					
Konfigurale Invarianz	654.2/234	.936	.916	.076 [.069 -.083]	.048
Metrische Invarianz	705.5/266	.933	.923	.073 [.067 -.080]	.062
Skalare Invarianz	722.9/280	.932	.926	.072 [.065 -.078]	.062
Strikte Invarianz	790.1/298	.925	.923	.073 [.067 -.079]	.072
Sportverein					
Konfigurale Invarianz	633.4/234	.934	.914	.073 [.066 -.080]	.041
Metrische Invarianz	688.9/266	.930	.920	.070 [.064 -.077]	.071
Skalare Invarianz	722.3/280	.927	.920	.070 [.064 -.077]	.075
Strikte Invarianz	768.0/298	.922	.920	.070 [.064 -.076]	.086

Sowohl für das Geschlecht als auch für die Mitgliedschaft im Sportverein belegen die Ergebnisse konfigurale, metrische, skalare und strikte Invarianz: Vergleicht man die Modelle, zeigt sich kein – orientiert an den vorgeschlagenen Cut-off-Werten von $-.01$ (CFI & TLI) und $.015$ (RMSEA) (Chen, 2007) – verschlechterter Fit (Tab. 6).

Kriteriumsvalidität

Es wurde angenommen, dass Jungen und Vereinsmitglieder über ein höheres sportliches Interesse verfügen als Mädchen und Nichtmitglieder. Wie bereits in der Studie von Heim und Sohnsmeier

er (2016), zeigten sich signifikante Geschlechtsunterschiede: Die latenten Faktormittelwerte der männlichen Probanden waren signifikant höher sowohl für die Skala *Rezeption* als auch für den *G-Faktor*. Demgegenüber ergaben sich keine Unterschiede für die Skalen *Aktivität* und *Vermittlung*. Zudem zeigten sich signifikante Unterschiede zwischen Vereinsmitgliedern und Nichtmitgliedern: Vereinsmitglieder weisen für die Skala *Aktivität* und den *G-Faktor* signifikant höhere Werte auf (Tab. 7). Für die *Geschlechtsunterschiede* sind die Effektstärken als klein bis mittel einzustufen, die Unterschiede zwischen Vereinsmitgliedern und Nichtmitgliedern zeigen mittlere bis hohe Effekte.

Tab. 7: Latente Faktormittelwerte der Skalen nach Geschlecht und Vereinsmitgliedschaft

	Geschlecht	
	Weiblich	Männlich
Rezeption	.000	.46; p < .001
Aktivität	.000	.18; p = .138
Vermittlung	.000	-.04; p = .675
G-Faktor-iSpo	.000	.23; p = .012
	Vereins-sport	
	Mitglied	Nichtmitglied
Rezeption	.000	-.09; p = .441
Aktivität	.000	-.46; p = .019
Vermittlung	.000	-.21; p = .112
G-Faktor-iSpo	.000	-.89; p < .001

Anmerkung: Parameterschätzungen aus Modell 3 (skalare Invarianz), in dem die Itemkonstanten und Faktorladungen invariant über Geschlecht bzw. Vereinsmitgliedschaft sind. Mittelwertsunterschiede entsprechen Cohens *d* (Steinmetz et al., 2008). Die Faktormittelwerte für Mädchen resp. Vereinsmitglieder dienen als Referenz (= 0) und die Faktormittelwerte für Jungen resp. Nichtmitglieder wurden frei geschätzt.

Entsprechen der theoretischen Annahmen zeigt sich darüber hinaus ein starker Zusammenhang zwischen dem Sportengagement und der Aktivitätsskala ($r = .59$), während die Korrelationen mit der Skala *Rezeption* nur gering ausgeprägt sind ($r = .19$).

Konvergente Validität

Zur Überprüfung der konvergenten Validität wurden die Korrelationen der iSpo-Skalen mit den beschriebenen Messinstrumenten simultan bestimmt.

