

Messung naturwissenschaftlich-experimenteller Problemlösefähigkeit und deren Zusammenhang mit kognitiven und nicht-kognitiven Persönlichkeitsmerkmalen von Schülerinnen und Schülern

Kerstin Höner & Dieter Wenzel

Technische Universität Braunschweig, Institut für Fachdidaktik der Naturwissenschaften, Abteilung Chemie und Chemiedidaktik

Kurzfassung

Die vorliegende Studie untersucht anhand einer interaktiven naturwissenschaftlichen Aufgabensequenz die Zusammenhänge zwischen experimentellem Problemlöseerfolg und kognitiven, motivationalen sowie leistungsfördernden Persönlichkeitsmerkmalen von Schülerinnen und Schülern.

Bei der Messung experimenteller Problemlösefähigkeiten hat sich in vielen Studien gezeigt, dass die Ergebnisse zwischen den Leistungen in Paper-Pencil-Tests und denen beim realen Experimentieren nur gering miteinander korrelieren, sodass die Aussagekraft rein schriftlicher Tests nur sehr bedingt ist. Deshalb wurde hier eine experimentelle Aufgabensequenz für die Diagnose naturwissenschaftlicher Problemlösefähigkeit verwendet, die schriftliche Aufzeichnungen und reales Experimentieren miteinander verbindet und die quantitativ ausgewertet werden kann. Darüber hinaus findet mit dieser Aufgabensequenz eine individuelle Betrachtung der erbrachten Leistungen der Schülerinnen und Schüler in Einzelarbeit statt.

Da angenommen werden kann, dass der Problemlöseerfolg durch kognitive Fähigkeiten und nicht-kognitive Aspekte beeinflusst wird, wurden diese als Kovariablen mit erhoben. An der Untersuchung nahmen 275 Schülerinnen und Schüler der Jahrgangsstufen 5 bis 7 niedersächsischer Gymnasien teil. Es zeigten sich positive signifikante Zusammenhänge zwischen der naturwissenschaftlich-experimentellen Problemlösefähigkeit und den kognitiven Fähigkeiten, den motivationalen Aspekten sowie zur Beurteilung ausgewählter Persönlichkeitsmerkmale der Schülerinnen und Schüler durch die Lehrkräfte.

Schlüsselwörter: naturwissenschaftlich-experimentelles Problemlösen; kognitive Fähigkeiten; Schulnoten; Motivation

Measurement of experimental problem solving ability and its relationship to personality characteristics of pupils

Summary

The study presented here examines the relationship between experimental problem solving success and cognitive, motivational and certain personality traits that promote performance of pupils using an interactive science task sequence.

Regarding the measurement of experimental problem solving capacities, many studies have confirmed that students' performance in paper-pencil tests correlates only slightly with their performance in practical experimentation. Hence, the informative value of written tests as an exclusive research data base is limited.

Consequently, the study presented here used an experimental task sequence combining written records with practical experimentation in order to diagnose science problem solving ability via a quantitative analysis. In addition, individual examination of the pupils' achievements in individual work is conducted with the task sequence.

Since it can be assumed that the problem-solving success is influenced by cognitive abilities and non-cognitive aspects, these were also included as covariates.

A total of 275 pupils attending Germany "Gymnasiums" (secondary school) in grades 5 to 7 in Lower Saxony took part in the study.

Positive correlations were found between experimental problem solving ability and cognitive abilities, motivational aspects as well as teachers' assessment of selected personality traits of their pupils.

Key Words: experimental problem solving; cognitive abilities; marks; motivation

Einleitung

Problemlösekompetenz wird in fast allen Lern- und Leistungsbereichen als relevante Qualifikation angesehen. Sie gehört besonders in den MINT-Fächern zu den zu erwerbenden Kompetenzen im Unterricht. Im naturwissenschaftlichen Unterricht spielt dabei das experimentelle Problemlösen im Sinne der Erkenntnisgewinnung eine zentrale Rolle.

Mit der Zielsetzung des Aufbaus entsprechender Kompetenzen ist die Frage der Diagnostik experimenteller Problemlösekompetenz eng verbunden, um den Status festzustellen und entsprechende Fördermaßnahmen abzuleiten.

Für die Messung experimenteller Problemlösekompetenz wurden vielfältige Verfahren entwickelt und evaluiert. Die Verfahren reichen von Multiple-Choice-Tests über Realexperimente, Protokolle und Interviews bis hin zu Computersimulationen (vgl. Emden 2011; Gut 2012; Klos et al. 2008; Maiseyenko 2014; Nehring 2014; Rieß und Robin 2012; Schreiber 2012; Schreiber et al. 2014; Shavelson et al. 1999; Theyßen et al. 2013; Vorholzer et al. 2016). Insgesamt hat sich gezeigt, dass ein Messverfahren allein in der Regel nicht ausreicht, um die experimentelle Problemlösekompetenz von Schülerinnen und Schülern ausreichend abzubilden. So wurden z. T. nur geringe Korrelationen zwischen den Ergebnissen schriftlicher Tests und denen von Realexperimenten festgestellt (Hamman et al. 2008; Theyßen et al. 2013; Schreiber et al. 2014), sodass bis jetzt noch ungeklärt ist, welche Methoden valide den Kompetenzzustand erfassen (Maiseyenko 2014; Schreiber et al. 2014; Schecker et al. 2016).

Verschiedene mögliche Testverfahren wurden von Baur (2015) im Hinblick auf deren Eignung für die Schulpraxis analysiert und beschrieben. Die Datenerhebung mittels Videografie und anschließender Analyse liefert zwar authentische Daten, ist aber für die alltägliche Praxis zu zeitintensiv. Schülerprotokolle, welche während des Experimentierens von den Lernenden angefertigt werden und später von der Lehrkraft kriteriengeleitet ausgewertet werden, liefern zwar auch keine kongruente Abbildung der Realexperimentierphase, erscheint aber als einziges Verfahren für die Schulpraxis geeignet (Baur 2015, 32). Auch Emden (2011, 111) kommt zu der Aussage, dass die Eignung der Protokollmethode als alternative und valide Leistungsmessung unter Vorbehalt angenommen werden kann. Bei der Videografie realen Unterrichts werden in der Regel Schülergruppen beim Experimentieren aufgenommen (z. B. Emden 2011), sodass eine Individualdiagnostik nur sehr eingeschränkt möglich ist. Dies trifft dann auch auf die angefertigten zugehörigen Protokolle zu.

Für die Diagnostik naturwissenschaftlich-experimenteller Problemlösekompetenzen einzelner Schülerinnen und Schüler sind Tests notwendig, die das reale Experimentieren mit begleitenden prozessorientierten schriftlichen Aufzeichnungen in Einzelarbeit zusammen mit zielgerichteten Beobachtungen in der Schulpraxis ermöglichen.

Insgesamt reicht natürlich eine Einzelmessung zu einem Zeitpunkt nicht aus, sondern es müssen in zeitlicher Abfolge oder parallel mehrere Testverfahren eingesetzt werden, um generalisierbare Aussagen über die Leistungsfähigkeit der jeweiligen Schülerinnen und Schüler zu erlauben.

Ein weiterer wichtiger Aspekt bei der Diagnose experimenteller Fähigkeiten ist das notwendige fachinhaltliche Vorwissen für die Bearbeitung der Aufgaben. In der Regel können domänenspezifische Tests nur für bestimmte Jahrgangsstufen eingesetzt werden, wenn das erforderliche Fachwissen vorhanden ist.

In der vorliegenden Studie wurde eine naturwissenschaftlich-experimentelle Aufgabensequenz für die Diagnose naturwissenschaftlicher Problemlösefähigkeit verwendet, die schriftliche Aufzeichnungen (prozessorientierte Protokolle) und reales Experimentieren miteinander verbindet und in Einzelarbeit von den Schülerinnen und Schülern durchgeführt wird. In vorangegangenen Pilotstudien wurden die Schülerinnen und Schüler bei ihrer Aufgabenbearbeitung parallel videografiert. Es ergaben sich durch die Videografie keine zusätzlichen, fundamentalen Informationen, da die Schülerinnen und Schüler bei der Einzelarbeit in der Regel nicht sprechen. Es zeigte sich, dass die schriftlichen Aufzeichnungen in Kombination mit zielgerichteten, sehr kurzen Beobachtungen ausreichen, um zu beurteilen, ob die Schülerinnen und Schüler mit den Materialien in geeigneter Art und Weise umgehen können. Die Aufgabenstellungen sind klar formuliert und strukturiert und das erforderliche Experimentiermaterial ist überschaubar. Die Aufgabensequenz ist in mehreren Jahrgangsstufen einsetzbar und lässt sich zeitökonomisch durchführen sowie auswerten.

Da aus theoretischer Perspektive anzunehmen ist, dass auch für das naturwissenschaftlich-experimentelle Problemlösen kognitive Fähigkeiten als Prädiktoren und nicht-kognitive Aspekte als Moderatoren eine wichtige Rolle spielen (Heller, Perleth 2000), wurden diese als Kovariablen mit erhoben und auf korrelative Zusammenhänge untersucht. Diese Daten können auch im Hinblick auf die Validität der Aufgabensequenz verwendet werden.

Theoretischer Hintergrund

Experimentelle Problemlösekompetenz und deren Modellierung

Problemlösen lässt sich als „zielorientiertes Denken und Handeln in Situationen, für deren Bewältigung keine routinierten Vorgehensweisen verfügbar sind“ definieren (Klieme et al. 2001, 185). In ähnlicher Form wird das Problemlösen von Dörner (1976) als das Beseitigen von Barrieren zwischen dem Ausgangs- und dem Zielzustand beschrieben, welches stets mit kognitiven Anstrengungen einhergeht.

Problemlösen im naturwissenschaftlichen Unterricht lässt sich von anderen Unterrichtsfächern abgrenzen, wenn das Experimentieren im Fokus steht. Das experimentelle Problemlösen hat dabei einen besonders hohen Stellenwert, da dadurch der Prozess der Erkenntnisgewinnung in den Naturwissenschaften mit spezifischen Strategien durchlaufen wird (vgl. Emden 2011, 117). Wichtige Schritte sind in dem Prozess das Aufstellen von Hypothesen, das Planen und Durchführen von Experimenten, die Auswertung und die Dokumentation der Ergebnisse.

