

van den Ham, Ann-Katrin; Ehmke, Timo; Hahn, Inga; Wagner, Helene; Schöps, Katrin
Mathematische und naturwissenschaftliche Kompetenz in PISA, im IQB-Ländervergleich und in der National Educational Panel Study (NEPS) – Vergleich der Rahmenkonzepte und der dimensionalen Struktur der Testinstrumente

Bundesministerium für Bildung und Forschung [Hrsg.]: Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments. Stand August 2016. Berlin : Bundesministerium für Bildung und Forschung 2016, S. 140-160. - (Bildungsforschung; 44)



Empfohlene Zitierung/ Suggested Citation:

van den Ham, Ann-Katrin; Ehmke, Timo; Hahn, Inga; Wagner, Helene; Schöps, Katrin: Mathematische und naturwissenschaftliche Kompetenz in PISA, im IQB-Ländervergleich und in der National Educational Panel Study (NEPS) – Vergleich der Rahmenkonzepte und der dimensionalen Struktur der Testinstrumente - In: Bundesministerium für Bildung und Forschung [Hrsg.]: Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments. Stand August 2016. Berlin : Bundesministerium für Bildung und Forschung 2016, S. 140-160 - URN: urn:nbn:de:0111-pedocs-126776

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

*Ann-Katrin van den Ham, Timo Ehmke, Inga Hahn,
Helene Wagner, Katrin Schöps*

Mathematische und naturwissenschaftliche Kompetenz in PISA, im IQB-Ländervergleich und in der National Educational Panel Study (NEPS) – Vergleich der Rahmenkonzepte und der dimensionalen Struktur der Testinstrumente

1 Einleitung

Bildung gilt als eine wesentliche Voraussetzung für die aktive Partizipation als verantwortungsvolle(r) Bürgerin bzw. Bürger in einer demokratischen Gesellschaft. Bisher ist allerdings noch wenig über die kumulative Entwicklung von Kompetenzen über die Lebensspanne bekannt. Um Einsicht in den Bildungsprozess und die Kompetenzentwicklung zu erhalten, wurde in Deutschland die National Educational Panel Study (NEPS) ins Leben gerufen (Blossfeld, Maurice & Schneider, 2011). NEPS untersucht, wie sich Kompetenzen von Kindern, Jugendlichen und Erwachsenen über den Lebenslauf entwickeln, wie diese Kompetenzen die Bildungskarriere beeinflussen und inwiefern die Kompetenzentwicklung von Lerngelegenheiten mitbestimmt wird. Im Jahr 2010 wurde in NEPS erstmalig eine Kohorte von ca. 16.500 Neuntklässlerinnen und Neuntklässlern ausgewählt.

Im Rahmen des Bildungsmonitorings wird in Deutschland zusätzlich regelmäßig das Erreichen der nationalen Bildungsstandards am Ende der Sekundarstufe durch das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) geprüft (Pant, Stanat, Schroeders, Roppelt, Siegle & Pöhlmann, 2013). Außerdem werden die Kompetenzen von Schülerinnen und Schülern in Deutschland im internationalen Vergleich durch die regelmäßige Teilnahme an Programme for International Student Assessment (PISA: Organisation for Economic Cooperation and Development [OECD], 2013b; Prenzel, Sälzer, Klieme & Köller, 2013) erfasst.

Obwohl die Testkonzeptionen aus NEPS im Bereich der mathematischen und naturwissenschaftlichen Kompetenzmessungen am Ende der Sekundarstufe (NEPS-K9) große Überschneidungen mit dem IQB-Ländervergleich (LV 2012) und PISA 2012 aufweisen, sind die Ergebnisse der Studien aufgrund von unterschiedlichen Berichtsskalen nicht direkt vergleichbar. Eine Verlinkung der Berichtsskalen der Mathematik- und Naturwissenschaftsmessungen in NEPS-K9 mit den Mathematik- und Naturwissenschaftsmessungen im LV 2012 und in PISA 2012 würde jedoch eine Interpretation der Befunde aus NEPS in einem nationalen bzw. internationalen Referenzrahmen ermöglichen.

Eine Verlinkung von Berichtsskalen aus unterschiedlichen Large-Scale-Assessments erfordert nach Kolen und Brennan (2010) und van de Vijver (1998), dass sich die Studien hinsichtlich der (1) Schlussfolgerungen, (2) Population, (3) Messeigenschaften und -bedingungen sowie der (4) operationalisierten Konstrukte (konzeptionell, dimensional und skalenbezogen) hinreichend ähnlich sind (vgl. Ehmke et al., 2014). Es wird also vorausgesetzt, dass zwischen einem latenten, also nicht direkt beobachtbaren Konstrukt, den verwendeten Testmodellen und den eingesetzten Testitems gleiche Beziehungen bestehen. Dies wird erreicht, wenn neben der konzeptuellen Äquivalenz auch dimensionale Äquivalenz und Skalenäquivalenz gegeben sind.

Ziel des vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Projektes „PISA, Bildungsstandards und die National Educational Panel Study (NEPS). Vergleich der Rahmenkonzepte und Validierung der NEPS-Testinstrumente in den Naturwissenschaften und in der Mathematik“ war es daher, unter anderem die Naturwissenschafts- und Mathematiktests aus NEPS-K9 mit PISA 2012 und LV 2012 anhand der oben genannten Kriterien zu vergleichen und die Berichtsskalen gegebenenfalls zu verankern. In diesem Beitrag werden wesentliche Ergebnisse zur konzeptionellen und dimensional Äquivalenz der Tests vorgestellt.

2 Theoretischer Hintergrund – Überblick über die Studien

2.1 IQB-Ländervergleich (LV 2012)

Das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) wurde von der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) beauftragt, die Einhaltung der Bildungsstandards regelmäßig und flächendeckend zu überprüfen, die Bildungsstandards weiterzuentwickeln und Kompetenzstufenmodelle zu entwerfen. Am Ende der Sekundarstufe I findet daher alle drei Jahre, alternierend für die Bereiche Deutsch und erste Fremdsprache sowie Mathematik und Naturwissenschaften, der LV statt. Der LV bietet insbesondere die Möglichkeit, die Kompetenzniveaus von Schülerinnen und Schülern zwischen den Bundesländern zu vergleichen. In der Erhebung in 2012 wurden die Bereiche Mathematik und Naturwissenschaften erfasst. Insgesamt wurden ca. 45.000 Neuntklässlerinnen und Neuntklässler getestet (Pant et al., 2013).

Der Mathematiktest des LV 2012 basiert auf dem Rahmenkonzept der Bildungsstandards im Fach Mathematik am Ende der Sekundarstufe I (Roppelt, Blum & Pöhlmann, 2013) und unterscheidet drei Teilbereiche. So werden zum Lösen einer mathematischen Aufgabe Kenntnisse über mathematische Inhalte benötigt, die fünf mathematischen Inhaltsbereichen (Leitideen) zugeordnet werden können: Zahl, Messen, Raum und Form, Funktionaler Zusammenhang und Daten und Zufall. Außerdem müssen beim Lösen der Aufgaben bestimmte mathematische Prozesse angewendet werden, die sich durch sechs allgemeine mathematische Kompetenzen (Prozesse) beschreiben lassen: „Mathematisch argumentieren“, „Probleme mathematisch lösen“, „Mathematisch modellieren“, „Mathematische Darstellungen verwenden“,

„Formal-technisches Arbeiten“ und „Mathematisch kommunizieren“. Die mathematischen Aktivitäten werden auf unterschiedlichen kognitiven Anforderungsniveaus gefordert, die in drei Anforderungsbereichen beschrieben werden: „Reproduzieren“, „Zusammenhänge herstellen“ und „Verallgemeinern und Reflektieren“ (KMK, 2003; van den Ham et al., 2014).

Der Naturwissenschaftstest basiert auf dem Kompetenzmodell der Bildungsstandards in den Fächern Biologie, Physik und Chemie (Pant, Stanat, Schroeders et al., 2013). Die naturwissenschaftliche Kompetenz wird in vier Kompetenzbereiche unterteilt: „Umgang mit Fachwissen“ (fachspezifisch für Biologie, Chemie und Physik), „Erkenntnisgewinnung“, „Kommunikation“ und „Bewertung“. Weiterhin werden für jeden Kompetenzbereich fünf Komplexitätsstufen unterschieden: „ein Fakt“, „zwei Fakten“, „ein Zusammenhang“, „zwei Zusammenhänge“ und „übergeordnetes Konzept“. Eine weitere Dimension zur Untersuchung der naturwissenschaftlichen Kompetenz stellen die kognitiven Prozesse dar, die in „reproduzieren“, „selegieren“, „organisieren“ und „integrieren“ unterteilt werden (KMK, 2004a, 2004b, 2004c; Wagner, Schöps, Hahn, Pietsch & Köller, 2014).