Erwartungsgemäß ergaben sich positive Zusammenhänge mit allen erhobenen Variablen. Diese erreichten hohe Werte für die Subskala *Aktivität*, während das Interesse an rezeptiven Handlungsmodalitäten nur schwache Korrelationen aufweist. Die Assoziationen zwischen der Vermittlungskomponente und den hier berücksichtigten Skalen bewegen sich überwiegend im Bereich moderater

Tab. 8: Korrelationen zwischen den iSpo-Skalen und SESSW, Wichtigkeit von Sport, Sportfreude sportlichem Fähigkeitsselbstkonzept, angegeben ist der Pearson-Korrelationskoeffizient r , * = signifikant ($< .05$); ** = hochsignifikant ($< .01$)

	G-Faktor-iSpo	Rezeption	Aktivität	Vermittlung
SESSW Intrinsische Werte Sport	.65**	.24**	.78**	.54**
SESSW Nützlichkeit Sport	.62**	.27**	.65**	.55**
SESSW Wichtigkeit Sport	.61**	.25**	.67**	.50**
Wichtigkeit von Sport	.61**	.24**	.70**	.51**
Freude am Sporttreiben	.56**	.15*	.68**	.51**
Fähigkeitsselbstkonzept Sport	.62**	.27**	.67**	.51**

Größenordnungen. Dieses Muster spricht einerseits für die konvergente Validität des iSpo, weil vor allem die Handlungsmodalitäten der Aktivität und Vermittlung hohe korrelative Werte zeigen. Die gering ausgeprägten Zusammenhänge mit der Skala rezeptiven Interesses stützen andererseits die Annahme diskriminanter Validität.

Prädiktive Validität

Berechnet wurde eine schrittweise logistische Regression, in der eine Schätzung der Wahrscheinlichkeit erfolgte, mit der Sport als vierstündiges Fach in der Kursstufe gewählt wird (in Anlehnung an Steinmayr & Spinath, 2010).

Tab. 9: Ergebnisse schrittweise logistischer Regressionsanalysen zur Vorhersage der Kurswahl Sport als vierstündiges Fach in der Kursstufe (N = 82, Wahl Sport vierstündig N = 16)

Prädiktor	Modell 1			Modell 2			Modell 3		
	B	p	Exp[b]	B	p	Exp[b]	B	p	Exp[b]
Sportnote	2.55	< .001	12.74	1.75	.023	5.74	1.73	.028	5.62
PSK				1.85	.038	6.35	1.77	.085	5.90
G-Faktor-iSpo							.11	.891	1.11
Sportnote	2.55	< .001	12.74	1.75	.023	5.74	1.90	.021	6.71
PSK				1.85	.038	6.35	1.87	.037	6.51
Rezeption							-.218	.554	.804
Sportnote	2.55	< .001	12.74	1.75	.023	5.74	1.54	.056	4.66
PSK				1.85	.038	6.35	1.56	.112	4.76
Aktivität							.473	.462	1.60
Sportnote	2.55	< .001	12.74	1.75	.023	5.74	1.80	.022	6.03
PSK				1.85	.038	6.35	1.68	.082	5.35
Vermittlung							.149	.709	1.16

In Modell 1 zeigt sich, dass die Sportnote ein signifikanter Prädiktor für die Kurswahl darstellt. Die Hinzunahme des physischen Selbstkonzepts in Modell 2 schwächt diesen Effekt leicht ab, die Sportnote und das physische Selbstkonzept erklären die Kurswahl in etwa zu gleichen Anteilen. Die iSpo-Skalen haben schließlich über die Sportnote und das physische Selbstkonzept hinaus keine Erklärungskraft für die Kurswahl, wie Modell 3 belegt (Tab. 9).