Das eigenständige experimentelle Problemlösen kann als relativ komplexer, kognitiver Problemlöseprozess verstanden werden, der eine besondere Herausforderung für Schülerinnen und Schüler darstellt (Hamman 2004; Hamman et al. 2008; Mayer et al. 2003).

Es gibt verschiedene Modelle, die versuchen, die experimentelle Problemlösefähigkeit in Teilkompetenzen auszudifferenzieren (vgl. Greiff 2012; Grube 2011; Hamman et al. 2008; Henke 2007; Klos 2008; Mayer 2007; Nehring 2014; Wellnitz und Mayer 2013; für eine Übersicht siehe Emden 2011). Häufig orientieren sich diese an dem von Klahr und Dunbar (1988) vorgestellten Modell *Scientific Discovery as Dual Search* (SDDS). Das SDDS-Modell beschreibt das Experimentieren als eine Suche in einem Hypothesen- und einem Experimentierraum sowie die Analyse von Evidenz. Insgesamt lassen sich auch bei unterschiedlicher Ausdifferenzierung in den einzelnen Modellen drei zentrale Prozessschritte extrahieren: Fragen, Vermutungen und Hypothesen generieren, Experimente planen und durchführen, Daten auswerten und interpretieren (vgl. auch Vorholzer et al. 2016, 26). Die Befunde zur Dimensionierung dieser drei Teilkompetenzen zu einer oder mehreren Dimensionen sind bisher inkonsistent (Vorholzer et al. 2016, 27). Dies scheint u. a. damit zusammenzuhängen, dass zwischen fachinhaltlichen und prozessbezogenen Kompetenzen nicht ausreichend differenziert wird bzw. die beiden Konstrukte nicht genügend voneinander abgegrenzt werden (z. B. Grube 2011; Hamman et al. 2008; Henke 2007; Vorholzer et al. 2016). Für eine valide Kompetenzmessung naturwissenschaftlich-experimentellen Problemlösens sollte fachinhaltliches Wissen kein schwierigkeiterzeugendes

Merkmal der Aufgabe darstellen (vgl. Vorholzer et al. 2016; Messick 1995). Dies bedeutet, dass ein Test benötigt wird, der vorrangig die prozessbezogenen Kompetenzen des Experimentierens im Sinne der Erkenntnisgewinnung als eigenständiges Konstrukt erfasst (Vorholzer et al. 2016, 26).

Problemlösekompetenz und deren Zusammenhang mit kognitiven und nicht-kognitiven Kovariablen

Ein entscheidendes Merkmal des Problemlösens ist das Überwinden von Barrieren, um von dem Ausgangszustand zu dem Zielzustand zu gelangen, was immer mit kognitiven Anstrengungen verbunden ist (Dörner 1976). Aus theoretischer Perspektive ist deshalb anzunehmen, dass es einen Zusammenhang zwischen der Problemlösekompetenz und den kognitiven Fähigkeiten des Problemlösers gibt (s. auch Nehring 2014, 96). Die empirische Befundlage zu dieser Relation ist jedoch divers, wenn zwischen analytischem und komplexem Problemlösen unterschieden wird (Greiff 2012; Leutner et al. 2005; Leutner et al. 2012; Süß 1996; Wirth et al. 2005) und hängt damit von der Art der Problemstellung ab. Positive Zusammenhänge zum analytischen Problemlösen scheinen bewiesen (Putz-Osterloh 1981; Wittmann und Süß 1999; Leutner 2002). Ebenso gibt es konsistente, positive Zusammenhänge zu Kompetenzen im Bereich der Erkenntnisgewinnung (Henke 2007; Kampa 2012; Klos 2008). Für experimentelle Problemlösefähigkeiten¹ zeigen sich in neueren Studien zum naturwissenschaftlich-experimentellen Problemlösen positive statistische Zusammenhänge zum anschauungsgebundenen Denken (Scherer 2014a, 2014b) und zu arithmetischen Grundkompetenzen (Kretzschmar et al. 2014). Heitmann (2012) identifizierte auch sprachliche Fähigkeiten als starken Prädiktor für naturwissenschaftliche Problemlösefähigkeiten. Dieses Ergebnis stimmt zum Teil mit den beschriebenen kognitiven Teilkompetenzen von Prenzel et al. (2001) überein, da dort unter anderem beschrieben wird, dass sprachliche Fähigkeiten beim „naturwissenschaftlichen Denken, Verstehen und Schlussfolgern“ (ebd., 209) von großer Bedeutung sind. Aufgrund dieser Zusammenhänge werden kognitive Grundfertigkeiten in Validierungsstudien im Hinblick auf die Konstruktvalidität mit aufgenommen (Kampa 2012).

Häufig werden bei Untersuchungen zum Problemlösen auch Schulnoten mit erhoben. Schulnoten bilden primär das Ausmaß des erreichten schulischen Erfolgs ab, gelten aber auch als indirektes Maß für intellektuelle Leistungsfähigkeit (Greiff 2012) und werden deshalb für die Validierung von IQ-Tests mit herangezogen (Stern, Hardy 2004; Heller, Perleth 2000). Schulnoten haben auch eine hohe prädiktive Validität für zukünftige Leistungsfähigkeit (Greiff 2013; Greiff und Fischer 2013; Heublein et al. 2012; Köller 2016; Meidinger 2014; Schuler 2000). Die Erhebung findet oft durch Selbstauskünfte der Schülerinnen und Schüler statt (z. B. Schuler 2000; Greiff und Fischer 2013), da sie so einfacher zu ermitteln sind. Trost (1975) berichtet über Studien mit hohen Übereinstimmungen zwischen selbstberichteten und tatsächlichen Noten. Untersuchungen zu Zusammenhängen zwischen Schulnoten und der Problemlösekompetenz zeigen, dass die Schulnoten in den Fächern Mathematik, Chemie und Deutsch bzw. die entsprechende Muttersprache von gutem prädiktivem Wert sind (vgl. Nehring 2014; Kretzschmar et al. 2014; Scherer 2014a, 2014b; Wüstenberg et al. 2014). Die Mathematiknote erscheint dabei als bester Prädiktor mit Korrelationen von durchschnittlich 0,30 und die Korrelationen mit den Naturwissenschaftsnoten liegen zwischen 0,26 und 0,31 (Schuler 2006). Nehring weist zwischen den Kompetenzen im Bereich der Erkenntnisgewinnung im Fach Chemie und den Schulnoten in den Fächern Biologie, Chemie, Physik, Mathematik und Deutsch mittlere Korrelationen von $0,22 \leq r \leq 0,46$ nach (Nehring 2014, 182).

Schulnoten spiegeln neben allgemeinen kognitiven Fähigkeiten auch einige nichtkognitive Kompetenzen wieder. Dazu gehören z. B. die Lernbereitschaft und die Leistungsmotivation (Trappmann et

¹ Diese lassen sich u. E. nach im Sinne der Pädagogischen Psychologie zwischen dem analytischen und dem komplexen Problemlösen einordnen. In der naturwissenschaftsdidaktischen Forschung wird diese Unterscheidung in der Regel so nicht thematisiert.

al. 2007; Rindermann und Oubaid 1999). Es ist bekannt, dass ein erheblicher Anteil (zwischen 25 % und 50 %) der Unterschiede zwischen Schülerinnen und Schülern bezüglich ihrer schulischen Leistungen durch motivationale Faktoren aufgeklärt wird (Helmke und Weinert 1997; zur Übersicht Wisniewski 2013). Auch die viel zitierte Kompetenzdefinition von Weinert (2001, 27) beinhaltet neben kognitiven Fähigkeiten² motivationale und volitionale Bereitschaften und Fähigkeiten. Diese Definition lässt sich auf die naturwissenschaftlich-experimentelle Problemlösekompetenz übertragen, so dass andere leistungsförderliche Persönlichkeitsmerkmale damit verknüpft sind. Dies wurde für den Bereich der Begabtenforschung im Fach Mathematik mehrfach nachgewiesen (Käpnick 2001; 2013). Gestützt wird diese Annahme auch durch die Untersuchungsergebnisse von Heitmann (2012) und Wüstenberg et al. (2014), in denen fachspezifische bzw. schulische Motivation als wichtige Indikatoren für Problemlöseleistungen im naturwissenschaftlichen Bereich identifiziert werden. Auch Nehring (2014) verweist auf eine Reihe von Studien, die positive Korrelationen nachweisen. Eine hohe Korrelation bedeutet zwar nicht eindeutig, dass zwei Variablen kausal miteinander zusammenhängen, sie liefert aber einen Hinweis darauf, dass dies der Fall sein könnte.

Insgesamt ist die Befundlage hinsichtlich der Zusammenhänge zwischen den Kompetenzen im Bereich der Erkenntnisgewinnung und anderen Kovariablen nicht eindeutig geklärt. Aufgrund der eingeschränkten Vergleichbarkeit bisheriger Studien durch verschiedene fachliche Kontexte und auch im Hinblick auf die diversen Beweislagen hinsichtlich der Zusammenhänge zwischen dem naturwissenschaftlich-experimentellem Problemlösen und anderen Kovariablen besteht weiterhin Forschungsbedarf in diesem Bereich (Nehring 2014, 36-39).

Forschungsfragen

Die vorliegende Studie dient der weiteren Erprobung eines Testinstruments zur Messung naturwissenschaftlich-experimenteller Problemlösefähigkeit und dessen Validierung im Hinblick auf die Konstruktvalidität. Zu diesem Zweck werden korrelative Zusammenhangsanalysen zwischen den Leistungen in der naturwissenschaftlich-experimentellen Aufgabensequenz und den kognitiven Fähigkeiten, Schulnoten, Fachbeliebtheiten und anderen motivationalen Aspekten sowie leistungsförderlichen Persönlichkeitsmerkmalen berechnet. Dabei werden sowohl Selbsteinschätzungen der Schülerinnen und Schüler als auch die Beurteilungen der entsprechenden Naturwissenschaftslehrkräfte einbezogen. Es werden Variablen berücksichtigt, die in der Diskussion um Kompetenzen im Bereich der Erkenntnisgewinnung eine Rolle spielen (vgl. Nehring 2014, 96). Derartige Zusammenhangsanalysen können auch für eine externe Validierung herangezogen werden (Hartig et al. 2012). Beim naturwissenschaftlich-experimentellen Problemlösen wird hier nicht in Teilkompetenzen ausdifferenziert, da der gesamte Prozess als ein Konstrukt betrachtet werden soll.