2.2 Programme for International Student Assessment (PISA 2012)

PISA wird von der Organisation for Economic Co-operation and Development (OECD) geleitet. PISA erfasst international vergleichend Schülerleistungen in einem Zyklus von drei Jahren. Es werden die drei Domänen Lesen, Mathematik und Naturwissenschaften unterschieden, wobei an jedem Messzeitpunkt eine der Domänen alternierend schwerpunktmäßig erhoben wird. In der jeweiligen Hauptdomäne werden umfassendere Tests durchgeführt. Die Zielgruppe der PISA-Studie sind Jugendliche im Alter von 15 Jahren, wodurch die Ergebnisse Aussagen über grundlegende Kompetenzen von Schülerinnen und Schülern am Ende der Pflichtschulzeit ermöglichen. Zusätzlich werden Bezugspunkte für die Weiterentwicklung und eine Wissensbasis für die politische Steuerung des Bildungssystems gewonnen (OECD, 2013b; Prenzel et al., 2013).

In PISA 2012 wurden in Deutschland ca. 5.000 15-Jährige im Rahmen der Studie getestet (Prenzel et al., 2013). Die Hauptdomäne in 2012 war die mathematische Kompetenz der Schülerinnen und Schüler. Die zugrunde liegende mathematische Rahmenkonzeption basiert auf dem mathematischen Grundbildungskonzept (Mathematical Literacy). PISA definiert Mathematical Literacy wie folgt:

„Mathematical literacy is an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens“ (OECD, 2013a, 25).

Diese Definition bringt zum Ausdruck, dass PISA auf eine enge Orientierung an Lehrplänen verzichtet und vor allem die funktionale Nutzung von Mathematik in den Vordergrund stellt. Mathematische Probleme sollen dementsprechend in unter-

schiedlichen Kontexten und Situationen gelöst werden (Neubrand et al., 2001). In der mathematischen Rahmenkonzeption werden die Dimensionen Inhalt, Prozess und Kontext beschrieben. Hierzu gehören vier mathematische Inhaltsbereiche: „Quantität“, „Veränderung und Beziehungen“, „Raum und Form“ und „Unsicherheit und Daten“, die mit drei mathematischen Prozessen verbunden sind, welche die grundlegenden mathematischen Fähigkeiten konkretisieren: „Situationen mathematisch formulieren“, „mathematische Konzepte, Fakten, Prozeduren und Schlussfolgerungen anwenden“ und „mathematische Ergebnisse interpretieren, anwenden und bewerten“. Außerdem werden sieben fundamentale mathematische Fähigkeiten definiert, die zum Lösen einer Aufgabe benötigt werden (kommunizieren, mathematisieren, repräsentieren, argumentieren, Problemlösestrategien entwickeln, mit Mathematik symbolisch, formal und technisch umgehen sowie mathematische Hilfsmittel verwenden). Die Aufgaben werden zusätzlich in unterschiedlichen Kontexten verortet: persönliche, berufliche, gesellschaftliche und wissenschaftliche Kontexte (OECD, 2013a).

Die naturwissenschaftliche Rahmenkonzeption basiert auf dem naturwissenschaftlichen Grundbildungskonzept (scientific literacy). In PISA wird scientific literacy wie folgt definiert:

„For the purposes of PISA, scientific literacy refers to an individual’s:

- Scientific knowledge and use of that knowledge to identify questions, acquire new knowledge, explain scientific phenomena and draw evidence-based conclusions about science-related issues.
- Understanding of the characteristic features of science as a form of human knowledge and enquiry.
- Awareness of how science and technology shape our material, intellectual and cultural environments.
- Willingness to engage in science-related issues, and with the ideas of science, as a reflective citizen“ (OECD, 2013a, 100).

In der Rahmenkonzeption werden für die Naturwissenschaften drei Teilkompetenzen unterschieden: „naturwissenschaftliche Fragestellungen erkennen“, „naturwissenschaftliche Phänomene erklären“ und „naturwissenschaftliche Evidenz nutzen“. Die Basis für diese Teilkompetenzen bilden das (objektbezogene) naturwissenschaftliche Wissen und (Meta-)Wissen über die Naturwissenschaften. Der naturwissenschaftlichen Kompetenz werden vier Wissenssysteme untergeordnet: „Physikalische Systeme“, „Lebende Systeme“, „Erd- und Weltraumsysteme“ und „Technologische Systeme“. Dem Bereich „Wissen über die Naturwissenschaften“ werden zwei Aspekte untergeordnet: „naturwissenschaftliches Forschen“ und „naturwissenschaftliche Erklärungen“. Die motivationalen Orientierungen und Einstellungen einer Person bilden eine weitere Komponente, auf der die Teilkompetenzen der naturwissenschaftlichen Grundbildung basieren. Die naturwissenschaftliche Kompetenz wird bei PISA situations- bzw. kontextgebunden untersucht. In der Rahmenkonzeption werden folgende fünf Kontexte differenziert: „Gesundheit“, „natürliche Ressourcen“, „Umwelt, Risiken/Gefahren“ und zuletzt „Grenzen von Naturwissenschaften und Technik“ (OECD, 2006; Prenzel et al., 2007; Wagner et al., 2014).

2.3 Das Nationale Bildungspanel (NEPS)

NEPS wird unter Federführung des Leibniz-Instituts für Bildungsverläufe (IfBi) durchgeführt. Die Studie differenziert fünf inhaltliche Untersuchungsbereiche (Säulen) (Kompetenzentwicklung, Lernumwelten, Bildungsentscheidungen, Migrationshintergrund und Bildungsrenditen), die im Längsschnitt über acht Lebensstapen (von den Neugeborenen und der frühkindlichen Betreuung bis hin zur Erwachsenenbildung und dem lebenslangen Lernen) untersucht werden. Bezüglich der Kompetenzentwicklung werden die Fähigkeiten in den Bereichen Mathematik, Sprache (Orthografie, Lese- und Hörverstehen in deutscher und englischer Sprache, Kenntnisse in der Erstsprache bei Schülerinnen und Schülern mit Migrationshintergrund), Naturwissenschaften, ICT-Literacy und Problemlösen erfasst. Die Daten werden mithilfe eines Multi-Kohorten-Sequenz-Designs erhoben und der Wissenschaft in Form von scientific use files zu Forschungszwecken zur Verfügung gestellt (Blossfeld, Maurice & Schneider, 2011; Weinert, Artelt, Prenzel, Senkbeil, Ehmke & Carstensen, 2011). Im Herbst 2010 startete zum ersten Mal eine Kohorte von ca. 15.000 Schülerinnen und Schülern der neunten Jahrgangsstufe (NEPS-K9).

In NEPS wird mathematische Kompetenz als das Ausmaß beschrieben, „in dem Schülerinnen und Schüler, aber auch Erwachsene, die in der Schule gelernte Mathematik in problemhaltigen, vorwiegend außermathematischen Situationen flexibel anwenden können“ (Ehmke et al., 2009, 317). Diese Definition basiert auf dem mathematischen Grundbildungskonzept, wie es im Rahmen von PISA definiert wurde (Weinert et al., 2011). Damit steht auch im NEPS-Mathematiktest die Bearbeitung kontextbezogener Probleme im Vordergrund. Die mathematische Rahmenkonzeption differenziert zwischen einer inhaltlichen und einer prozessbezogenen Komponente. Es werden vier mathematische Inhaltsbereiche definiert: „Quantität“, „Veränderung und Beziehungen“, „Raum und Form“ und „Daten und Zufall“. Die prozessorientierte Komponente umfasst mathematische und kognitive Denkanforderungen, die beim Lösen mathematischer Aufgaben erforderlich sein können. Zudem wird zwischen sechs kognitiven Komponenten mathematischer Denkprozesse differenziert (vgl. Niss, 2003; KMK, 2003): „mathematisch argumentieren“, „mathematisch kommunizieren“, „modellieren“, „mathematische Probleme lösen“, „repräsentieren“ (Darstellungen verwenden) und „technische Fertigkeiten einsetzen“ (Ehmke et al., 2009).

Das Rahmenkonzept der Naturwissenschaften basiert auf dem naturwissenschaftlichen Grundbildungskonzept, welches von der American Association for the Advancement of Science (AAAS, 1993, 2009) und PISA (OECD, 2006) definiert wurde. Dabei wird davon ausgegangen, dass grundlegendes Verständnis von naturwissenschaftlichen Konzepten und Prozessen für das alltägliche Leben relevant ist und die Basis für lebenslanges Lernen formt: „Rather than focusing on memorized knowledge, scientific literacy reflects the ability to apply one's existing scientific knowledge in different everyday life contexts and situations“ (Hahn et al., 2013). Die naturwissenschaftliche Kompetenz wird ähnlich wie in PISA in die objektbezogene Komponente (naturwissenschaftliches Wissen) und prozessbezogene Komponente (Wissen über die Naturwissenschaften) unterteilt. Das „naturwissenschaftliche Wis-

sen“ umfasst in der NEPS-Rahmenkonzeption die inhaltsbezogenen Komponenten „Stoffe“, „Systeme“, „Entwicklung“ und „Wechselwirkungen“. Innerhalb des „Wissens über die Naturwissenschaften“ werden die prozessbezogenen Komponenten „Naturwissenschaftliche Denk- und Arbeitsweisen“ unterschieden. Die Erfassung der naturwissenschaftlichen Kompetenz erfolgt eingebettet in die ausgewählten Kontexte: „Gesundheit“, „Umwelt“ und „Technologie“ (Hahn et al., 2013).