5 | DISKUSSION

Ziel der vorliegenden Arbeit war die Validierung des vor Kurzem in dieser Zeitschrift vorgestellten Fragebogens zur Messung von Interesse am Sport (Heim & Sohnsmeier, 2016). Die dimensionale Struktur des für diese Studie verfeinerten Instruments konnte an einer Stichprobe von $N = 643$ repliziert werden (faktorielle Validität). Auf Basis der bisherigen Analysen lässt sich das Interesse am Sport im Hinblick auf die Handlungsmodalitäten Rezeption, Aktivität und Vermittlung gut erfassen. Diese Beschränkung auf drei Facetten des Interesses am Sport muss jedoch kritisch diskutiert werden. Zum einen sind weitere Dimensionen denkbar, die relevante Aspekte betreffen, aber derzeit unberücksichtigt bleiben – zum Beispiel verstehend-reflektierende Handlungsmodalitäten. Zum anderen konnte die theoretisch postulierte Differenzierung von gefühls- und wertbezogener Valenz sowie intrinsischer Motivation (Heim & Sohnsmeier, 2016) empirisch nicht bestätigt werden. Während dieses Ergebnis im Einklang mit den Analysen von Schiefele, Krapp, Wild und Winteler (1993) ist, finden sich für das situationale Interesse Hinweise, dass eine gefühls- und eine wertbezogene Komponente unterschieden werden kann (Linnenbrink-Garcia et al., 2010). Zudem betonen O’Keefe und Linnenbrink-Garcia (2014) die Bedeutung einer gefühls- und wertbezogene Komponente des individuellen Interesses in Leistungssituationen: „As expected, high levels of affect- and value-related interest were associated with the optimization of both performance and self-regulatory resources“ (O’Keefe & Linnenbrink-Garcia, 2014, S. 76). Vor diesem Hintergrund erscheint uns diese Differenzierung auch im Hinblick auf das individuelle Interesse weiterhin sinnvoll.

Die bifaktorielle Analyse identifiziert zudem einen G-Faktor, der die gesamte Breite des Konstrukts Interesse am Sport widerspiegelt, wenngleich es sich nicht um ein eindimensionales Instrument handelt, da der G-Faktor und die Gruppenfaktoren etwa zu gleichen Teilen zur Varianzaufklärung beitragen. Mit Blick auf die Konstruktion einer eindimensionalen Kurzskaala Interesse am Sport liefern die I-ECV-Werte zusätzlich wichtige Hinweise. So empfehlen Stucky, Thissen und Edelen (2013) zur Konstruktion von eindimensionalen Skalen die Auswahl der Items mit den höchsten I-ECV-Werten.

Die internen Konsistenzen und die Test-Retest-Reliabilitäten sind sehr zufriedenstellend. Die Analysen zur Messinvarianz stützen die Annahme strikter Invarianz für das Geschlecht und die Vereinsmitgliedschaft, d. h., das Messmodell ist nicht durch das Geschlecht der Probanden oder die Zugehörigkeit zu einem Sportverein konfundiert. Die nachfolgend untersuchten Gruppenunterschiede entsprechen im Wesentlichen den Erwartungen. Überraschend ist, dass für die Vereinsmitgliedschaft keine Unterschiede für die Vermittlungsskala gefunden wurden. Mit Blick auf das

Geschlecht zeigten sich nur Unterschiede für den G-Faktor und die Skala Rezeption. Dies weist auf die Relevanz weiterer Faktoren jenseits des Interesses hin, die das Sportengagement bestimmen¹², könnte aber zudem auch mit der hier untersuchten gymnasialen Stichprobe verknüpft sein.

Die Zusammenhangsanalysen mit verschiedenen Instrumenten, die eine konzeptuelle Nähe zum Interessenkonstrukt aufweisen, stützen die konvergente und – mit Einschränkungen – die diskriminante Validität. Insbesondere die durchgängig hohen Korrelationen mit dem G-Faktor und der Skala Aktivität belegen die konvergente Validität, während die geringen Zusammenhänge mit der Skala Rezeption für die diskriminante Validität sprechen. Im strengen Sinne wären für den Nachweis diskriminanter Validität Konstrukte heranzuziehen, die inhaltlich zwar Nähe zu individuellem Interesse am Sport aufweisen, aber konzeptionell eigenständig sind. In zukünftigen Studien könnten daher Zusammenhangsanalysen der iSpo-Skalen mit kognitiv akzentuierten Motivationskonstrukten hier tiefere Erkenntnisse liefern.