Die Arbeit schließt an zwei Pilotstudien im Rahmen von Masterarbeiten an, in denen untersucht wurde, inwiefern ein Zusammenhang zwischen mathematischer und naturwissenschaftlicher Leistungsfähigkeit besteht und in welcher Beziehung dieser Zusammenhang zu leistungsfördernden Persönlichkeitsmerkmalen bzw. zu den Noten in den Unterrichtsfächern Mathematik und Chemie steht. Für die Erfassung der naturwissenschaftlichen Kompetenzen wurde eine experimentelle Aufgabensequenz als Testinstrument entwickelt und erprobt (Giessel und Höner 2016). Das Testinstrument erwies sich dabei als geeignet für die individuelle Diagnose naturwissenschaftlicher Problemlösekompetenz und ist vorwissensneutral in den Jahrgangsstufen 4 bis 9 einsetzbar. Die Inhaltsvalidität wurde argumentativ aufgrund der Aufgabeninhalte in einer Expertenrunde bestätigt. Dasselbe Testinstrument wurde auch in dieser Studie verwendet. Die begleitenden Testinstrumente (Lehrer- und Schülerfragebögen) wurden in den vorangegangenen Pilotstudien erprobt und werden auch für diese Untersuchung verwendet.

Für die Untersuchung von Zusammenhängen ergeben sich folgende Fragestellungen:

² Die experimentellen Fertigkeiten verstanden als sensomotorische Ebene sollen hier nicht weiter betrachtet werden; experimentelle Fähigkeiten beziehen sich hier auf die intellektuelle Ebene.

Forschungsfrage 1a: Korrelieren die Schülerleistungen bei der naturwissenschaftlichen Aufgabensequenz mit denen der kognitiven Aufgabensequenz?

Forschungsfrage 1b: Korrelieren die Schülerleistungen bei der naturwissenschaftlichen Aufgabensequenz und den kognitiven Aufgabensequenzen mit den Schulnoten in den Fächern Chemie, Mathematik und Deutsch?

Forschungsfrage 2: Korrelieren die fachspezifische Motivation mit den Noten der Schülerinnen und Schüler und mit den Leistungen in der naturwissenschaftlichen Aufgabensequenz und den kognitiven Aufgabensequenzen?

Forschungsfrage 3: Korrelieren die Lehrerbeurteilungen ausgewählter, leistungsförderlicher Persönlichkeitsmerkmale der Schülerinnen und Schüler mit den Leistungen in der naturwissenschaftlichen Aufgabensequenz und den kognitiven Aufgabensequenzen?

Untersuchungsdesign

Das Untersuchungsdesign lehnt sich eng an die beiden Pilotstudien an. Insgesamt wurden jeweils eine naturwissenschaftlich-experimentelle Aufgabensequenz, drei kognitive Subtests des Kognitiven Fähigkeitstests von Heller und Perleth (2000) als kognitive Aufgabensequenz (KFT4-12+R), ein Schülerfragebogen sowie ein Lehrerfragebogen eingesetzt. Die naturwissenschaftliche Aufgabensequenz soll dabei die Fähigkeit des experimentellen Problemlösens im Bereich der Naturwissenschaften erfassen. Hierbei muss beachtet werden, dass es nur eine naturwissenschaftliche Aufgabensequenz gibt, die von allen Klassen und allen Jahrgängen unter gleichen Bedingungen bearbeitet wird, während die Anforderungen der kognitiven Aufgaben jahrgangsabhängig sind. Die Bearbeitung durch die Schülerinnen und Schüler erfolgt in Einzelarbeit.

Testinstrumente

Für die vorliegende Studie wurde eine kritische Sichtung des Forschungsstands zur Thematik der Messung naturwissenschaftlich-experimentellen Problemlösens vorgenommen.

Die naturwissenschaftlich-experimentelle Aufgabensequenz wurde an die Prozessschritte des SDDS-Modells angelehnt, sodass die wesentlichen Prozesse der Erkenntnisgewinnung in den Naturwissenschaften damit erfasst werden und darin neben der Expertenrunde ein weiterer Beitrag zur inhaltlichen Validität gesehen werden kann. Da sich für eine konvergente Validierung keine entsprechenden vorwissensneutralen Vergleichsinstrumente in mehreren Jahrgangsstufen einsetzen ließen, wurden hier andere Konstrukte und deren Operationalisierung herangezogen (vgl. Glug 2009), die aus theoretischer Sicht Einfluss auf das Testergebnis haben. Die Validierung erfolgt hier durch Vergleich des Leistungskonstrukts „Schulnote und kognitive Fähigkeiten“, des Motivationskonstrukts „Interesse“ sowie des Moderatorenkonstrukts „leistungsförderliche Persönlichkeitsmerkmale“³.

Die herangezogenen Testinstrumente (Lehrer- und Schülerfragebögen) wurden in zwei Pilotstudien erprobt und ergaben gute Reliabilitäten der Skalen. Die Validität des Kognitiven Fähigkeitstests nach Heller und Perleth (2000) ist gegeben.

Naturwissenschaftliche Aufgabensequenz

Dieses Testinstrument verbindet individuelle, schriftliche Aufzeichnungen der Schülerinnen und Schüler mit ihren experimentellen Handlungen und stellt von der Aufgabenstellung her eine Problemlöseaufgabe im Sinne des Prozesses der Erkenntnisgewinnung in den Naturwissenschaften dar. In dieser Aufgabensequenz geht es um die Härtebestimmung von Feststoffen nach Mohs, der für einen qualitativen Vergleich zweier Feststoffe die Methode des gegenseitigen Ritzens nutzt (Graubner 1980; Schumann 2002). Die Härte als Stoffeigenschaft wird im Anfangsunterricht Chemie in der Regel

³ Im Rahmen der Begabungsforschung werden diese auch als „begabungsstützende Persönlichkeitsqualitäten“ bezeichnet, denen eine Hilfs- und Stützfunktion zur Intelligenz zugeschrieben wird (Käpnick 2013).

nicht ausdrücklich behandelt (Niedersächsisches Kultusministerium 2007, 2012, 2013). Es ist deshalb davon auszugehen, dass Schülerinnen und Schüler zwar Alltagsvorstellungen dazu haben, welche Stoffe eher hart sind, aber keine Vorkenntnisse im Sinne der Mohs'schen Härteskala und der experimentellen Bearbeitung, wie sie in der Aufgabensequenz gefordert ist.

Bei der naturwissenschaftlichen Aufgabensequenz sollen verschiedene Feststoffe bzgl. ihrer Härte untersucht und in eine Rangfolge gebracht werden. Es gibt zwei Arbeitsblätter mit insgesamt fünf zu bearbeitenden Teilaufgaben. Für die zwei nacheinander zu bearbeitenden, aufeinander aufbauenden experimentellen Problemstellungen gibt es je eine Experimentierbox mit verschiedenen Materialien (vgl. Abb. 1 und 2).

Das prozessorientierte Versuchsprotokoll ist vorstrukturiert und folgende Prozessmerkmale sind berücksichtigt: Vermutungen generieren und begründen (s. Abb. 1 und 2: 1a, 2a), Experimente planen, durchführen, dokumentieren und Vermutungen reflektieren (1b, 2b). Aufgabe 2 baut auf die Aufgabe 1 auf, da das erhaltene Ergebnis aus 1 dort berücksichtigt werden soll. Mit der Bearbeitung des Aufgabenteils 2c soll ermittelt werden, ob die Schülerinnen und Schüler erkennen, dass für ein zielgerichtetes Experimentieren nur wenige Versuchsschritte notwendig sind.

Aufgabe 1

Zu vergleichen sind die beiden Feststoffe Eisen und Marmor. Es soll festgestellt werden, welcher Feststoff härter ist.

Merke: Ein härterer Feststoff ritzt immer einen weicheren.

- a) Stelle eine begründete Vermutung auf, welcher der beiden Feststoffe härter ist!
- b) Versuche, deine Vermutung in einem Experiment zu bestätigen. Beschreibe genau, wie du vorgegangen bist! Kannst du deine Vermutung bestätigen?



Abb. 1 Aufgabe 1 und Experimentierbox 1

Die Experimentierbox für die erste Aufgabe enthält folgende Materialien: einen Spatel, ein Eisenblech, ein Stück Marmor, ein Aluminiumblech, Schleifpapier sowie den Hinweis, dass man auch seinen Fingernagel benutzen kann (Abb. 1).

Für diese erste Aufgabe haben die Schülerinnen und Schüler 10 Minuten Zeit.

Aufgabe 2

Zu vergleichen sind nun die vier Feststoffe Eisen, Marmor, Quarz und Gips. Es sind keine weiteren Materialien zugelassen.

- a) Stelle eine begründete Vermutung auf, wie die Reihenfolge der Feststoffe nach zunehmender Härte aussieht.
- b) Versuche deine Vermutung mit einem Experiment zu bestätigen. Jeder Vergleich von zwei Feststoffen untereinander zählt dabei als ein Untersuchungsschritt. Beschreibe genau, wie du vorgegangen bist, und notiere deine Untersuchungsschritte! Kannst du deine Vermutung bestätigen?
- c) Wie kannst du in möglichst wenig Schritten eine eindeutige Reihenfolge der Feststoffe bezüglich der Härte festlegen, wenn nur dein Ergebnis aus Aufgabe 1 bekannt ist?