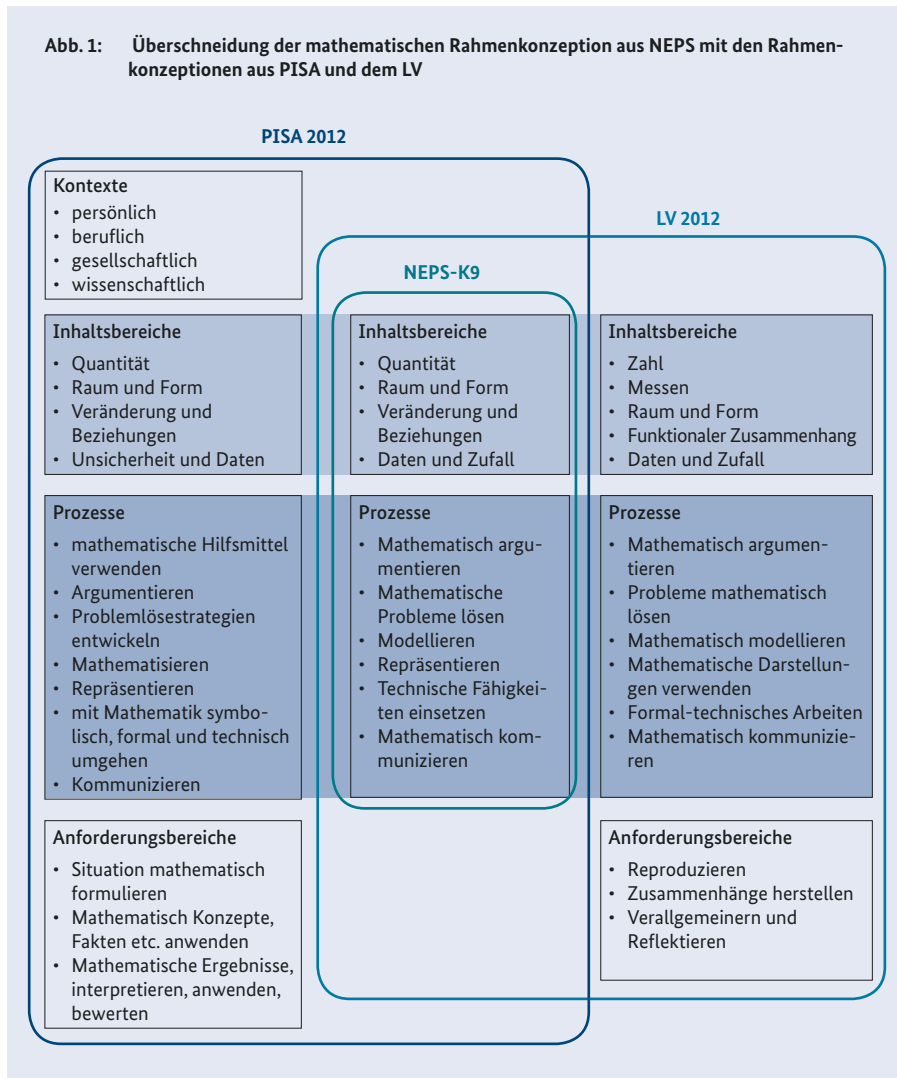
2.4 Vergleich der Rahmenkonzeptionen

Alle drei Studien beanspruchen für sich, sowohl die mathematische als auch die naturwissenschaftliche Kompetenz von Schülerinnen und Schülern am Ende der Sekundarstufe I zu messen. Es lassen sich jedoch Unterschiede in den Zielstellungen und den damit verbundenen Schlussfolgerungen zwischen den Studien NEPS-K9, PISA 2012 und LV 2012 finden. So überprüft der LV 2012, inwieweit die in den Bildungsstandards formulierten Lernziele für Mathematik und die Naturwissenschaften am Ende der Sekundarstufe I erreicht wurden, und ermöglicht einen Vergleich der Kompetenzniveaus der Schülerinnen und Schüler zwischen den Bundesländern. PISA hingegen schafft eine Verortung der Kompetenzen von Schülerinnen und Schülern aus Deutschland an einem internationalen Referenzmaßstab (im Vergleich zwischen OECD-Staaten). Durch die zyklischen Durchführungen ermöglichen beide Studien außerdem Aussagen zu Trendentwicklungen (LV erst ab 2018). Im Unterschied dazu erlaubt NEPS bislang keine Verortung an Referenzmaßstäben. Das Ziel dieser Studie ist eine längsschnittliche Analyse der Entwicklung von Kompetenzen der Stichprobe von Schülerinnen und Schülern. Während anhand von NEPS Aussagen zu den längsschnittlichen Entwicklungen getroffen werden sollen, lassen die PISA-Ergebnisse Aussagen auf Ebene der Staaten und die LV-Ergebnisse Aussagen auf Ebene der Länder in Deutschland zu. Unterschiede in den Stichproben zwischen den Studien kommen dadurch zustande, dass diese im LV 2012 und in NEPS klassenbasiert sind, in PISA 2012 dagegen altersbasiert gezogen werden (vgl. van den Ham et al., 2014). Wobei die Mehrheit der Neuntklässlerinnen und Neuntklässler etwa 15 Jahre alt ist und es dadurch zwischen den Zielpopulationen einen hohen Überlappungsbereich gibt (Sälzer, Prenzel & Klieme, 2013).

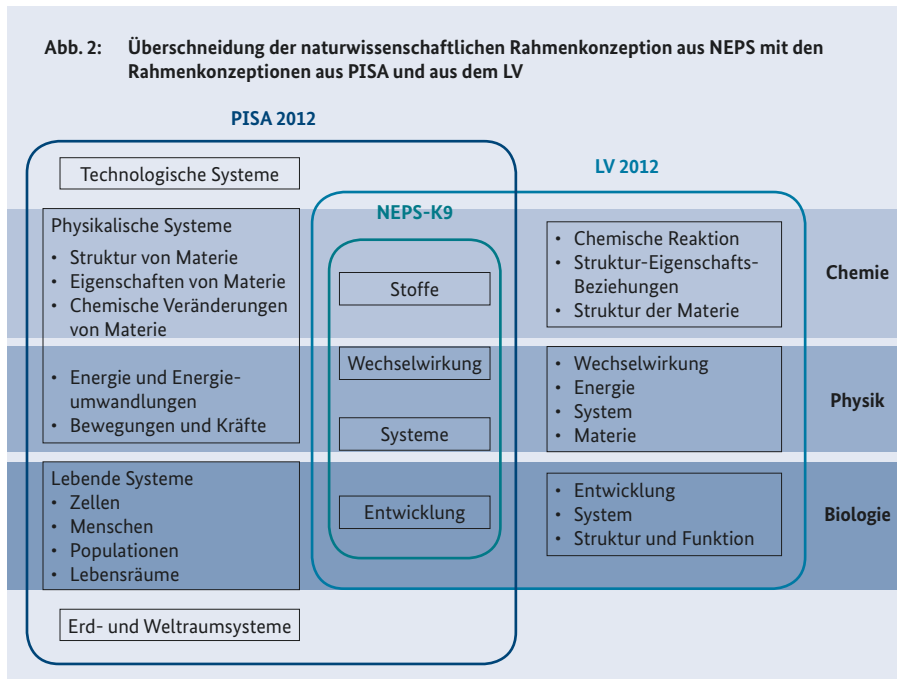
Sowohl die mathematischen als auch die naturwissenschaftlichen Rahmenkonzepte aus PISA 2012 und NEPS-K9 basieren auf einem Literacy-Ansatz, der explizit nicht curriculumorientiert ist. Die Definition von mathematischer und naturwissenschaftlicher Kompetenz im LV 2012 basiert hingegen auf den KMK-Bildungsstandards. Diese geben vor, welche Kompetenzen etwa am Ende der Sekundarstufe I entwickelt sein sollten, und sind damit Richtschnur für die Kerncurricula der Länder (Pant et al., 2013; Köller, 2008).

Das Konstrukt der mathematischen Kompetenz wird sowohl in PISA 2012 als auch im LV 2012 und in NEPS-K9 durch die Definition von Teilbereichen weiter ausdifferenziert (vgl. Abbildung 1). Die NEPS-Rahmenkonzeption Mathematik orientiert sich bei der Gliederung der Inhaltsdimension vor allem an dem Rahmenkonzept von PISA. Der LV definiert einen Inhaltsbereich („Messen“) mehr als NEPS und

PISA. Die Inhalte des Bereichs „Messen“ fallen in NEPS unter den Bereich „Zahl“. Bei der Prozessdimension greift das NEPS-Rahmenkonzept die Bereiche aus dem LV auf. Sowohl die PISA-Rahmenkonzeption als auch die LV-Rahmenkonzeption definieren außerdem sogenannte Anforderungsniveaus. Diese beschreiben die kognitiven Anforderungen einer Aufgabe. PISA unterscheidet zudem unterschiedliche Aufgabenkontexte. Diese werden weder in der Rahmenkonzeption des LV noch in der Rahmenkonzeption der NEPS-Studie bestimmt.



