

Interdependentes Entscheiden in der Unterrichtsplanung – Entwicklung eines generischen Rahmenmodells und eines vignettenbasierten Fähigkeitstests

Daniel Scholl · Simon Küth  · Christoph Schüle

Eingegangen: 15. September 2021 / Überarbeitet: 25. April 2022 / Angenommen: 23. Juni 2022 / Online publiziert: 11. August 2022
© Der/die Autor(en) 2022

Zusammenfassung Didaktische Ansätze, Studien des Decision-Making-Paradigmas und gegenwärtige Kompetenzmodelle erachten die wechselseitige Verknüpfung aller Planungsentscheidungen (Interdependenz) als wichtige Anforderung an die Unterrichtsplanung. Ein generisches Rahmenmodell und einen Test für interdependentes Planungsentscheiden gibt es allerdings noch nicht. Deshalb wird in diesem Beitrag zunächst ein zweidimensionales Rahmenmodell vorgeschlagen, in dem die entsprechende kognitive Operation (1) formal als konditionales, begrenztes und potenziell unbegrenztes interdependentes Entscheiden und (2) inhaltlich als wechselseitige Verknüpfung intentionalen, thematischen, methodischen und medialen Entscheidens spezifiziert wird. Anschließend wird die Entwicklung eines raschhomogenen, standardisierten Vignettentests mit akzeptablen Reliabilitäten und ersten Validierungsbefunden berichtet. Die Pilotierungsstichprobe umfasst $N=547$ Teilnehmer*innen, davon 313 angehende und praktizierende Lehrkräfte.

Schlüsselwörter Unterrichtsplanung · Professionelle Kompetenz · Mehrdimensionales Entscheidungsmodell · Raschhomogener Planungstest

Prof. Dr. Daniel Scholl · Simon Küth (✉)
Erziehungswissenschaft mit dem Schwerpunkt Schulpädagogik, Universität Siegen,
Adolf-Reichwein-Straße 2a, 57068 Siegen, Deutschland
E-Mail: simon.kueth@uni-siegen.de

Prof. Dr. Daniel Scholl
E-Mail: daniel.scholl@uni-siegen.de

Dr. Christoph Schüle
Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung,
Keßlerstraße 42, 31134 Hildesheim, Deutschland
E-Mail: christoph.schuele@nlq.niedersachsen.de

Interdependent planning decisions in lesson planning—Development of a generic framework and a corresponding vignette test

Abstract Didactic approaches, studies of the decision-making paradigm, and current competency models consider the reciprocal linkage of all planning decisions (interdependence) as an important requirement for instructional planning. However, a generic framework model and corresponding test for interdependent planning decision-making does not yet exist. Therefore, this paper first proposes a two-dimensional framework model in which the corresponding cognitive operation is specified in two ways; (1) formally as conditional, bounded, and potentially unbounded interdependent decision making and, (2) substantively as reciprocal linking of intentional, thematic, methodological, and medial decision-making. Subsequently, we report the development of a Rasch-homogeneous, standardized vignette test with acceptable reliabilities and indicators of its validity, piloted on a sample of $N=547$ participants, where 313 were prospective and practicing teachers.

Keywords Lesson planning · Professional competence · Multi-dimensional decision-making model · Rasch-homogeneous planning test

1 Einleitung

Die erlernbare, professionelle Unterrichtsplanung (König et al. 2015) gilt als essenzielle Voraussetzung für einen lernwirksamen Unterricht (Yinger und Hendricks-Lee 1995). Denn seit den didaktischen Planungsansätzen der 1950er-Jahre (Scholl 2018) wird angenommen, dass die vorab getroffenen Planungsentscheidungen zur Qualitätsentwicklung der folgenden Unterrichtsdurchführung beitragen können (Wengert 1989). Ein solcher Beitrag wird diesen Entscheidungen unter anderem dann unterstellt, wenn sie die „zentrale generische Anforderung“ (König et al. 2015, S. 378) der Interdependenz erfüllen (Küth et al. 2021). Im Sinne dieser Anforderung, die inzwischen zu einem der bedeutendsten Metakriterien der Unterrichtsplanung geworden ist (Vogelsang und Riese 2017), soll in der Unterrichtsplanung eine „widerspruchsfreie Wechselwirkung der Planungsmomente“ (Schulz 1972, S. 45) durch die wechselseitige Abstimmung von Ziel-, Inhalts-, Methoden und Medienentscheidungen und deren jeweiligen Verschränkungen mit den situativen und individuellen Unterrichtsvoraussetzungen (z. B. mit dem Vorwissen oder den Kompetenzvorgaben des Kerncurriculums) hergestellt werden.