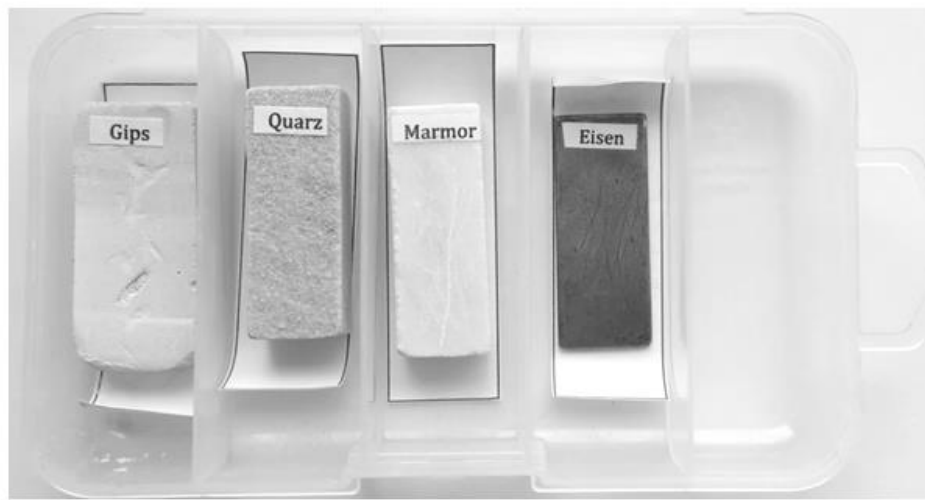


Abb. 2 Aufgabe 2 und Experimentierbox 2

Die Experimentierbox für die zweite Aufgabe enthält folgende Materialien: ein Eisenblech, jeweils ein Stück Marmor, Quarz und Gips (Abb. 2).

Für die Bearbeitung dieser zweiten Aufgabe haben die Schülerinnen und Schüler 15 Minuten Zeit.

Die Bewertung der Aufgaben erfolgt rein quantitativ anhand eines Auswertungsmanuals, in dem die zu bewertenden Aspekte aufgelistet sind, sodass vorher festgelegte Kriterien vorliegen (Tab. 1). Diese wurden auch schon in den beiden Pilotstudien verwendet⁴ und haben sich als geeignet hinsichtlich der Auswertungsobjektivität erwiesen⁵ (Müller und Büttner 1994; Wirtz und Caspar 2007). Die Punktevergabe orientiert sich dabei an den Prozessschritten der Erkenntnisgewinnung in den Naturwissenschaften. Ausgehend von einer begründeten Vermutung soll die experimentelle Durchführung sinnvoll (Durchführung, Beobachtung und Schlussfolgerung sowie Rückbezug zur Hypothese) beschrieben werden, damit sich die Vorgehensweise nachvollziehen lässt. Bei Aufgabe 2a wird ein Punkt für die Berücksichtigung des Ergebnisses aus Aufgabe 1 vergeben. Dies wird in der Aufgabenstellung

⁴ Die Bewertung orientiert sich an den Bewertungsempfehlungen zum Indikator-Aufgaben-Test von Käpnick (2001, 180ff). Die Punkteverteilung ist so gewählt, dass die Punkte innerhalb der Aufgaben ausgewogen verteilt sind.

⁵ Die Interrater-Reliabilität (zwei Rater) wurde als Intraklassen-Korrelationskoeffizient berechnet (two-way random) und ergab einen Wert von 0,976, was einer guten Beurteilerübereinstimmung entspricht.

nicht gefordert, sodass darin eine eigenständige Leistung der Schülerinnen und Schüler zu sehen ist, inwiefern sie selbstständig einen Transfer von Wissen auf neue Aufgaben durchführen.

Bei der Bewertung ist zu beachten, dass es keinen Punktabzug gibt, wenn die Lösung fachlich nicht korrekt ist, also z. B. Marmor härter als Eisen eingestuft wird. Vielmehr geht es darum, die in sich konsistente Argumentation und Dokumentation zu bewerten.

Tab. 1 Auswertungsmanual - Punktevergabe (Übersicht) für die naturwissenschaftliche Aufgabensequenz

Aufgabe	Erwartungshorizont	Punktzahl
1a	<ul style="list-style-type: none"> • Angeben einer Vermutung • Begründung für die Vermutung 	1 1
1b	<ul style="list-style-type: none"> • Sinnvolle Beschreibung der experimentellen Vorgehensweise (Untersuchungsschritte) • Erwähnung der Methode des gegenseitigen Ritzens • Nennung eines Ergebnisses, einer Deutung • Erneutes Eingehen auf die Vermutung 	1 1 1 1
1 gesamt		6
2a	<ul style="list-style-type: none"> • Aufstellen einer vermuteten Reihenfolge • Berücksichtigung des individuellen Ergebnisses aus Aufgabe 1 • Begründung für die vermutete Reihenfolge 	1 1 1
2b	<ul style="list-style-type: none"> • Sinnvolle Beschreibung der experimentellen Vorgehensweise • Benennung der Untersuchungsschritte • Nennung eines Ergebnisses, einer Deutung (Reihenfolge) • Erneutes Eingehen auf die Vermutung 	1 1 1 1
2c	<ul style="list-style-type: none"> • Erläuterung über das Weglassen des Vergleichs von Eisen und Marmor 	1
2 gesamt		8
gesamt		14

Kognitive Aufgabensequenz

Zur Erfassung der kognitiven Fähigkeiten wurde eine verkürzte Form des „KFT 4-12+R“ von Heller und Perleth (2000) verwendet, da die Bearbeitung aller Aufgaben aus zeitlichen Gründen nicht möglich war. Auch in anderen Studien werden verkürzte Formen des KFT verwendet (z. B. Emden 2011; Maisyenko 2014; Küsting 2007; Schulz 2011; Seidel et al. 2003). Der KFT besteht insgesamt aus drei Teilen - einem verbalen, einem quantitativen und einem nonverbalen Teil mit jeweils drei Subskalen. Für die vorliegende Studie wurde aus jedem Teil nur eine Subskala verwendet (Tab. 2).

Tab. 2 Aufgabenauswahl KFT (Heller und Perleth 2000)

Testteil KFT	Subtest KFT	Zeitbedarf (in Min:Sek)	Erfasste Fähigkeiten
Verbaler Teil	V3 Wortanalogien	09:30	Sprachgebundenes Denken, Wortschatz, Sprachverständnis
Quantitativer Teil	Q2 Zahlenreihen	11:00	Zahlengebundenes Denken, logisches Denken

Für die Auswertung der kognitiven Subtests wurden die Musterlösungen des KFT verwendet (vgl. Heller und Perleth 2000, 123ff).

Schüler- und Lehrerfragebogen

Der Schülerfragebogen (Giessel und Höner 2016) beinhaltet Angaben zum Geschlecht und zu den Schulnoten in den Unterrichtsfächern Chemie, Mathematik und Deutsch. In den Fächern Mathematik und Deutsch wurden die Noten nur durch Selbstauskünfte der Schülerinnen und Schüler erhoben, die Chemienoten wurden dagegen zusätzlich durch die Lehrkräfte angegeben, sodass diese Information zweifach vorlag⁶.

Darüber hinaus gibt es Fragen zu den Fachbeliebtheiten der Fächer Chemie, Biologie, Mathematik, Deutsch und Musik sowie zu motivationalen Aspekten drei verschiedener Unterrichtsfächer auf einer 5-stufigen Likertskala (Abb. 3), um das fachspezifische Interesse und das Fähigkeitsselbstkonzept der Schülerinnen und Schüler durch Selbstauskunft zu erfragen.

Bewerte die folgenden Aussagen! Die Ziffern 1 – 5 stehen nun für Ablehnung bzw. Zustimmung einer Aussage.					
	1 „trifft nicht zu“	2 „trifft eher nicht zu“	3 „weder noch“	4 „trifft eher zu“	5 „trifft zu“
Ich bin sehr am Fach Chemie interessiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich bin sehr am Fach Mathematik interessiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich bin sehr am Fach Deutsch interessiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Für mich ist Chemie eines der leichteren Fächer.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Für den Chemieunterricht mache ich meine Hausaufgaben besonders gerne.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich kann mir vorstellen, freiwillig an einer Chemie-AG teil zu nehmen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Im Fach Chemie fällt es mir sehr leicht, mir Dinge zu merken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Im Fach Mathematik fällt es mir sehr leicht, mir Dinge zu merken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mir fällt es sehr leicht, mir Vokabeln zu merken.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich beschäftige mich auch in meiner Freizeit gern mit Chemie.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abb. 3 Beispielfragen aus dem Schülerfragebogen mit 5er-Likert-Skala

⁶ Die durch die Schülerinnen und Schüler angegebenen Chemienoten und die Angaben der Lehrkräfte korrelieren hoch signifikant miteinander ($r = 0,872$; $p < 0,001$) und zeigen damit eine statistische Übereinstimmung. Deshalb wird auch für die Mathematik- und die Deutschnoten aus den Schülerangaben angenommen, dass sie den tatsächlichen Noten entsprechen.

Das Hauptaugenmerk des Schülerfragebogens richtet sich demnach, trotz jeweils zwei Fragen zum Mathematik- und Deutschunterricht, überwiegend auf die persönlichen Interessen und Fähigkeitseinschätzungen zum Unterrichtsfach Chemie.

Der Lehrerfragebogen wurde im Wesentlichen aus vorangegangenen Untersuchungen übernommen (Giessel und Höner 2016). Ziel des Lehrerfragebogens ist die Erhebung verschiedener Persönlichkeitsmerkmale der Schülerinnen und Schüler durch Beurteilungen der jeweiligen Naturwissenschaftslehrkräfte. Die abgefragten Persönlichkeitsmerkmale (Abb. 4) haben sich als leistungsförderliche Merkmale erwiesen (Käpnick 2001; Käpnick 2013; Giessel und Höner 2016) bzw. wirken sich positiv auf schulische Leistungen aus (Wiesniewski 2013).

Wie stark sind die folgenden Merkmale bei dem Schüler/der Schülerin ausgeprägt?						
	gering	eher gering	mittel	eher hoch	hoch	nicht zu beurteilen
Selbstständigkeit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Konzentrationsfähigkeit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Durchhaltevermögen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Anstrengungsbereitschaft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eigeninitiative	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Interesse an naturwissenschaftlichen Phänomenen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ausdrucksvolle Sprache	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kreativität	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abb. 4 Beispielfragen aus dem Lehrerfragebogen

Die Lehrkräfte beurteilen die Schülerinnen und Schüler auf einer 5-stufigen Likertskala (mit der Möglichkeit „nicht zu beurteilen“).

Stichprobe

An der Untersuchung nahmen insgesamt 275 Schülerinnen und Schüler der Jahrgangsstufen 5 bis 7 von drei niedersächsischen Gymnasien teil. Es handelt sich um eine Gelegenheitsstichprobe, die entsprechend der Tabelle 3 zusammengesetzt ist.