Die NEPS-Rahmenkonzeption für den Bereich naturwissenschaftlicher Kompetenz wurde in Anlehnung an die PISA-Rahmenkonzeption und die Konzeption der Bildungsstandards entwickelt (Hahn et al., 2013). Abbildung 2 zeigt die inhaltliche Überschneidung der Rahmenkonzeption aus NEPS mit der Rahmenkonzeption aus PISA im Bereich des naturwissenschaftlichen Wissens und mit der Rahmenkonzeption aus dem LV im Kompetenzbereich Fachwissen. Demnach kann das NEPS-Konzept „Stoffe“ in den Umgang mit Fachwissen „Chemie“ (Bildungsstandards) und „Physikalische Systeme“ (PISA) eingeordnet werden. Ein Äquivalent für das Konzept „Wechselwirkungen“ aus NEPS findet sich in der Rahmenkonzeption der Bildungsstandards im Kompetenzbereich Umgang mit Fachwissen „Physik“ und in der Rahmenkonzeption aus PISA im Wissenssystem „Physikalische Systeme“. In der NEPS-Rahmenkonzeption umfasst das Konzept „Systeme“ biologische und technologische Systeme. Aus diesem Grund steht dieses Konzept an der Schnittstelle zwischen dem Umgang mit Fachwissen „Biologie“ und „Physik“ in den Bildungsstandards und den „Physikalischen und lebenden Systemen“ bei PISA. Schließlich kann auch eine Überschneidung des NEPS-Konzepts Entwicklung mit dem Kompetenzbereich Umgang mit Fachwissen „Biologie“ (Bildungsstandards) einerseits und dem Wissenssystem „Lebende Systeme“ (PISA) andererseits postuliert werden (vgl. Wagner et al., 2014).



Fazit: Die mathematischen und naturwissenschaftlichen Rahmenkonzeptionen aus den Studien PISA 2012, NEPS-K9 und LV 2012 weisen große inhaltliche Überschneidungen auf. So sind die Definition der mathematischen Inhaltsbereiche und kognitiven Prozesse sehr ähnlich, und es gibt deutliche Überschneidungen der NEPS-Rahmenkonzeption mit den Beschreibungen des Bereiches naturwissenschaftliches

Wissen in PISA 2012 und mit der Rahmenkonzeption aus dem LV 2012 im Kompetenzbereich Fachwissen. Daraus allein lässt sich jedoch nicht unmittelbar schlussfolgern, dass die drei Studien dieselben mathematischen bzw. naturwissenschaftlichen Kompetenzen messen.

3 Fragestellungen

In diesem Beitrag soll daher untersucht werden, ob die konzeptionellen und dimensionalen Voraussetzungen für ein Linking der Berichtsskalen aus den drei Studien in beiden Domänen gegeben sind. Folgende Forschungsfragen lassen sich ableiten:

- (1) Inwieweit sind die mathematischen und naturwissenschaftlichen Testkonzeptionen in PISA 2012, im LV 2012 und in NEPS-K9 konzeptionell vergleichbar?
- (2) Inwieweit ist die dimensionale Struktur für die latenten Konstrukte „mathematische Kompetenz“ und „naturwissenschaftliche Kompetenz“ in PISA 2012, im LV 2012 und in NEPS-K9 ähnlich?

4 Methode

4.1 Vergleichbarkeit der mathematischen und naturwissenschaftlichen Testkonzeptionen

4.1.1 Expertenreview der Testaufgaben in Mathematik

Für die Beantwortung der ersten Forschungsfrage wurden die 22 Mathematikaufgaben aus dem NEPS-Mathematiktest anhand eines Expertenreviews klassifiziert.¹ Drei in der Aufgabenkonstruktion erfahrene Expertinnen und Experten aus der Mathematikdidaktik wurden gebeten, alle 22 Items des NEPS-K9-Mathematiktests in die Rahmenkonzeptionen des LV-Mathematiktests 2012 und des PISA-Mathematiktests 2012 einzuordnen. Für diese Einordnung wurden Fragebögen erstellt, welche die Zuordnung der NEPS-Items zu den jeweiligen Dimensionen (Inhaltsbereiche, Prozesse, Anforderungsbereiche und Kontexte) vom LV 2012 und von PISA 2012 erfassen. Des Weiteren wurden umfangreiche Informationen zu den Hintergründen und Rahmenkonzeptionen der Studien zur Verfügung gestellt. Um zu überprüfen, inwieweit die Expertinnen und Experten in ihren Klassifikationen übereinstimmen, wurde die prozentuale Übereinstimmung zwischen den Ratern berechnet. Der Median der prozentualen Übereinstimmung zwischen den drei Ratern beträgt über alle Items 80 Prozent und ist damit als gut zu bewerten. Zusätzlich wurde das Cohens κ (Median der κ -Werte der drei Raterpaare) als zufallskorrigiertes Übereinstimmungs-

¹ Befunde in diesem Abschnitt und in Abschnitt 5.1.1 sind publiziert in van den Ham et al. (2014).

maß berechnet. Die κ -Werte liegen im Mittel bei $\kappa = 0.5$ und sind damit ebenfalls in einem akzeptablen Bereich (Wirtz & Caspar, 2002). Anschließend wurden die 22 NEPS-Aufgaben in die Teilbereiche der PISA- und LV-Rahmenkonzeptionen eingeordnet. Um dies zu ermöglichen, wurde zuerst für jedes NEPS-K9-Mathematikitem von drei Expertinnen und Experten eingeschätzt, in welchen Inhaltsbereich, Prozess, Anforderungsbereich und Kontext das Item in der PISA-Rahmenkonzeption und in welchen Inhaltsbereich, Prozess und Anforderungsbereich das Item in der LV-Rahmenkonzeption fallen würde. Die Items wurden schließlich dem Teilbereich der PISA-2012- bzw. LV-2012-Rahmenkonzeption zugeordnet, welcher von mindestens zwei Expertinnen und Experten für dieses Item eingeschätzt wurde. Auf diese Weise konnten alle Items den Teilbereichen zugeordnet werden. Die so entstehenden Verteilungen der NEPS-Aufgaben in den PISA- und LV-Rahmenkonzeptionen wurden den Verteilungen der NEPS-Aufgaben in den NEPS-Inhaltsbereichen gegenübergestellt. Aufgrund der nicht vorliegenden Zuordnung der NEPS-Aufgaben zu den NEPS-Prozessen konnte für diese kein Vergleich zwischen den Studien durchgeführt werden.

4.1.2 Expertenreview der Testaufgaben in den Naturwissenschaften

Im Bereich Naturwissenschaften wurden die 28 Items des NEPS-K9-Naturwissenschaftstests von sieben Expertinnen und Experten in die PISA- und LV-Rahmenkonzeptionen eingeordnet.² Da das Ziel der Studie darin bestand, NEPS-Items anhand ihrer naturwissenschaftlichen Inhalte einzuschätzen, war es wichtig, Personen mit einem entsprechenden fachlichen und methodischen Hintergrund zu finden. Dementsprechend wurden fünf Expertinnen und Experten mit einem fachdidaktischen Hintergrund, eine Person mit einem erziehungswissenschaftlichen und psychologischen Hintergrund und eine Person mit einem Lehramtshintergrund ausgewählt. Weiterhin war es wichtig, dass die ausgewählten Personen über umfangreiche Kenntnisse hinsichtlich mindestens einer der Rahmenkonzeptionen und eines Kompetenztests verfügten. So waren vier der Expertinnen und Experten besonders mit PISA vertraut und drei von ihnen mit dem LV.

Zur Einordnung der NEPS-Items erhielten die Expertinnen und Experten alle nötigen Hintergrundinformationen zu den Studien sowie ein Reviewsheet, anhand dessen sie die Zuordnung der NEPS-Items zu den jeweiligen Wissens- und Kompetenzbereichen der anderen Studien vornehmen konnten (Wagner et al., 2014). Um zu überprüfen, inwieweit die Expertinnen und Experten in ihren Klassifikationen übereinstimmen, wurde die sogenannte Generalisierbarkeitstheorie angewandt (Cronbach, Gleser, Nanda & Rajaratnam, 1972). Die Auswertung wurde mit dem Programm G-String IV (Bloch & Norman, 2011) sowie mit dem Programm IASGA (Li & Lautenschlager, 1999) vorgenommen. Der berechnete Generalisierbarkeitskoeffizient darf als Cohens κ interpretiert werden. Hierbei werden Werte ≥ 0.6 als substantielle Übereinstimmung zwischen den Ratern beurteilt (Landis & Koch, 1972). Bis auf

2 Befunde in diesem Abschnitt und in Abschnitt 5.1.2 sind publiziert in Wagner et al. (2014).

einen Bereich (Bildungsstandards, Bereich Komplexität), für den die Konsistenz bei gerade 0.6 liegt, zeigen alle Bereiche Konsistenzwerte zwischen 0.75 und 0.92, sodass die Expertenurteile in ihrer Übereinstimmung als substantiell bezeichnet werden können. Dieses Ergebnis stellt die Grundlage für die Beurteilung und Interpretation der weiteren Ergebnisse der Studie dar.