Ausdrücklich zum Gegenstand der empirischen Beschreibung von Planungsoperationen wurde das interdependente Entscheiden im Decision-Making-Paradigma der 1970er und 80er-Jahre (z. B. Bromme 1981; Shavelson und Stern 1981; Yinger 1980). Auch gegenwärtig wird die Anforderung der Interdependenz (bzw. Adaptivität als Teilaspekt) in die fachübergreifenden (s. zur Übersicht Rothland 2021) und fachbezogenen (z. B. Sachunterricht: Kirsch 2020; Physik: Schröder et al. 2020) Kompetenzmodelle zur Unterrichtsplanung aufgenommen.

Trotz der weiten Verbreitung dieser Planungsanforderung wurde das interdependente Planungsentscheiden bisher noch nicht näher in einem Rahmenmodell zu

kennzeichnen und über einen spezifischen Test zu operationalisieren versucht, um es einer gründlicheren empirischen Untersuchung zugänglich zu machen. Eine solche Untersuchung wäre aber wichtig, um bekannte Probleme der Planungspraxis von (angehenden) Lehrkräften genauer aufzuklären, von der die „empirische Forschung zeigt [...], dass u. a. die Interdependenz didaktischer Entscheidungen oft nur unzureichend berücksichtigt wird“ (Schrader und Schöb 2016, S. 331; Koeppen 1998; Mutton et al. 2011). Außerdem könnte eine theoretische und empirische Ausschärfung der Interdependenzanforderung helfen, dem Desiderat der empirisch zu wenig gestützten Berechtigung dieser Anforderung zu begegnen. Deshalb wird in diesem Beitrag ein Rahmenmodell interdependenten Entscheidens theoretisch hergeleitet und die Pilotierung und Validierung eines darauf basierenden Vignettentests berichtet. Ein Schwerpunkt liegt dabei auf der Prüfung der faktoriellen Validität als Teilaspekt von Konstruktvalidität (Jenßen et al. 2015) über die dimensionale Struktur eines multidimensionalen within-item Raschmodells.

2 Theoretischer Hintergrund

2.1 Interdependentes Planungsentscheiden in aktuellen Planungskompetenzmodellen

In den vorliegenden Planungskompetenzmodellen wird die Unterrichtsplanung zu den „*kontextspezifische[n] kognitive[n] Leistungsdispositionen*“ gezählt, „die sich funktional auf Situationen und Anforderungen in bestimmten *Domänen* beziehen“ (Klieme und Leutner 2006, S. 879). So modellieren z. B. König et al. (2017) die Planungskompetenz in Anlehnung an Blömeke et al. (2015) als wissensbasierte situationsspezifische Fähigkeit des Wahrnehmens von Planungsoptionen, deren Interpretation und Entscheidung. Die in diesem Modell fokussierten adaptiven Planungsentscheidungen, die sich auf „verschiedene Momente der Planung“ (König et al. 2015, S. 377) wie die Ziele, Inhalte, Methoden oder Medien beziehen können, werden als „ein fachunspezifischer Aspekt der Planungskompetenz“ (König et al. 2015, S. 377) und als Spezialfall interdependenten Entscheidens aufgefasst (König et al. 2017).