Mit Vorsicht sind die Ergebnisse zur prädikativen Validität zu betrachten. Denn es handelt sich einerseits um eine sehr kleine Teilstichprobe, in der nur 16 Schüler(innen) angaben, Sport als vierstündiges Fach in der Kursstufe zu wählen. Da durchgängig die Sportnote ein wesentlicher Prädiktor ist, scheint eher das Aufwand-Nutzen-Kalkül die Kurswahl zu bedingen. Andererseits bedürfte eine belastbarere Prüfung der prädiktiven Validität eines längsschnittlichen Untersuchungsdesigns.

Abgesehen von den genannten Grenzen, bietet das optimierte und hier vorgestellte Instrument eine ökonomisch und psychometrisch überzeugende Möglichkeit zur Analyse des Interesses am Sport bei Jugendlichen. Der Fragebogen iSpo ist in der derzeitigen Form besonders gut geeignet zur Erforschung des außerschulischen Sportengagements und seiner Entwicklung im Lebenslauf. Untersuchungen im Kontext des Sportunterrichts sind vor dem Hintergrund des Angebots-Nutzungs-Modells möglich. So ist z. B. die Frage nach der Bedeutung des individuellen Interesses für die Wahrnehmung der Unterrichtsqualität aus Schülersicht auch im Sportunterricht von großem Interesse (Rahn, Gruehn, Keune & Fuhrmann, 2016). Darüber hinaus ist der Fragebogen u. E. zur Prüfung von Wirkungen einer sportunterrichtlichen Erziehung zum Sport geeignet und kann im Rahmen von Mediator- oder Moderatormodellen unter den Prämissen einer Erziehung durch Sportunterricht eingesetzt werden. Mit Blick auf die curricularen Programmatiken wäre aus unserer Sicht allerdings eine Ergänzung des iSpo durch verstehend-reflektierende Aktivitäten angezeigt.

12 In die gleiche Richtung weisen die korrelativen Befunde. Trotzdem zeigt sich vor dem Hintergrund der moderaten Korrelation die Fruchtbarkeit des Interessenansatzes.

LITERATUR

- Abel, J. (2002). Kurswahl aus Interesse? Wahlmotive in der gymnasialen Oberstufe und Studienwahl. *Die Deutsche Schule*, 94 (2), 192-203.
- Burmann, U., Seyda, M., Heim, R. & Konowalczyk, S. (2016). Individualisierungstendenzen im Sport von Heranwachsenden – revisited. *Sport und Gesellschaft*, 13 (2), 113-143.
- Chen, A., Darst, P. W. & Pangrazi, R. P. (1999). What constitutes situational interest? Validating a construct in physical education. *Measurement in Physical Education and Exercise Science*, 3 (3), 157-180.
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J. & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, 36 (4), 462-494.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464-504.
- Dickhäuser, O. & Schrahe, K. (2006). Sportliches Fähigkeitsselbstkonzept und allgemeiner Selbstwert. *Zeitschrift für Sportpsychologie*, 13 (3), 98-103.
- Eccles, J. S., S., A. T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco: Freeman.
- Gerlach, E., Wilsmann, F., Kehne, M., Oesterreich, C. & Stucke, C. (2005). *Sportunterricht in Deutschland. Dokumentation der Erhebungsinstrumente der SPRINT-Studie*. Paderborn: Department Sport und Gesundheit.
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 143-171). Berlin: Springer.
- Heim, R. & Sohnsmeier, J. (2016). Interesse am Sport – Theoretische Konzeptualisierung und Konstruktion eines Fragebogens. *Zeitschrift für sportpädagogische Forschung*, 4 (2), 21-60.
- Horn, J. L. & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117-144.
- Jekauc, D., Voelke, M., Wagner, M. O., Mewes, N. & Woll, A. (2013). Reliability, validity, and measurement invariance of the German version of the Physical Activity Enjoyment Scale. *Journal of Pediatric Psychology*, 38 (1), 104-115.
- Kendzierski, D. & DeCarlo, K. J. (1991). Physical activity enjoyment scale: Two validation studies. *Journal of Sport and Exercise Psychology*, 13 (1), 50-64.
- Köller, O., Daniels, Z., Schnabel, K.U. & Baumert, J. (2000). Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse. *Zeitschrift für Pädagogische Psychologie*, 14 (1), 26-37.
- Krapp, A. (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie in Erziehung und Unterricht*, 44, 185-201.
- Krapp, A. (1999). Intrinsische Lernmotivation und Interesse. Forschungsansätze und konzeptuelle Überlegungen. *Zeitschrift für Pädagogik*, 45, 387-406.
- Krapp, A. (2010). Interesse. In D. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 311-323). Weinheim: Beltz-PVU.
- Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., et al. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*, 70 (4), 647-671.
- Marsh, H. W., Hau, K. T. & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11 (3), 320-341.
- Marsh, H. W., Morin, A. J. S., Parker, P. D. & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, 10 (1), 85-110.
- Möller, J. & Trautwein, U. (2015). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 177-199). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Morin, A. J. S., Arens, A. K. & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, 23 (1), 116-139.