4.2 Dimensionale Zusammenhänge

Die Basis für die Analyse der dimensional Zusammenhänge bildeten Daten der Validierungsstudie, welche im Rahmen des Projektes im Frühjahr 2012 durchgeführt wurde.

Die Stichprobe bestand aus $N = 1.965$ Neuntklässlerinnen und Neuntklässlern aus 80 Schulen. Am ersten Testtag bearbeiteten $N = 1.679$ Schülerinnen und Schüler Mathematik- und Naturwissenschaftsaufgaben aus PISA. Am zweiten Testtag beantworteten $N = 825$ Testpersonen Mathematikaufgaben und $N = 862$ Naturwissenschaftsaufgaben aus dem Ländervergleich. Anschließend wurde der NEPS-Mathematiktest von $N = 1.330$ und der NEPS-Naturwissenschaftstest von $N = 1.335$ Schülerinnen und Schülern bearbeitet. Insgesamt nahmen $N = 634$ Schülerinnen und Schüler an allen drei Mathematiktests und $N = 673$ Schülerinnen und Schüler an allen drei Naturwissenschaftstests teil.

Die Zusammenhänge der mit den verschiedenen Tests erfassten mathematischen und naturwissenschaftlichen Kompetenzen wurden als latente Korrelationen jeweils in dreidimensionalen Rasch-Modellen mit der Software ConQuest (Wu, Adams & Haldane, 2007) geschätzt. In einem ersten Modell wurden die mathematische Kompetenz aus NEPS, die mathematische Kompetenz aus PISA und die mathematische Kompetenz aus dem LV auf jeweils einer separaten Dimension modelliert. In einem zweiten Modell wurden die naturwissenschaftlichen Kompetenzen aus NEPS, PISA und dem LV auf jeweils eigenen Dimensionen spezifiziert.

5 Ergebnisse

5.1 Ergebnisse zur konzeptionellen Äquivalenz der Mathematiktests aus NEPS-K9, PISA 2012 und dem LV 2012

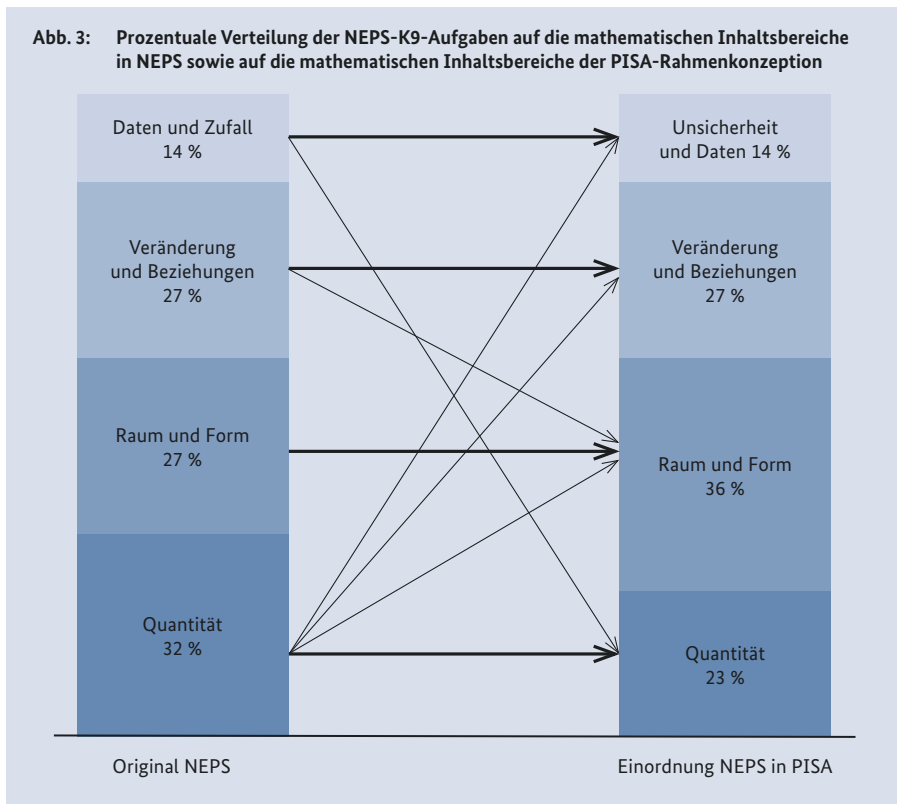
In einem ersten Schritt wurde die Vergleichbarkeit der Kompetenzmessungen in Mathematik und in den Naturwissenschaften in NEPS-K9, PISA 2012 und dem LV 2012 auf der konzeptionellen Ebene untersucht. Dazu wurden die Testkonzeptionen analysiert: a) hinsichtlich der Einordbarkeit der Aufgaben aus dem Mathematik- und Naturwissenschaftstest aus NEPS-K9 in die entsprechenden Teildimensionen der Mathematik- und Naturwissenschaftstests aus PISA 2012 und LV 2012 sowie b) hinsichtlich der Unterschiede zwischen den Einordnungen der NEPS-K9-Aufgaben in die Rahmenkonzeptionen aus NEPS-K9 sowie aus PISA 2012 bzw. des LV 2012.

5.1.1 Vergleich der mathematischen Testkonzeptionen aus NEPS-K9 und PISA 2012

Das Expertenreview zeigt, dass die Expertinnen und Experten alle Aufgaben aus NEPS-K9 sowohl in einen Inhaltsbereich, in mindestens einen mathematischen Prozess- und Anforderungsbereich als auch in einen Aufgabenkontext aus der PISA-Rahmenkonzeption einordnen können. Es gibt also keine NEPS-Aufgaben, die einen Inhaltsbereich, Prozess oder Kontext erfassen, welcher nicht in der PISA-Rahmenkonzeption definiert wird.

Die prozentuale Verteilung der NEPS-K9-Aufgaben in den Inhaltsbereichen aus NEPS unterscheidet sich nicht signifikant von der durch die Expertinnen und Experten vorgenommenen Einordnung der NEPS-Aufgaben in die gleichnamigen PISA-Inhaltsbereiche ($\chi^2 = 1.8$, $df = 3$, ns). Tendenziell werden jedoch einige Unterschiede zwischen den Inhaltsbereichen deutlich (siehe Abbildung 3). Der größte Unterschied befindet sich zwischen den Inhaltsbereichen „Quantität“ in NEPS und „Quantität“ in PISA. Die Expertinnen und Experten ordnen drei der sieben NEPS-Aufgaben aus dem Bereich „Quantität“ anderen PISA-Inhaltsbereichen zu. Ein weiterer Unterschied zeigt sich beim Inhaltsbereich „Raum und Form“. Hier werden zwar alle „Raum und Form“-Aufgaben aus dem NEPS-K9-Test auch dem Bereich „Raum und Form“ in PISA zugeordnet. Allerdings wird auch jeweils eine Aufgabe aus den NEPS-Inhaltsbereichen „Quantität“ und „Veränderung und Beziehungen“ dem PISA-Inhaltsbereich „Raum und Form“ zugeordnet. Die Aufgaben aus den NEPS-Inhaltsbereichen „Veränderung und Beziehungen“ und „Daten und Zufall“ werden größtenteils den gleichnamigen Bereichen in PISA zugeordnet, nur jeweils eine Aufgabe wird einem anderen Inhaltsbereich zugesprochen (vgl. van den Ham et al., 2014).

In den NEPS-K9-Mathematikaufgaben lassen sich außerdem fünf der sechs PISA-Prozesse identifizieren. Keine der 22 NEPS-Mathematikaufgaben wird von den Expertinnen und Experten dem Prozess „mathematische Hilfsmittel verwenden“ zugeordnet. Dieser Befund stimmt mit der NEPS-Rahmenkonzeption überein. In dieser wird explizit angegeben, dass mit dem NEPS-Test keine separate Kompetenz zum Umgang mit mathematischen Hilfsmitteln gemessen werden soll. Auf Basis des Expertenreviews werden zudem alle PISA-Anforderungsbereiche durch mindestens vier und alle Kontexte durch mindestens drei NEPS-Aufgaben abgedeckt.