Tab. 3 Zusammensetzung der Stichprobe

Jahrgang	Anzahl	%	Anzahl Mädchen	Anzahl Jungen
5	56	20,4	22	34
6	80	29,1	40	39
7	139	50,5	79	60
Summe	275	100	141	133

Durchführung

Die Schülerinnen und Schüler wurden nicht auf den Test vorbereitet. Die Testung wurde standardisiert durchgeführt und fand jeweils in einer Doppelstunde von 90 Minuten im Rahmen des regulären Unterrichts an den Schulen statt. Die Teilnahme war für alle Schülerinnen und Schüler verpflichtend. Die Bearbeitung der Testinstrumente erfolgte in Einzelarbeit in der folgenden Reihenfolge: naturwis-

senschaftliche Aufgabensequenz (Aufgabe 1, erst nach deren Abschluss Aufgabe 2), kognitive Aufgabensequenzen (V3, Q2, N3), Schülerfragebogen. Die Lehrerfragebögen wurden von den verschiedenen Naturwissenschaftslehrkräften der jeweiligen Klassen zeitlich parallel ausgefüllt, sodass den Lehrkräften die Ergebnisse der Tests nicht bekannt waren.

Statistische Analysen

Die Daten wurden mithilfe der Statistiksoftware IBM SPSS Statistics 23 bzw. 24 eingegeben und ausgewertet. Die Daten der Aufgabensequenzen und der Fragebögen wurden dabei einheitlich in eine Richtung codiert, sodass hohe Zahlenwerte einer guten Leistung bzw. einer hohen Ausprägung entsprechen.

Der Kolmogorov-Smirnow-Anpassungstest zeigte für die meisten Variablen keine Normalverteilung. Der korrelative Zusammenhang zweier Variablen wurde deshalb nach Spearman (beidseitig) berechnet. Für die Untersuchung von Jahrgangsunterschieden wurde der Kruskal-Wallis-Test verwendet. Für die Skalen des Lehrerfragebogens und des Schülerfragebogens wurden Reliabilitäten berechnet. Für die Items des Schülerfragebogens wurde eine Faktorenanalyse (Hauptkomponenten) durchgeführt. Die Höhe der jeweiligen Korrelationskoeffizienten wird in Anlehnung an Cohen (1988) interpretiert.

Ergebnisse

Die Schülerinnen und Schüler erreichen bei der naturwissenschaftlichen Aufgabensequenz einen Mittelwert von $9,4 \pm 2,29$ von maximal 14 zu erreichenden Punkten. Dies ergibt einen Schwierigkeitsindex von 0,67 und liegt damit für die Gymnasialschülerinnen und –schüler etwas unter der mittleren Itemschwierigkeit von 0,5 (Maier 2015).

Da bei allen Auswertungen keine signifikanten Jahrgangsunterschiede festzustellen waren, werden die Ergebnisse jahrgangsübergreifend dargestellt. Auch im Hinblick auf das Geschlecht gab es keine signifikanten Differenzen.

Zusammenhänge zwischen den Leistungen in der naturwissenschaftlichen Aufgabensequenz und der kognitiven Aufgabensequenz sowie den Schulnoten

Eine Untersuchung des korrelativen Zusammenhangs zwischen den erreichten Punktzahlen in den aufeinander aufbauenden naturwissenschaftlichen Aufgaben zeigt, dass es einen signifikanten Zusammenhang ($r = 0,418$; $p < 0,001$) gibt. Die beiden Aufgaben werden bei den folgenden Auswertungen als eine Einheit betrachtet, da die zweite Aufgabe auf die erste aufbaut und sehr ähnliche Problemlöseprozesse nötig sind. Es werden die Gesamtpunkte verwendet.

Die erreichten Punktzahlen in den drei kognitiven Aufgabenteilen korrelieren ebenfalls signifikant miteinander (Abb. 5). Die Korrelationskoeffizienten liegen dabei im mittleren Bereich. Die drei Subtests werden im Weiteren trotzdem getrennt betrachtet, da sie unabhängig voneinander bearbeitet werden können und sie unterschiedlichen Bereichen der kognitiven Fähigkeiten entsprechen (s. Tab. 2).

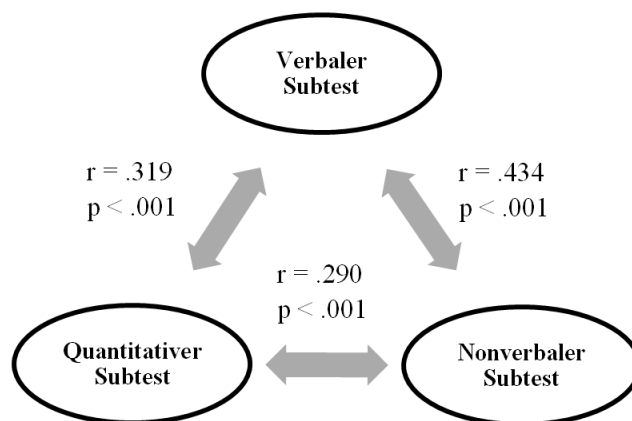


Abb. 5 Korrelationen zwischen den kognitiven Aufgaben

Werden die Zusammenhänge zwischen den erreichten Punkten der naturwissenschaftlichen Aufgabensequenz (NW Aufgaben) untersucht, so ergeben sich die in der Tabelle 4 angegebenen positiven signifikanten Zusammenhänge mit den drei kognitiven Aufgabenteilen.

Tab. 4 Korrelationen zwischen den erreichten Punkten in den Aufgabensequenzen (N = 275)

			Verbale Aufgabe V3	Quantitative Aufgabe Q2	Nonverbale Aufgabe N3
Spearman-	NW	Korrelationskoeffizient	0,327	0,176	0,228
Rho	Aufgaben	Sig. (2-seitig)	0,000	0,003	0,000

Die Korrelationskoeffizienten liegen nur im mittleren Bereich für den verbalen Aufgabenteil und im geringen Bereich für die anderen beiden Aufgabenteile. Dies weist, darauf hin, dass es einen gewissen Zusammenhang zwischen den Leistungen in allen vier Bereichen gibt.

Beantwortung der Forschungsfrage 1a: Die Schülerleistungen bei der naturwissenschaftlichen Aufgabensequenz korrelieren (positiv) signifikant mit denen der kognitiven Aufgabensequenz.

Zusätzlich wurden die Zusammenhänge zwischen den Noten in den Unterrichtsfächern Chemie, Mathematik und Deutsch mit den Testleistungen überprüft. Es ergeben sich ausschließlich signifikante positive Zusammenhänge mit den Testleistungen in der naturwissenschaftlichen Aufgabensequenz und mit den Ergebnissen der kognitiven Aufgaben (Tab. 5).

Tab. 5 Korrelationen der naturwissenschaftlichen Aufgabensequenz und der kognitiven Aufgaben mit den Noten (N = 275)

			Note Chemie	Note Mathematik	Note Deutsch
Spearman-Rho	Verbale	Korrelationskoeffizient	0,342	0,345	0,368
	Aufgabe V3	Sig. (2-seitig)	0,000	0,000	0,000
	Quantitative	Korrelationskoeffizient	0,255	0,300	0,189
	Aufgabe Q2	Sig. (2-seitig)	0,000	0,000	0,002
	Nonverbale	Korrelationskoeffizient	0,243	0,354	0,284
	Aufgabe N3	Sig. (2-seitig)	0,000	0,000	0,000
	NW Aufgaben	Korrelationskoeffizient	0,313	0,407	0,299
		Sig. (2-seitig)	0,000	0,001	0,001

Die Korrelationskoeffizienten liegen bis auf drei Ausnahmen im mittleren Bereich bei ca. 0,3 oder etwas höher. Die Mathematiknote ($r > 0,3$) erweist sich dabei als bester Indikator für die gezeigten Leistungen. Die Deutschnote weist im Zusammenhang mit dem quantitativen Testteil Q2 die schwächste Korrelation auf und erwartungsgemäß die höchste mit dem verbalen Testteil V3. Tendenziell haben Schülerinnen und Schüler mit guten Noten in den drei Fächern bessere Leistungen in allen vier Tests erbracht.

Beantwortung der Forschungsfrage 1b: Die Schülerleistungen bei der naturwissenschaftlichen Aufgabensequenz und den kognitiven Aufgabensequenzen korrelieren (positiv) signifikant mit den Noten in den Fächern Chemie, Mathematik und Deutsch.

Die drei Fachnoten korrelieren untereinander hoch mit $0,514 < r < 0,553$ miteinander. Dieses Ergebnis deutet darauf hin, dass es sich prinzipiell um insgesamt bessere (schlechtere) Schülerinnen und Schüler handelt.

Zusammenhänge zwischen der Motivation, den Noten und den Leistungen in den Aufgaben

Die verschiedenen Items des Schülerfragebogens weisen untereinander eine Vielzahl signifikanter Korrelationen für die Items mit den Schwerpunkten Mathematik und Chemie und den zugehörigen Fachbeliebtheiten auf. Die Reliabilitätsanalyse ergibt jeweils einen Cronbachs α von $> 0,83$ und Trennschärfen zwischen 0,56 und 0,75. Bei der anschließenden Faktorenanalyse lassen sich zwei Faktoren identifizieren (Tab. 6). Aufgrund der den beiden Faktoren jeweils zugeordneten Items wurden die Faktoren mit „Motivation Chemie“ und „Motivation Mathematik“ bezeichnet. Es wurde ein Wert von 0,81 für das Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin ermittelt sowie eine erklärte Gesamtvarianz von 64,88 %. Als Extraktionsmethode wurde die „Hauptkomponentenanalyse“ und als Rotationsmethode die „Varimax mit Kaiser-Normalisierung“ gewählt. Beide Faktoren wurden als neue Variable gespeichert und für weitere Auswertungen verwendet.