- Muthén, L. K. & Muthén, B. O. (1998-2012). *Mplus user's guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.
- Myers, N. D., Martin, J. J., Ntoumanis, N., Celimli, S. & Bartholomew, K.J. (2014). Exploratory bifactor analysis in sport, exercise, and performance psychology: A substantive-methodological synergy. *Sport, Exercise, and Performance Psychology*, 3 (4), 258-272.
- O'Keefe, P. A. & Linnenbrink-Garcia, L. (2014). The role of interest in optimizing performance and self-regulation. *Journal of Experimental Social Psychology*, 53 (0), 70-78.
- Prenzel, M. (1988). *Die Wirkungsweise von Interesse. Ein pädagogisch-psychologisches Erklärungsmodell*. Opladen: Westdeutscher Verlag.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32 (2), 163-173.
- 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.
- Rodriguez, A., Reise, S. P. & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21 (2), 137-150.
- Scalas, L. F., Marsh, H. W., Nagengast, B. & Morin, A. J. S. (2013). Latent-variable approaches to the Jameian model of importance-weighted averages. *Personality & Social Psychology Bulletin*, 39 (1), 100-114.
- Schiefele, U. & Schaffner, E. (2015). Motivation. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 153-175). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schiefele, U., Krapp, A., Wild, K. P. & Winteler, A. (1993). Der Fragebogen zum Studieninteresse (FSI). *Diagnostica*, 39, 335-351.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407-441.
- Sohnsmeier, J. & Heim, R. (2015). Selbstwert und sportliches Fähigkeitskonzept im Jugendalter: der moderierende Effekt subjektiver Wichtigkeit. *Zeitschrift für sportpädagogische Forschung*, 3 (1), S. 61-83.
- Steinmayr, R. & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW). *Diagnostica*, 56 (4), 195-211.
- Steinmetz, H. (2010). Estimation and comparison of latent means across cultures. In P. Schmidt, J. Billiet & E. Davidov (Eds.), *Cross-cultural analysis: Methods and applications* (pp. 85-116). Routledge: Taylor & Francis Group.
- Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S. & Schwartz, S. H. (2008). Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality & Quantity*, 43 (4), 599-616.
- Stucky, B. D., Thissen, D. & Edelen, M. O. (2013). Using logistic approximations of marginal trace lines to develop short assessments. *Applied Psychological Measurement*, 37 (1), 41-57.
- Wigfield, A. & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12 (3), 265-310.