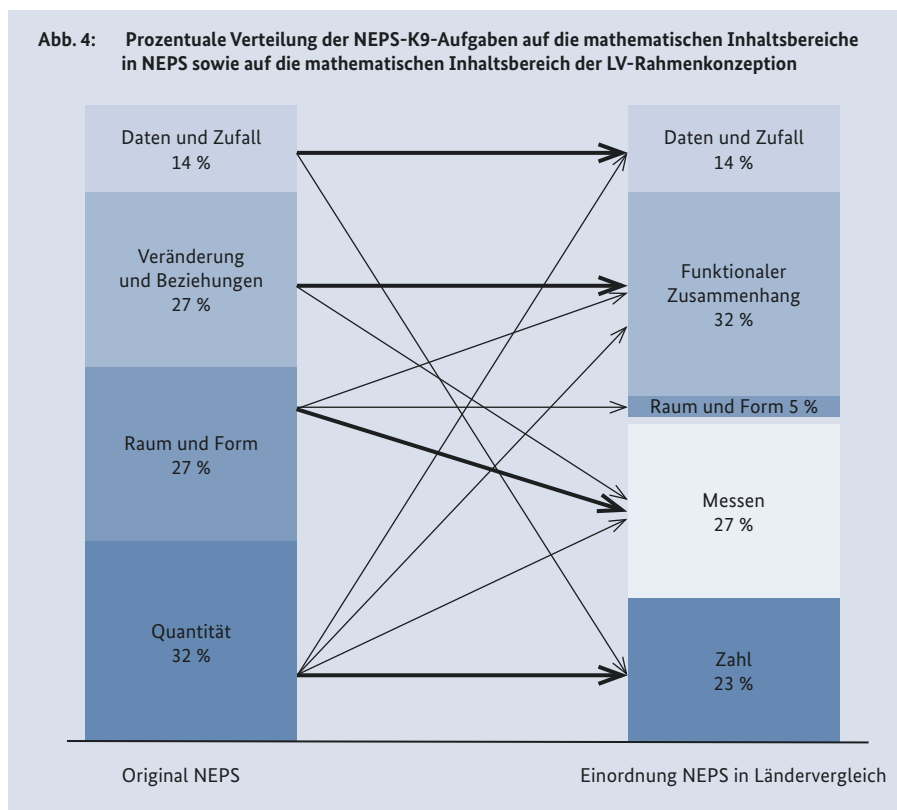


5.1.2 Vergleich der mathematischen Testkonzeptionen aus NEPS-K9 und LV 2012

Die Rahmenkonzeption des LV 2012 definiert einen Inhaltsbereich mehr als die NEPS-K9-Rahmenkonzeption. Die Inhalte dieses Bereiches „Messen“ fallen in der NEPS-Rahmenkonzeption unter den Inhaltsbereich „Quantität“. Bei der Einordnung der NEPS-Aufgaben in die LV-Rahmenkonzeption fällt jedoch auf, dass die meisten Aufgaben aus dem Bereich „Raum und Form“ von den Expertinnen und Experten in den Bereich „Messen“ der LV-Rahmenkonzeption eingeordnet werden. Bei diesen NEPS-Aufgaben müssen vor allem Flächen bzw. Winkel berechnet werden. Dementsprechend fallen im Vergleich weniger Aufgaben unter den Inhaltsbereich „Raum und Form“ nach der Definition des LV. Dem Inhaltsbereich „Quantität“ in NEPS werden zwei Aufgaben mehr zugeordnet als dem Inhaltsbereich „Zahl“. Die Inhaltsbereiche „Veränderung und Beziehungen“ aus NEPS und „Funktionaler Zusammenhang“ aus dem LV sowie „Daten und Zufall“ aus NEPS und „Daten und Zufall“ aus dem LV unterscheiden sich kaum voneinander. Die Unterschiede in der Zuordnung lassen sich jedoch nicht statistisch absichern (vgl. van den Ham et al., 2014).

In allen NEPS-K9-Mathematikaufgaben konnten die Expertinnen und Experten Prozesse und Anforderungsbereiche der LV-Rahmenkonzeption identifizieren. Dabei werden alle Prozesse aus der LV-Rahmenkonzeption auf Basis des Expertenreviews

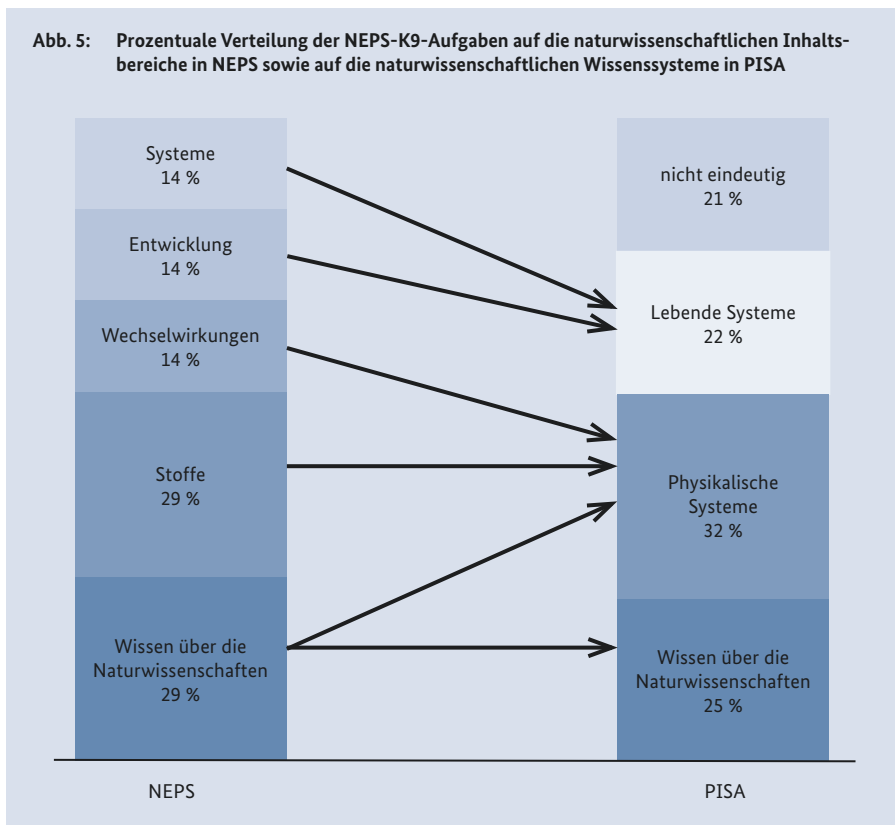
durch mindestens sechs Aufgaben abgedeckt. Die Anforderungsbereiche „Reproduzieren“ und „Zusammenhänge herstellen“ werden durch acht respektive 14 Aufgaben abgedeckt. Jedoch wurde keine der NEPS-Aufgaben dem Anforderungsbereich „Verallgemeinern und Reflektieren“ zugeordnet.



5.2 Vergleich der naturwissenschaftlichen Testkonzeptionen aus NEPS-K9 und PISA 2012

Das Expertenreview der Aufgaben des NEPS-K9-Naturwissenschaftstests zeigt, dass sich die NEPS-Items über die Wissenssysteme „lebende Systeme“, „physikalische Systeme“ und den Wissensbereich „Wissen über die Naturwissenschaften“ verteilen (siehe Abbildung 5). Die Zuordnungsrate der NEPS-Items zu den PISA-Wissensbereichen fällt mit 79 Prozent sehr hoch aus. Die in Abbildung 5 dargestellten Pfeile geben darüber Auskunft, wohin jeweils die Mehrzahl der Rater die meisten Items zugeordnet hat. Die Items der inhaltsbezogenen Komponenten „Systeme“ und „Entwicklung“ können vorwiegend in den PISA-Bereich „lebende Systeme“ eingeordnet werden. Weiterhin werden die Komponenten „Wechselwirkungen“ und „Stoffe“ beide vorwiegend dem PISA-Bereich der „physikalischen Systeme“ zugeordnet. Die Items der prozessbezogenen NEPS-Komponente „Wissen über die Naturwissenschaften“

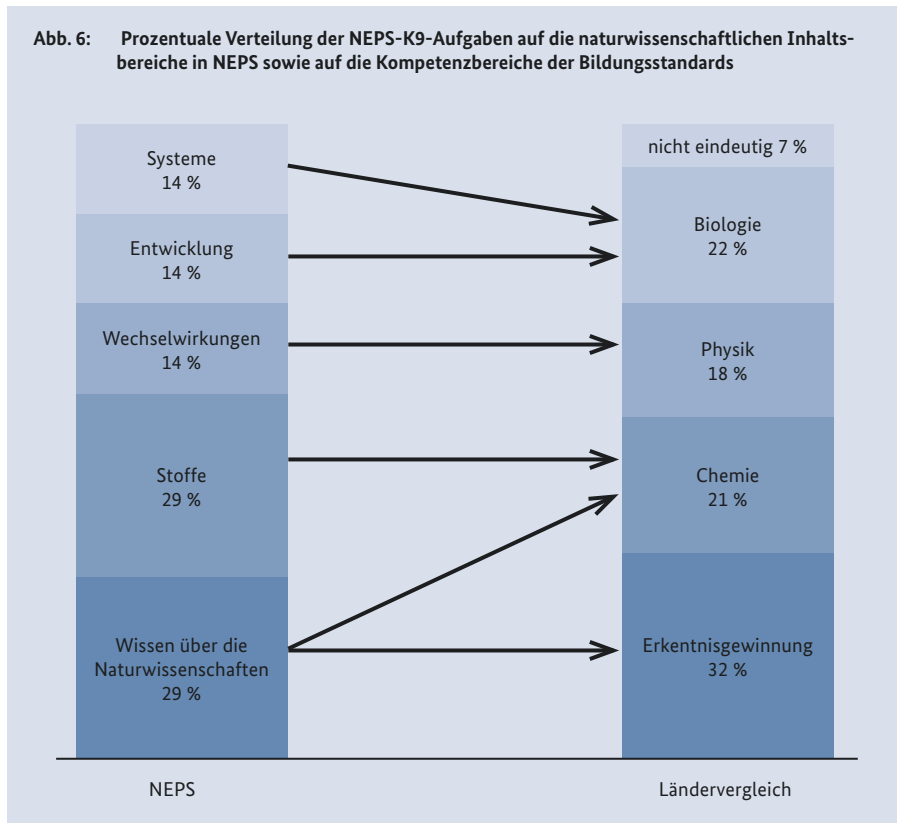
fallen sowohl in die „physikalischen Systeme“ als auch in den inhaltlich äquivalenten Bereich des „Wissens über die Naturwissenschaften“ in PISA. Hinsichtlich der Zuordnung der NEPS-K9-Items auf die Teilkompetenzen „naturwissenschaftliche Fragestellungen erkennen“, „naturwissenschaftliche Phänomene beschreiben, erklären und vorhersagen“ sowie „naturwissenschaftliche Evidenz nutzen, um Entscheidungen zu treffen“ konnte eine Zuordnungsrate von 96 Prozent erreicht werden. Die meisten NEPS-Items (67 Prozent der eingeordneten Items) werden dabei der Kompetenz „naturwissenschaftliche Phänomene erklären“ zugeordnet. Für die Mehrheit der NEPS-Komponenten (vier von fünf) ist eine eindeutige Zuordnung zu den Inhalten der Rahmenkonzeptionen von PISA möglich. Dabei werden die NEPS-Konzepte jeweils meist ähnlich bezeichneten Kompetenz- und Wissensbereichen zugeordnet. Für den NEPS-Inhaltsbereich „Systeme“ fällt eine eindeutige Einordnung bezüglich der PISA-Rahmenkonzeption schwer.