Tab. 6 Faktorenanalyse (Hauptkomponenten) Schülerfragebogen (Rotierte Komponentenmatrix)

Bezeichnung	Items	Komponente	
		1	2
Motivation Chemie	Ich bin sehr am Fach Chemie interessiert.	0,845	0,202
	Fachbeliebtheit Chemie	0,842	0,097
	Für mich ist Chemie eines der leichteren Fächer.	0,764	-0,090
	Im Fach Chemie fällt es mir sehr leicht, mir Dinge zu merken.	0,674	0,059
	Für den Chemieunterricht mache ich meine Hausaufgaben besonders gerne.	0,681	0,076
	Ich kann mir vorstellen, freiwillig an einer Chemie-AG teilzunehmen.	0,688	0,138
	Ich beschäftige mich auch in meiner Freizeit gern mit Chemie.	0,664	0,118
Motivation Mathematik	Fachbeliebtheit Mathematik	0,123	0,934
	Ich bin sehr am Fach Mathematik interessiert	0,141	0,928
	Im Fach Mathematik fällt es mir sehr leicht, mir Dinge zu merken.	0,067	0,869

Die beiden Faktoren „Motivation Chemie“ und „Motivation Mathematik“ unterscheiden sich zum Teil in Bezug auf die Korrelationen zu den erbrachten Leistungen in den Aufgabensequenzen sowie zu den

Noten (Tab. 7). Aus Tabelle 7 geht hervor, dass die beiden Faktoren „Motivation Chemie“ und „Motivation Mathematik“ zwar signifikante Zusammenhänge zu den Leistungen in der naturwissenschaftlichen Aufgabensequenz aufweisen, diese aber hinsichtlich der Korrelationskoeffizienten gering sind. Dies trifft auch auf die Zusammenhänge mit den drei kognitiven Aufgabenteilen zu. Der Zusammenhang zwischen der „Motivation Chemie“ und der verbalen Aufgabe V3 ist dabei nicht signifikant.

Schülerinnen und Schüler, die einen hohen (niedrigen) Wert in den Faktoren „Motivation Chemie“ oder „Motivation Mathe“ aufweisen, erzielen demnach tendenziell häufiger hohe (niedrige) Punkte bei der naturwissenschaftlichen Aufgabensequenz und den kognitiven Subtests.

Tab. 7 Zusammenhänge mit den Motivationsfaktoren (jahrgangsübergreifend)

			Motivation Chemie	Motivation Mathematik
Spearman-Rho	NW Aufgaben	Korrelationskoeffizient	0,147	0,190
		Sig. (2-seitig)	0,018	0,003
		N	261	261
Verbale Aufgabe V3	Verbale Aufgabe V3	Korrelationskoeffizient	0,098	0,155
		Sig. (2-seitig)	0,117	0,013
		N	259	259
Quantitative Aufgabe Q2	Quantitative Aufgabe Q2	Korrelationskoeffizient	0,149	0,283
		Sig. (2-seitig)	0,016	0,000
		N	261	261
Nonverbale Aufgabe N3	Nonverbale Aufgabe N3	Korrelationskoeffizient	0,142	0,282
		Sig. (2-seitig)	0,021	0,000
		N	261	261
Note Deutsch	Note Deutsch	Korrelationskoeffizient	0,117	0,251
		Sig. (2-seitig)	0,063	0,000
		N	254	254
Note Chemie	Note Chemie	Korrelationskoeffizient	0,491	0,174
		Sig. (2-seitig)	0,000	0,007
		N	239	255
Note Mathematik	Note Mathematik	Korrelationskoeffizient	0,106	0,702
		Sig. (2-seitig)	0,09	0,000
		N	255	255

Im Hinblick auf die Noten in Chemie und Mathematik ergeben sich jeweils hohe Korrelationen zum jeweiligen Motivationsfaktor.

Beantwortung der Forschungsfrage 2: Die fachspezifische Motivation korreliert hoch mit den jeweiligen Fachnoten, aber nur schwach mit den Leistungen in der naturwissenschaftlichen und den kognitiven Aufgabensequenzen.

Zusammenhänge zwischen den Persönlichkeitsmerkmalen und den Leistungen in den Aufgaben

Bei den Korrelationsanalysen des Lehrerfragebogens zeigte sich, dass alle Items (Abb. 5) zu den beurteilten Persönlichkeitsmerkmalen durch die Lehrkräfte signifikant positiv miteinander korrelieren

($r = 0,46$ bis $0,835$; $p < 0,001$). Die Reliabilitätsanalyse ergibt einen Cronbachs α von $0,923$ mit Trennschärfen zwischen $0,651$ und $0,802$. Deshalb wurde eine neue Variable „Summe Lehrerbeurteilungen“ als Summenscore berechnet und für die weitere Auswertung verwendet.

Der Score korreliert in mittlerer Höhe mit den erreichten Punkten in der naturwissenschaftlichen Aufgabensequenz ($r = 0,347$; $p < 0,001$) und etwas geringer aber signifikant auch mit den erreichten Punkten in der verbalen ($r = 0,263$; $p < 0,001$) und nonverbalen ($r = 0,220$; $p < 0,001$) Aufgabe, aber nicht mit der quantitativen Aufgabe des KFT. Insgesamt kann deshalb ein positiver Zusammenhang zwischen den von den Lehrkräften eingeschätzten leistungsförderlichen Persönlichkeitsmerkmalen der Schülerinnen und Schüler und den erbrachten Leistungen in fast allen getesteten Bereichen festgestellt werden.

Beantwortung der Forschungsfrage 3: Die Lehrerbeurteilungen leistungsförderlicher Persönlichkeitsmerkmale der Schülerinnen und Schüler korrelieren signifikant mit den Leistungen in den naturwissenschaftlichen und zwei der drei kognitiven Aufgabensequenzen.

Weiterhin gibt es signifikante Zusammenhänge zwischen dem Summenscore der Lehrerbeurteilungen und den beiden Faktoren „Motivation Chemie“ ($r = 0,260$; $p < 0,001$) und „Motivation Mathematik“ ($r = 0,361$; $p < 0,001$).

Diskussion und Ausblick

In der vorliegenden Studie werden korrelative Zusammenhänge zwischen der naturwissenschaftlich-experimentellen Problemlösefähigkeit und kognitiven, motivationalen und leistungsfördernden Persönlichkeitsmerkmalen von Schülerinnen und Schülern beschrieben. Die Diagnose experimenteller Problemlösefähigkeit erfolgte mit einer naturwissenschaftlichen Aufgabensequenz, die reales Experimentieren und schriftliche Aufzeichnungen miteinander verbindet und in Einzelarbeit von den Schülerinnen und Schülern bearbeitet wird. Der Test ist einfach auszuwerten und lässt sich vorwissensneutral in mehreren Jahrgangsstufen einsetzen.

Um die kognitiven Teilkompetenzen zu erheben, wurden drei Subtests des Kognitiven Fähigkeitstests (KFT) von Heller und Perleth (2000) ausgewählt. Außerdem wurden verschiedene leistungsförderliche Persönlichkeitsmerkmale (Lehrerfragebogen) und motivationale Aspekte sowie Schulnoten (Schülerfragebogen) erfasst, um sie mit den kognitiven Fähigkeiten bzw. mit den Problemlösefähigkeiten auf korrelative Zusammenhänge hin zu überprüfen, was auch der Validierung dient.

In der Auswertung der Ergebnisse ($N = 275$) zeigten sich nur schwache positive signifikante Zusammenhänge zwischen den erreichten Punkten in der naturwissenschaftlichen Aufgabensequenz und den erreichten Punkten in allen drei kognitiven Subtests. Dabei liegt der Korrelationskoeffizient für den verbalen Subtest mit $r = 0,327$ am höchsten. D. h., es lässt sich nur die Tendenz bestätigen, dass höhere kognitive Fähigkeiten mit besseren Leistungen beim naturwissenschaftlichen Problemlösen einhergehen. Dies passt auch zu den Ergebnissen von Nehring (2014, 176), der Zusammenhänge zwischen den Kompetenzen im Bereich der Erkenntnisgewinnung und den Kovariablen Intelligenz⁷ ($r = 0,36$) und Leseverständnis ($r = 0,19$) feststellt. Weitere Parallelen können zu den Ergebnissen von Wirth et al. (2005) hergestellt werden, die einen positiven Zusammenhang zwischen naturwissenschaftlichem Problemlösen und kognitiven Fähigkeiten festgestellt haben. Ebenso konnten positive Zusammenhänge zu den kognitiven Fähigkeiten im verbalen Bereich ermittelt werden, welches Ähnlichkeiten mit den Ergebnissen von Prenzel et al. (2001) und Heitmann (2012) aufweist.

Betrachtet man die signifikanten Korrelationen zwischen den erreichten Punktzahlen in den Aufgabensequenzen und den hier berücksichtigten Schulnoten, so kann geschlossen werden, dass die Noten in den Fächern Mathematik, Chemie und Deutsch gute Indikatoren für mögliche naturwissenschaftliche Leistungen sind, wie es auch viele andere Studien (vgl. z. B. Schuler 2006; Heublein et al. 2012; Greiff 2013; Meidinger 2014) gezeigt haben. Die Mathematiknote erwies sich dabei als bester

⁷ Die Intelligenz wurde nur mit dem N2-Subtest (Figurenanalogien) des KFT ermittelt.

Indikator für die erbrachte Leistung. Dies lässt sich so interpretieren, dass es sich insgesamt um bessere Schülerinnen und Schüler handelt und passt zu den Ausführungen von Köller (2016) im Hinblick auf die Ergebnisse verschiedener Schulstudien. Die Korrelationskoeffizienten sind dabei nicht sehr hoch ($0,189 \leq r \leq 0,407$), liegen aber im Bereich der Ergebnisse anderer Studien (vgl. Nehring 2014, Schuler 2006). Ähnliche korrelative Zusammenhänge zwischen arithmetischen Grundkompetenzen und der Note im Unterrichtsfach Mathematik mit Problemlösekompetenzen stellen beispielsweise Kretzschmar et al. (2014) fest. Bei Scherer (2014a, 2014b) konnten zudem statistische Zusammenhänge zwischen der Problemlösekompetenz und der Note in Chemie bzw. mit kognitiven Fähigkeiten im figural-bildhaften Bereich ermittelt werden.