Stellvertretend lässt sich an diesem Modell verdeutlichen, dass das Planungsentscheiden einerseits auf einem bestimmten inhaltlichen Konzeptions- und Planungswissen als kognitiver Leistungsdisposition beruht (Baumert und Kunter 2006) und dementsprechend eine *inhaltliche Dimension* hat. In seinen fachübergreifenden, konzeptuellen Anteilen beinhaltet dieses Wissen beispielsweise allgemein- und fachdidaktische Planungsansätze (Schüle et al. 2017). Für die konkrete Anwendung dieser Ansätze umfasst dieses Wissen außerdem situative Anteile, z. B. die Kenntnis des lernbereichsspezifischen Vorwissens von Schüler*innen zu einem bestimmten Thema (De Jong und Ferguson-Hessler 1996). Andererseits wird an diesem Modell deutlich, dass dieses Wissen erst durch die wechselseitige *Verknüpfung* der jeweiligen Wissensanteile funktional für die situationsbezogene Unterrichtsplanung genutzt werden kann und das Planungsentscheiden folglich eine *formale Dimension* aufweist. Diese Dimension bezieht sich explizit auf die Fähigkeit zur interdepen-

denen Inbeziehungsetzung einzelner Planungselemente. So wären beispielsweise für die Planung einer bestimmten Unterrichtsstunde die Methodenentscheidungen auf die Lernvoraussetzungen der Schüler*innen abzustimmen.

In bisherigen fachübergreifenden (z. B. Beck et al. 2008; König et al. 2020; Rey et al. 2018) und fachbezogenen (z. B. Kirsch 2020; Schröder et al. 2020) Planungsmodellen und -tests werden allerdings erst ausgewählte inhaltliche Entscheidungsaspekte (insbesondere Ziele bezogen auf Lernvoraussetzungen) oder das grundsätzliche Wissen über die Interdependenz (z. B. Schüle et al. 2017) berücksichtigt. Formal sind die Besonderheiten interdependenten Entscheidens in diesen Arbeiten kaum weiter differenziert. Methodisch wurden die Modelle und Tests zudem überwiegend nicht mit Stichproben praktizierender Lehrkräfte entwickelt, sondern mit Studierenden (Rothland 2021). Jene Stichproben wären aber ebenfalls bedeutsam, weil didaktische Planungsansätze in der Planungspraxis von einigen Lehrkräften nicht (Haas 1998) oder abhängig von deren berufsbiografischer Sozialisation (John 1991) aufgegriffen werden und die Planungskompetenz samt Planungsroutinen vermutlich lebenslang weiterentwickelt wird (Munthe und Conway 2017).

Aus diesen Gründen werden die inhaltliche und formale Dimension interdependenten Planungsentscheidens im Folgenden näher bestimmt, bevor ein Vignettentest mit entsprechenden Indikatoren vorgestellt und dessen Pilotierung an einer Stichprobe berichtet wird, die auch praktizierende Lehrkräfte umfasst.

2.2 Zur formalen Dimension des interdependenten Entscheidens

In einem weiten Sinne lässt sich Planen als konstruktiver Prozess des antizipierenden Reflektierens komplexer Interdependenzen von Entscheidungsoptionen unter Berücksichtigung jeweils relevanter Rahmenbedingungen zum Entwerfen von künftigen Handlungsmöglichkeiten definieren (Funke und Fritz 1995; Grunwald 2000; Mumford et al. 2001). Auf dieser Begriffsgrundlage kann die Unterrichtsplanung im engeren Sinne als Prozess der „anticipatory reflection“ (Conway 2001, S. 89) verstanden werden, in dem eine Lehrkraft „visualizes the future, inventories means and ends, and constructs a framework to guide his or her future action“ (Clark und Peterson 1986, S. 260).

Entsprechend kognitionspsychologischer Erkenntnisse des Decision-Making-Paradigmas umfasst das Planen als „instructional decisions made prior to the execution of plans during teaching“ (Sardo-Brown 1996, S. 519) zwei Stufen (Yinger 1980): Auf eine problemorientierte initiale Planungsidee und die Identifikation von Planungsoptionen hin (Stufe 1: problem-finding stage) folgt deren sukzessive Elaboration, gedankliche Testung sowie Modifikation zu einem Gesamtplan als zirkuläre und gestaffelte Abfolge von wechselwirkenden Planungsentscheidungen (Stufe 2: problem formulation and solution stage). Interdependentes Entscheiden findet folglich auf dieser zweiten Stufe statt.