5.2.1 Vergleich der naturwissenschaftlichen Testkonzeptionen aus NEPS-K9 und dem LV 2012

Vier und mehr der sieben Rater, die die NEPS-K9-Items einordneten, konnten 93 Prozent der Items einem bestimmten Teilbereich des „Umgangs mit Fachwissen“ oder der „Erkenntnisgewinnung“ zuordnen.

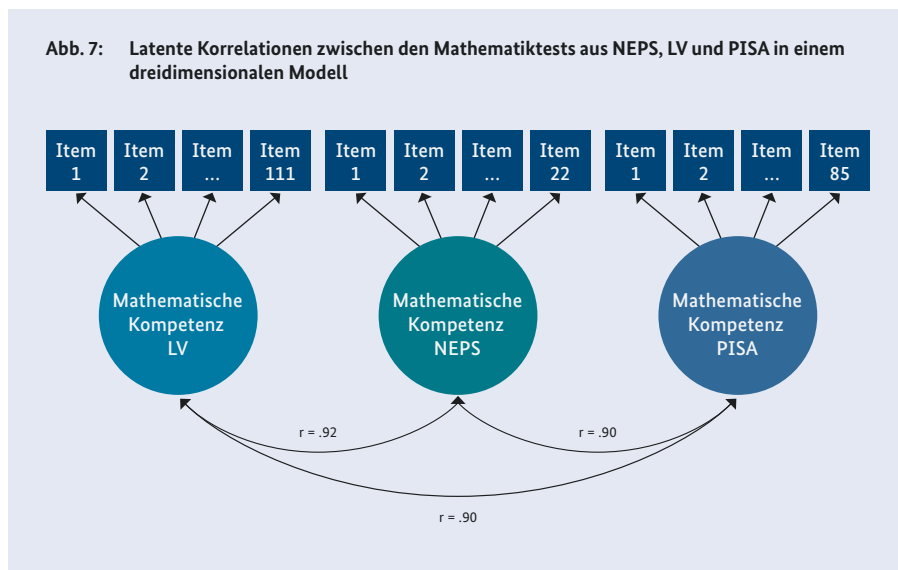
Die Pfeile in Abbildung 6 geben erneut die Kompetenzbereiche wieder, in welche die Mehrzahl der Rater die meisten Items zugeordnet hat. Insgesamt verteilen sich die NEPS-Items relativ gleichmäßig auf die Kompetenzbereiche der Bildungsstandards. Lediglich der Fachbereich „Physik“ ist leicht unterrepräsentiert. Analog zur Zuordnung in die PISA-Rahmenkonzeption werden die NEPS-Items der inhaltsbezogenen Komponenten „Systeme“ und „Entwicklung“ dem Fachbereich „Biologie“ zugeordnet. Die Items der NEPS-Komponente „Wechselwirkungen“ können eindeutig dem Fachbereich Physik und die Items der NEPS-Komponente „Stoffe“ ebenso eindeutig dem Fachbereich „Chemie“ zugeordnet werden. Im Hinblick auf die kognitiven Prozesse werden 90 Prozent der zugeordneten NEPS-Items den Prozessen „Organisieren“ und „Integrieren“ zugeordnet. Die Zuordnungsrate zur Dimension „kognitive Prozesse“ liegt insgesamt bei 75 Prozent. Auf die Dimension „Komplexität“ können 64 Prozent der NEPS-Items zugeordnet werden. Mit 72 Prozent liegt der Fokus der Items auf der Stufe „ein Zusammenhang“. Auch bei der Einordnung der NEPS-Aufgaben in die Rahmenkonzeption des LV gilt, dass für vier der fünf NEPS-Komponenten eine eindeutige Zuordnung zumeist ähnlich bezeichneter Kompetenz- und Wissensbereiche möglich war. Eine eindeutige Zuordnung des NEPS-Inhaltsbereiches Systeme bezüglich der LV-Konzeption fällt jedoch schwer.



5.3 Ergebnisse zur dimensionalen Äquivalenz

5.3.1 Zusammenhänge zwischen den mathematischen Kompetenztests aus PISA 2012, LV 2012 und NEPS-K9

Die aus der dreidimensionalen Modellierung entstehenden Zusammenhänge zwischen den Mathematiktests aus NEPS, LV und PISA werden in Abbildung 7 dargestellt. Die Höhe der Korrelationen deutet auf einen substantziellen Zusammenhang zwischen den Tests. Die latenten Korrelationen zwischen den drei Mathematikskalen fallen in vergleichbarer Höhe aus.

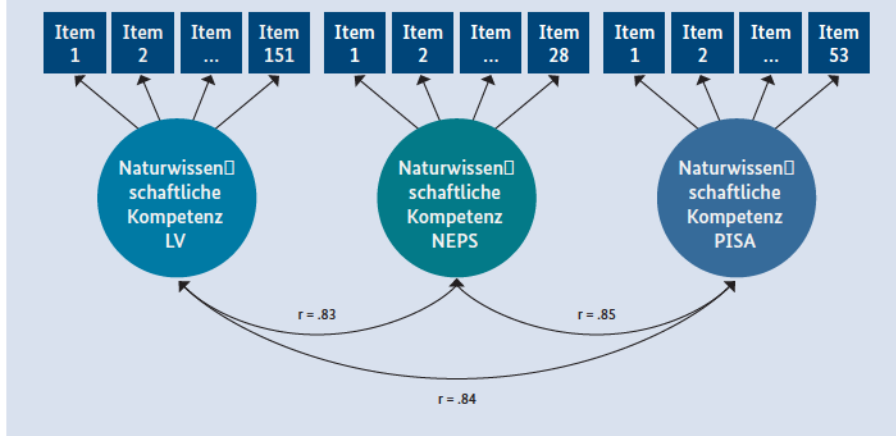


5.3.2 Zusammenhänge zwischen den Naturwissenschaftstests aus PISA 2012, LV 2012 und NEPS-K9

Um den korrelativen Zusammenhang der drei Naturwissenschaftstests aus dem Ländervergleich, aus PISA 2012 und aus NEPS-K9 zu bestimmen, wurde ebenfalls eine dreidimensionale Skalierung durchgeführt, in der jeder Kompetenztest als eigene Dimension angelegt wurde. Die Ergebnisse in Abbildung 8 zeigen substantzielle Korrelationen zwischen den drei Naturwissenschaftsskalen. Die Korrelation zwischen dem NEPS-Test und dem Ländervergleichstest ist in vergleichbarer Höhe wie die Korrelation zwischen dem NEPS-Test und dem PISA-Test.

Im Vergleich der Korrelationsmuster zwischen den Domänen Mathematik und Naturwissenschaften kann festgehalten werden, dass die Korrelationsmuster in Mathematik geringfügig höher ausfallen als in den Naturwissenschaften.

Abb. 8: Latente Korrelationen zwischen den drei Naturwissenschaftstests aus dem Ländervergleich, aus PISA und aus NEPS



6 Zusammenfassung und Diskussion

Dieser Beitrag hatte zum Ziel, die konzeptionelle und dimensionale Äquivalenz der in NEPS-K9, im LV 2012 und in PISA 2012 erfassten mathematischen und naturwissenschaftlichen Kompetenzen zu untersuchen. Dafür wurde geprüft, (1) inwieweit die mathematischen und naturwissenschaftlichen Testkonzeptionen in PISA 2012, in den länderübergreifenden Bildungsstandards und in NEPS konzeptionell vergleichbar sind und (2) inwieweit die dimensionale Struktur für die latenten Konstrukte „mathematische Kompetenz“ und „naturwissenschaftliche Kompetenz“ der Tests in PISA 2012, in den länderübergreifenden Bildungsstandards und in NEPS vergleichbar ausfällt.