Die positiven signifikanten Korrelationen (Tab. 7) zwischen der Problemlösekompetenz und den kognitiven Fähigkeiten mit den beiden Faktoren „Motivation Chemie“ und „Motivation Mathematik“ fallen nur schwach aus, zeigen aber dass motivationale Aspekte mit den erbrachten Leistungen zusammenhängen. Die Untersuchungsergebnisse von Heitmann (2012) können hier als Parallele hinsichtlich der Zusammenhänge zwischen Problemlösekompetenz und kognitiven Fähigkeiten, fachspezifischer Motivation und sprachlichen Aspekten angeführt werden. Heitmann (2012) identifizierte „kognitive Fähigkeiten, fachbezogene Motivation und (Mutter-)Sprache als stärkste Prädiktoren“ (Heitmann 2012, 154) für naturwissenschaftliche Problemlösefähigkeiten. Des Weiteren wurden in Untersuchungen von Wüstenberg et al. (2014) Komponenten der fluiden Intelligenz, der naturwissenschaftlichen Erkenntnisgewinnung (Scientific Reasoning) und der schulischen Motivation (Learning Orientation) als wichtige Prädiktoren für Problemlöseleistungen identifiziert. Bei Nehring (2014) ergeben sich mittlere positive Zusammenhänge zwischen den Kompetenzen im Bereich der Erkenntnisgewinnung und Kovariablen aus dem affektiven Bereich.

Bei den durch die Lehrkräfte beurteilten (leistungsfördernden) Persönlichkeitsmerkmalen ergibt sich ein signifikanter positiver Zusammenhang zu den erbrachten Leistungen in drei von vier Tests. Dies zeigt, dass die abgefragten Merkmale geeignet sind, um die Schülerinnen und Schüler hinsichtlich ihrer Leistungsdisposition einzuschätzen und deutet darauf hin, dass diese Merkmale auch als leistungsförderliche Merkmale für den Bereich der Naturwissenschaften gültig sind (Giessel und Höner 2016, Kämpnick 2013).

Die erreichten Punkte in der naturwissenschaftlichen Aufgabensequenz unterscheiden sich nicht signifikant für die Jahrgangsstufen 5 bis 7. Dies mag auf den ersten Blick überraschen, da eine Kompetenzentwicklung beim experimentellen Problemlösen zu den Zielen des naturwissenschaftlichen Unterrichts gehört und eigentlich erwartet werden sollte. Im Hinblick auf die verwendete naturwissenschaftliche Aufgabensequenz ist aber im Vergleich zu vielen anderen Testinstrumenten zu beachten, dass kein fachinhaltliches Vorwissen notwendig ist. Die Siebtklässler haben also in dieser Hinsicht keinen Vorteil gegenüber den Fünft- und Sechstklässlern. Ein anderer Aspekt sind die erforderlichen Prozessschritte für die erfolgreiche Bearbeitung der Aufgabensequenz. Wenn im regulären naturwissenschaftlichen Unterricht das Experimentieren im Sinne der Erkenntnisgewinnung und in Anlehnung z. B. an das Modell von Klahr und Dunbar (1988) geübt und gefestigt würde, sollten die höheren Jahrgänge besser abschneiden. Im Schulalltag lässt sich aber häufig immer noch feststellen, dass Experimente wie Kochvorschriften im Unterricht abgearbeitet werden, ohne dass den Schülerinnen und Schülern der Weg der Erkenntnisgewinnung mit den entsprechenden Strategien abgefordert wird. Da es sich bei der vorliegenden Stichprobe um eine Gelegenheitsstichprobe handelt, könnten aber auch selektive Aspekte die Ergebnisse mitbeeinflussen, was in weiteren Anschlussuntersuchungen überprüft werden soll. Hinsichtlich der Aufgabenschwierigkeit ergibt sich für diese Stichprobe aus Gymnasialschülerinnen und -schülern insgesamt ein Index von 0,67, der zufriedenstellend ist. Es lässt sich vermuten, dass der Schwierigkeitsindex mit Schülerinnen und Schülern anderer Schulformen niedriger ausfällt, was in Folgeuntersuchungen geklärt werden soll.

Damit eine richtige Einschätzung über die Aussagekraft der gewonnenen Erkenntnisse dieser Studie stattfinden kann, sollten folgende Aspekte berücksichtigt werden:

- Bei der zugrundeliegenden Stichprobe handelt es sich um eine Gelegenheitsstichprobe, sodass die Ergebnisse nicht dazu benutzt werden können, um repräsentative Aussagen über alle Schülerinnen und Schüler der untersuchten Altersgruppen zu treffen.
- Obwohl Vorkehrungen getroffen wurden, um eine Durchführungsobjektivität zu gewährleisten (gleiches Material, gleicher Ablauf, gleiche Testinstruktionen, etc.), konnten gewisse Unterschiede nicht vermieden werden. Die Testdurchführungen fanden beispielsweise an unterschiedlichen Tagen und zu verschiedenen Unterrichtszeiten statt.
- Die hier vorgestellte rein quantitative Auswertung der schriftlichen Aufzeichnungen der Schülerinnen und Schüler erlaubt kaum Rückschlüsse auf die angewendeten Problemlösestrategien bzw. auf das Methodenverständnis. Hier sollen sich in Zukunft weitere qualitative Auswertungen anschließen. Insgesamt lässt sich für den Moment ableiten, dass mit der hier vorgestellten naturwissenschaftlich-experimentellen Aufgabensequenz ein Testinstrument vorliegt, dass sich vorwissensneutral in den Klassen 5 bis 7⁸ einsetzen lässt, um Hinweise auf die experimentelle Problemlösekompetenz zu erhalten. Die Aufgabensequenz kann innerhalb einer Schulstunde bearbeitet werden und lässt sich im Anschluss einfach auswerten. Natürlich reicht eine einmalige Messung niemals aus, um eine Beurteilung abzugeben. Die verwendete Aufgabensequenz kann aber ergänzend zu anderen Testverfahren im Schulalltag eingesetzt werden, da sie hinsichtlich der Kriterien nach Baur (2015) geeignet ist.

Danksagung

Wir danken den Schulen, die uns bei der Erprobung des Testinstruments unterstützt haben.

Förderung der Studie

Die Studie wurde als Teilprojekt Diagonal-MINT des Gesamtprojektes TU4Teachers der Technischen Universität Braunschweig im Rahmen der Qualitätsoffensive Lehrerbildung vom Bundesministerium für Bildung und Forschung gefördert. Fördernummer: FKZ 01JA1609

Literatur

- Baur, A. (2015). Inwieweit eignen sich bisherige Diagnoseverfahren des Bereichs Experimentieren für die Schulpraxis? *Zeitschrift für Didaktik der Biologie*, 19, 26-37.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates 2.
- Dörner, D. (1976). *Problemlösen als Informationsverarbeitung* (1. Aufl.). Stuttgart, Berlin, Köln, Mainz: Kohlhammer.
- Emden, M. (2011). *Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*. Berlin: Logos Berlin (Studien zum Physik- und Chemie-lernen, 118).
- Giessel, A., & Höner, K. (2016). Mathematisch-naturwissenschaftliche Talente diagnostizieren – eine Fallstudie. In: K. Höner, M. Looß, R. Müller, A. Strahl (Hrsg.), *Naturwissenschaften vermitteln: Von der frühen Kindheit bis zum Lehrerberuf*. LIT Verlag, Münster, Band 5.
- Glug, I. (2009). *Entwicklung und Validierung eines Multiple-Choice-Tests zur Erfassung prozessbezogener naturwissenschaftlicher Grundbildung*. Dissertation. Christian-Albrechts-Universität Kiel.
- Graubner, R. (1980). *Lexikon der Geologie, Minerale und Gesteine*. Emil Vollmer: München.
- Greiff, S. (2012). *Individualdiagnostik komplexer Problemlösefähigkeit*. Waxmann: Münster.
- Greiff, S. (2013). Der Nutzen einer komplexen Problemlösekompetenz: Theoretische Überlegungen und empirische Befunde. *Zeitschrift für Pädagogische Psychologie*, 27(1-2), 27-39.

⁸ In den beiden Pilotstudien wurde sogar die Einsetzbarkeit für die Jahrgangsstufen 4 und 6 bis 9 gezeigt.

- Greiff, S., & Fischer, A. (2013). Der Nutzen einer komplexen Problemlösekompetenz. Theoretische Überlegungen und empirische Befunde. *Zeitschrift für Pädagogische Psychologie*, 27(1-2), 27–39.
- Grube, C. R. (2011). *Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung. Untersuchung der Struktur und Entwicklung des wissenschaftlichen Denkens bei Schülerinnen und Schülern der Sekundarstufe I. Dissertation*. Kassel: Universität Kassel. <https://kobra.bibliothek.uni-kassel.de/bitstream/urn:nbn:de:hebis:34-2011041537247/3/DissertationChristianeGrube.pdf>. Zugegriffen: 10.08.2017.
- Gut, C. (2012). Modellierung und Messung experimenteller Kompetenz: Analyse eines large-scale Experimentiertests. Band 134. LOGOS: Berlin.
- Hammann, M. (2004). Kompetenzentwicklungsmodelle. Merkmale und ihre Bedeutung - dargestellt anhand von Kompetenzen beim Experimentieren. *Der mathematische und naturwissenschaftliche Unterricht*, 57(4), 196-203.
- Hammann, M., Phan, T., Ehmer, M., & Grimm, T. (2008). Assessing pupils 'skills in experimentation. *Journal of Biological Education*, 42(2), 66-72.
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion*. Berlin, Heidelberg: Springer-Verlag, 2. Aufl., 143-171.
- Heitmann, P. (2012). *Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*. Dissertation. Humboldt-Universität zu Berlin.
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+R)* (3. Aufl.). Göttingen: Beltz Test.
- Helmke, A., & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In: F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 71-176). Göttingen: Hogrefe.
- Henke, C. (2007). *Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. Untersuchungen am Beispiel des HIGHSEA-Projekts in Bremerhaven*. Berlin: LOGOS Verlag.
- Hesse, I., & Latzko, B. (2011). *Diagnostik für Lehrkräfte* (2. Aufl.). Opladen [u.a.]: Budrich.
- Heublein, U., Richter, J., Schmelzer, R. & Sommer, D. (2012). *Die Entwicklung der Schwund- und Studienabbrecherquoten an den deutschen Hochschulen. Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2010* (HIS: Forum Hochschule 3 | 2012). Hannover: HIS.
- Käpnick, F. (2001). *Mathe für kleine Asse. Handbuch für die Förderung mathematisch interessierter und begabter Dritt- und Viertklässler*. Volk und Wissen: Berlin.
- Käpnick, F. (2013). Theorieansätze zur Kennzeichnung des Konstruktes "Mathematische Begabung" im Wandel der Zeit. In: *Mathematische Begabungen – Denkansätze zu einem komplexen Themenfeld aus verschiedenen Perspektiven*. Band 4, Münster, WTM-Verlag, 9-39.
- Kampa, N. (2012). *Aspekte der Validierung eines Tests zur Kompetenz in Biologie – Eine Studie zur Kompetenz in Biologie und ihren Teildimensionen Konzept und Prozesswissen*. Berlin: Humboldt-Universität zu Berlin, Zugriff am: 10.08.2017, verfügbar unter <http://edoc.hu-berlin.de/dissertationen/kampa-nele-2012-12-19/PDF/kampa.pdf>
- Klahr, D., & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science*, 12, 1–48.
- Klieme, E., Funke, J., Leutner, D., Reimann, P., & Wirth, J. (2001). Problemlösen als fächerübergreifende Kompetenz? Konzeption und erste Resultate aus einer Schulleistungsstudie. *Zeitschrift für Pädagogik*, 47, 179–200.
- Klos, S. (2008). *Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht – der Einfluss eines integrierten Unterrichtskonzepts*. Berlin: LOGOS Verlag.
- Klos, S., Henke, C., & Kieren, C. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen - zwei verschiedene Kompetenzen. *Zeitschrift für Pädagogik*, 54(3), 304-321.