Hinsichtlich der konzeptionellen Äquivalenz kann festgehalten werden, dass es große Übereinstimmungen zwischen den Testkonstrukten gibt. Für den NEPS-K9-Mathematiktest gilt, dass alle Aufgaben in das PISA- und LV-Rahmenkonzept eingeordnet werden können. Im NEPS-K9-Mathematiktest werden keine Aufgaben eingesetzt, die nicht diesen Rahmenkonzepten entsprechen. Lediglich ein Prozess der PISA-Rahmenkonzeption und ein Anforderungsbereich der LV-Rahmenkonzeption werden nicht durch die NEPS-Aufgaben repräsentiert. Bezüglich des NEPS-Naturwissenschaftstests kann festgehalten werden, dass für die Mehrheit der NEPS-Komponenten eine eindeutige Zuordnung zu den Inhalten der PISA- und LV-Rahmenkonzeption möglich ist. Die Zuordnungsrate von mindestens 79 Prozent zu den Kompetenz- bzw. Wissensbereichen der LV- und PISA-Rahmenkonzeption deutet auf eine hohe Vergleichbarkeit der theoretischen Konstrukte.

Die korrelativen Zusammenhänge bestätigen sowohl für die mathematische als auch für die naturwissenschaftliche Kompetenz eine deutliche Überschneidung zwischen den Tests. Dabei lassen sich die größeren Übereinstimmungen zwischen den mathematischen Rahmenkonzeptionen in höheren Korrelationen zwischen den mathematischen Konstrukten ($.90 \leq r \leq .92$) als zwischen den naturwissenschaftlichen Konstrukten ($.83 \leq r \leq .85$) wiederfinden.

Dennoch weisen die gefundenen Unterschiede zwischen den Rahmenkonzepten und auch die Höhe der Korrelationen darauf hin, dass es sich nicht um austauschbare Konstrukte handelt. Nach Kolen und Brennan (2010) ist jedoch auch eine Verknüpfung von unterschiedlichen Tests möglich. Die gefundenen Übereinstimmungen dieser Studien bilden eine wichtige Voraussetzung für eine mögliche Verlinkung der Berichtsskalen dieser Studien. Die aufgedeckten Unterschiede weisen auf die Relevanz tiefer gehender Analysen hin, da diese zu Unterschieden in den Testergebnissen und somit in den Ergebnissen eines Linking führen können (Kolen & Brennan, 2010). Es ist daher von Belang, die faktorielle Struktur der Tests noch intensiver auf ihre Vergleichbarkeit zu prüfen, z. B. durch einen Vergleich der Korrelationen zwischen den Subdimensionen innerhalb der Tests und zwischen den Tests. Ein weiterer wichtiger Schritt wäre es, zudem die Gruppeninvarianzen und deren Einfluss auf mögliche Entschlüsse, basierend auf einem Linking (Kolen & Brennan, 2010), zu analysieren. Die gefundenen Unterschiede in der sprachlichen Schwierigkeit der Tests (van den Ham et al., 2014) weisen vor allem auf die Wichtigkeit einer Analyse möglicher Invarianzen bezüglich der Subgruppe von Schülerinnen und Schülern mit Deutsch als Zweitsprache. Nur bei hinreichender Äquivalenz auf dimensionaler und skalenbezogener Ebene ist eine Verankerung der Skalen sinnvoll zu interpretieren. Zusammenfassend kann geschlossen werden, dass eine Verlinkung der mathematischen und naturwissenschaftlichen Berichtsskalen möglich ist, sofern die gefundenen Differenzen sowie mögliche Linkingfehler berücksichtigt werden, eine hierfür angemessene Linking-Methode gewählt wird und adäquate Interpretationen getätigt werden.

Literaturverzeichnis

- American Association for the Advancement of Science (1993). *Benchmarks for science literacy. Project 2061*. New York, NY: Oxford University Press.
- American Association for the Advancement of Science (2009). *Benchmarks for science literacy. Project 2061*. New York, NY: Oxford University Press. Abgerufen am 16.07.2015 von <http://www.project2061.org/publications/bsl/online/index.php>.
- Bloch, R. & Norman, G. (2011). *G String IV (Version 6.1.1). User Manual*. Abgerufen am 16.07.2015 von http://fhspcrd.mcmaster.ca/g_string/download/g_string_4_manual_611.pdf.
- Blossfeld, H.-P. (2008). *Education as a Lifelong Process. A Proposal for a National Educational Panel Study (NEPS) in Germany. Part B: Theories, Operationalizations and Piloting Strategies for the Proposed Measurements*. Unveröffentlichter BMBF-Antrag. Bamberg: Universität Bamberg.
- Blossfeld, H.-P., Maurice, J. von & Schneider, T. (2011). The National Educational Panel Study: need, main features, and research potential. *Zeitschrift für Erziehungswissenschaft*, 14, 5–18.
- Cronbach, L., Gleser, G. C., Nanda, H. & Rajaratnam, N. (1972). *The dependability of behavioral measurement. Theory of generalizability for scores and profiles*. New York, NY: Wiley.

- Ehmke, T., Duchhardt, C., Geiser, H., Grüßing, M., Heinze, A. & Marschick, F. (2009). Kompetenzentwicklung über die Lebensspanne – Erhebung von mathematischer Kompetenz im Nationalen Bildungspanel. In A. Heinze & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (S. 313–327). Münster: Waxmann.
- Ehmke, T., Köller, O., Nissen, A. & van den Ham, A.-K. (2014). Äquivalenz von Kompetenzmessungen in Schulleistungsstudien. *Unterrichtswissenschaft*, 42 (4), 290–300.
- Hahn, I., Schöps, K., Rönnebeck, S., Martensen, M., Hansen, S., Saß, S., Delahefte, J. M. & Prenzel, M. (2013). Assessing scientific literacy over the lifespan – A description of the NEPS science framework and the test development. *Journal for Educational Research Online*, 5 (2), 110–138.
- Kolen, M. J. & Brennan, R. L. (2010). *Test equating, scaling, and linking: Methods and practices*. New York, NY: Springer.
- Köller, O. (2008). Bildungsstandards in Deutschland: Implikationen für die Qualitätssicherung und Unterrichtsqualität. In M. A. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), *Perspektiven der Didaktik* (S. 47–59). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Landis J. R. & Koch G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Li, M. N. & Lautenschlager, G. J. (1999). IASGA: a SAS MACRO program for interrater agreement studies of qualitative data via generalizability approach. *Educational and Psychological Measurement*, 59 (3), 532–537.
- Neubrand, M., Biehler, R., Blum, W., Cohors-Freseborg, E., Flade, L., Knoche, N., Lind, D., Löding, W., Möller, G. & Wynands, A. (2001). Grundlagen der Ergänzung des internationalen PISA-Mathematiktests in der deutschen Zusatzhebung. *Zentralblatt für Didaktik der Mathematik – Berichtsteil*, 33, 45–59.
- Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatsis & S. Papastavridis (Hrsg.), *3rd Mediterranean Conference on Mathematical Education. Athens – Hellas 3–5 January 2003* (S. 115–124). Athen: The Hellenic Mathematical Society.
- Organisation for Economic Co-operation and Development (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*. Paris: OECD.
- Organisation for Economic Co-operation and Development (2013a). *PISA 2006 Science Competencies for Tomorrow's World*. Paris: OECD.
- Organisation for Economic Co-operation and Development (2013b). *What students know and can do: Student performance in mathematics, reading and science* (Bd. I, überarbeitete Auflage, Februar 2014). Paris: OECD.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.). (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (Hrsg.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (Hrsg.). (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

- Roppelt, A., Blum, W. & Pöhlmann, C. (2013). Beschreibung der untersuchten mathematischen Kompetenzen. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 23–37). Münster: Waxmann.
- Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, O. Köller & E. Klieme (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 155–187). Münster: Waxmann.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. Neuwied: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. Neuwied: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004b). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. Neuwied: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004c). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. Neuwied: Luchterhand.
- van den Ham, A.-K., Nissen, A., Ehmke, T., Sälzer, C. & Roppelt, A. (2014). Mathematische Kompetenz in PISA, IQB-Ländervergleich und NEPS – Drei Studien, gleiches Konstrukt? *Unterrichtswissenschaft, 42* (4), 321–341.
- van de Vijver, F. J. R. (1998). Towards a theory of bias and equivalence. In J. A. Harkness (Hrsg.), *ZUMA-Nachrichten Spezial* (Bd. 3, S. 41–65). Mannheim: ZUMA.
- Wagner, H., Schöps, K., Hahn, I., Pietsch, M. & Köller, O. (2014). Konzeptionelle Äquivalenz von Kompetenzmessungen in den Naturwissenschaften zwischen NEPS, IQB-Ländervergleich und PISA. *Unterrichtswissenschaft, 42* (4), 301–320.
- Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T. & Carstensen, C. H. (2011). Development of competencies across the life span. *Zeitschrift für Erziehungswissenschaft, 14*, 67–86.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Wu, M., Adams, R. & Haldane, S. (2007). *ConQuest [computer software]*. Melbourne: Australian Council for Educational Research.