- Köller, O. (2016). Leistungsstandards und Leistungsbewertung an Gymnasien und Universitäten – zur Passung in den mathematisch-naturwissenschaftlichen Fächern. In: S. Lin-Kitzing, D. Di Fuccia, T. Gaube (Hrsg.): *Leistungsstandards und Leistungsbewertung an Gymnasien und Universitäten. Beiträge zur (nicht) vorhandenen Passung*. Klinkhardt: Bad Heilbrunn, S. 37-56.
- Kretzschmar, A., Neubert, J. C., & Greiff, S. (2014). Komplexes Problemlösen, schulfachliche Kompetenzen und ihre Relation zu Schulnoten. *Zeitschrift für Pädagogische Psychologie*, 28(4), 205-215.
- Künsting, J. (2007). *Effekte von Zielqualität und Zielspezifität auf selbstreguliert-entdeckendes Lernen durch Experimentieren*. Dissertation, Universität Duisburg-Essen.
- Leutner, D. (2002). The fuzzy relationship of intelligence and problem solving in computer simulations. *Computers in Human Behavior*, 18 (6), 685–697.
- Leutner, D., Wirth, J., Klieme, E., & Funke, J. (2005). Ansätze zur Operationalisierung und deren Erprobung im Feldtest zu PISA 2000. In E. Klieme, D. Leutner & J. Wirth (Hrsg.). *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie* (1. Aufl., S.21-36). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Leutner D., Fleischer, J., Wirth, J., Greiff, S., & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien. Untersuchungen zur Dimensionalität. *Psychologische Rundschau*, 63(1), 34–42.
- Maier, U. (2015). *Leistungsdiagnostik in Schule und Unterricht*. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Maiseyenko, V. (2014). *Modellbasiertes Experimentieren im Unterricht. Praxistauglichkeit und Lernwirkungen*. (Studien zum Physik- und Chemielernen, 166). Berlin: Logos-Verlag.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.). *Theorien in der biologiedidaktischen Forschung* (S.177-186). Berlin, New York: Springer.
- Mayer, J., Keiner, K., & Ziemek, H.-P. (2003), Naturwissenschaftliche Problemlösekompetenz im Biologieunterricht. In: A. Bauer et al. (Hrsg.): *Entwicklung von Wissen und Kompetenzen im Biologieunterricht*. Internationale Tagung der Sektion Fachdidaktik im VdBiol. Berlin, S. 21-24. Kiel: IPN.
- Meidinger, H.-P. (2014). Abitur. Studienberechtigung, Studienbefähigung, Studienerfolg? In S., Lin-Kitzing, D., Di Fuccia & R., Stengl-Jörns (Hrsg.). *Abitur und Studierfähigkeit. Ein interdisziplinärer Dialog* (S. 27-35). Bad Heilbrunn: Klinkhardt.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- Müller, R., & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine*, 13, 2465-2476.
- Nehring, A. (2014). *Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung*. Berlin: LOGOS, Band 177.
- Niedersächsisches Kultusministerium (2007): *Kerncurriculum für das Gymnasium Schuljahrgänge 5-10: Naturwissenschaften*. Hannover. Online verfügbar unter http://db2.nibis.de/1db/cuvo/datei/kc_gym_nws_07_nib.pdf, zuletzt geprüft am 10.08.2017.
- Niedersächsisches Kultusministerium (2012): *Kerncurriculum für die Integrierte Gesamtschule Schuljahrgänge 5-10: Naturwissenschaften*. Hannover. Online verfügbar unter http://db2.nibis.de/1db/cuvo/datei/kc_2012_igs_nws_i.pdf, zuletzt geprüft am 10.08.2017.
- Niedersächsisches Kultusministerium (2013): *Kerncurriculum für die Oberschule Schuljahrgänge 5-10: Naturwissenschaften*. Hannover. Online verfügbar unter <http://db2.nibis.de/1db/cuvo/datei/kcobsnaturwissenschaften.pdf>, zuletzt geprüft am 10.08.2017.

- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P., & Klopp, A. (2001). Naturwissenschaftliche Grundbildung. Testkonzeption und Ergebnisse. In Deutsches PISA-Konsortium (Hrsg.). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S.192-248). Opladen: Leske + Budrich.
- Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. *Zeitschrift für Psychologie*, *189*, 79-100.
- Rieß, W., & Robin, N. (2012). Befunde aus der empirischen Forschung zum Experimentieren im mathematisch-naturwissenschaftlichen Unterricht. In W. Rieß, M. Wirtz, B. Barzel & A. Schulz (Hrsg.). *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht* (S. 129-152). Münster: Waxmann.
- Rindermann, H., & Oubaid, V. (1999). Auswahl von Studienanfängern durch Universitäten – Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *20*, 172-191.
- Schecker, H., Neumann, K., Theyßen, H., Eickhorst, B., & Dickmann, M. (2016). Stufen experimenteller Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, *22*, 197-213.
- Scherer, R. (2014a). *Analyse der Struktur, Messvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*. (Studien zum Physik- und Chemielernen, 141). Berlin: Logos-Verlag.
- Scherer, R. (2014b). Komplexes Problemlösen im Fach Chemie. Ein domänenspezifischer Zugang. *Zeitschrift für Pädagogische Psychologie*, *28*(4), 181–192.
- Schreiber, N. (2012). *Diagnostik experimenteller Kompetenz: Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*. LOGOS: Berlin, Bd. 139.
- Schreiber, N., Theyßen, H., & Schecker, H. (2014). Diagnostik experimenteller Kompetenz: Kann man Realexperimente durch Simulationen ersetzen? *Zeitschrift für Didaktik der Naturwissenschaften*, *20*, 161-173.
- Schuler, H. (2000). Berufseignungsdiagnostik. In: G. Wenninger (Hrsg.), *Lexikon der Psychologie*. Spektrum: Heidelberg, Berlin, 209-212.
- Schuler, H. (2006). Noten als Prädiktoren von Studien- und Berufserfolg. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Psychologie Verlags Union: Weinheim, 535-541.
- Schulz, A. (2011). *Experimentierspezifische Qualitätsmerkmale im Chemieunterricht. Eine Videostudie* (Studien zum Physik- und Chemielernen, 113). Berlin: Logos Verlag.
- Schumann, W (2002). *Edelsteine und Schmucksteine. Alle Arten und Varietäten der Welt. 1600 Einzelstücke*. BLV: München u.a., 13. Auflage.
- Seidel, T., Prenzel, M., Duit, R., & Lehrke, M. (2003). *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“*. Kiel: Institut für die Pädagogik der Naturwissenschaften (IPN).
- Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (1999). Note on sources of sampling variability in science performance assessments. *Journal of Educational Measurement*, *36*(1), 61-71.
- Stern, E., Hardy, I. (2004): *Differentielle Psychologie des Lernens in Schule und Ausbildung*. In: Birbaumer et al.: *Enzyklopädie der Psychologie - Themenbereich C: Theorie und Forschung - Serie VIII: Differentielle Psychologie und Persönlichkeitsforschung - Band 5 Theorien und Anwendungsfelder*. Hogrefe Verlag, S. 595.
- Süß, H.-M. (1996). *Intelligenz, Wissen und Problemlösen. Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen* (Lehr- und Forschungstexte Psychologie, n.F., 5). Göttingen [u.a.]: Hogrefe, Verlag für Psychologie.
- Theyßen, H., Schecker, H., Neumann, K., Dickmann, M., & Eickhorst, B. (2013). Messung experimenteller Kompetenz in Large Scale Assessments. In: S. Bernholt (Hg.): *Inquiry-based Learning – Forschendes Lernen*. GDGP Jahrestagung Hannover 2012. Kiel, 596-598.

- Trappmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 21, 11-27.
- Trost, G. (1975). *Vorhersage des Studienerfolgs*. Braunschweig: Westermann.
- Vorholzer, A., von Aufschnaiter, C., & Kirschner, S. (2016). Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses experimenteller Denk- und Arbeitsweisen. *Zeitschrift für Didaktik der Naturwissenschaften*, 22, 25-41.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in der Schule – eine umstrittene Selbstverständlichkeit. In: F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17-31). Weinheim: Beltz Verlag.
- Wellnitz, N., & Mayer, J. (2013). Erkenntnismethoden in der Biologie – Entwicklung und Evaluation eines Kompetenzmodells. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 335-345.
- Wirth, J., Leutner, D., & Klieme, E. (2005). Problemlösekompetenz - Ökonomisch und zugleich differenziert erfassbar? In E. Klieme, D. Leutner & J. Wirth (Hrsg.). *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie* (1. Aufl., S. 73-82). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Wirtz, M., A., & Caspar, F. (2007). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*, Göttingen: Hogrefe.
- Wisniewski, B. (2013). *Psychologie für die Lehrerbildung*, Bad Heilbrunn: Klinkhardt.
- Wittmann, W. W., & Süß, H.-M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik symmetry. In P. L. Ackerman, P. C. Kyllonen & R. D. Roberts (Eds.), *Learning and individual differences. Process, trait, and content determinants*. Washington: American Psychological Association.
- Wüstenberg, S., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The role of Strategy Knowledge for the Application of Strategies in Complex Problem Solving Tasks. *Tech Know Learn*, 19(1-2), 127-146.