Formal lässt sich die Operation des interdependenten Entscheidens als die Integration wenigstens zweier Entscheidungsoptionen (zum Beispiel zu Zielen und Methoden) definieren, die jeweils aus Entscheidungsalternativen gewählt und unter Berücksichtigung der Unterrichtssituation, der Lernvoraussetzungen der Schüler*innen sowie der angenommenen Effekte dieser Optionen wechselseitig aufeinander bezo-

gen wurden (Fox et al. 2013; Shavelson und Stern 1981). Diese Operation kann in drei formale Fähigkeitsdimensionen (FD) differenziert werden, die den Zukunfts- (FD 1.1) und Wechselbezug (FD 1.2 und 1.3) des Entscheidens betreffen:

- *konditional interdependentes Entscheiden* (Fähigkeitsdimension 1.1): Interdependenz heißt hier, die Folgen mindestens zweier verknüpfter Planungsentscheidungen in der Form vereinfachter Konditionalschlüsse (wenn a, dann b; von Wright 1974) antizipieren zu können.
- *begrenzt interdependentes Entscheiden* (FD 1.2): Interdependenz im Sinne einer dynamischen (Westerman 1991) und relationalen Denkstruktur (Wing-mui So 1997) bedeutet in dieser Dimension, eine begrenzte Menge von Planungsoptionen im Planungsentscheiden untereinander und mit den identifizierten situativen Unterrichts- und individuellen Lernbedingungen gedanklich zu einem kohärenten Teilplan verbinden zu können.
- *(potenziell) unbegrenzt interdependentes Entscheiden* (FD 1.3): Interdependenz in dieser Dimension als „complex interrelated decisions based on a series of reasons“ (Tang und Watkins 1994, S. 41) entspricht der generischen Interdependenzanforderung vollständig (Vogelsang und Riese 2017). Eine potenziell unbegrenzte Menge von Planungsoptionen kann im Planungsentscheiden durch Verknüpfungen untereinander und mit den identifizierten Unterrichts- und Lernbedingungen gedanklich zu einem Gesamtplan verbunden werden, „so that the whole has a coherent structure and meaning“ (Tang und Watkins 1994, S. 41).

2.3 Zur inhaltlichen Dimension des interdependenten Entscheidens

Übereinstimmend mit Shulmans (1987) Annahme werden Planungsentscheidungen auf der Grundlage eines spezifischen, allerdings teilweise impliziten (Stender 2014) Planungswissens mit fachlichen (z. B. Lui und Bonner 2016), fachdidaktischen und fachübergreifenden Bestandteilen (z. B. König et al. 2020) getroffen. Dieses Wissen gibt der formalen Dimension interdependenten Entscheidens gleichermaßen seine inhaltliche Richtung, wie es das Planen als „bounded rationality“ (Shavelson und Stern 1981, S. 456) einschränkt. Als solche fachübergreifenden und fachdidaktischen Bestandteile gelten national und international in beinahe allen didaktischen Planungsansätzen die Unterrichtsziele, -inhalte, -aktivitäten bzw. -methoden, -materialien bzw. -medien sowie die Unterrichts- und Lernbedingungen (Friesen 2010; Shavelson und Stern 1981; Zierer und Seel 2012). Obwohl Entscheidungen in der Planungspraxis oft nur unzureichend auf Details einzelner Planungsansätze bezogen werden (Bromme 1981; Zahorik 1975), lässt sich auch dort eine Systematik weniger und wechselwirkender Planungsbereiche empirisch identifizieren (z. B. Tillema 1984 oder jüngst König et al. 2021), die den angedeuteten übergreifenden Bereichen der Planungsansätze entspricht und als kategoriales Ordnungswissen für das Planungsentscheiden dient.

Unter Berücksichtigung der Schnittmengen in den Planungsmodellen und der empirischen Befunde zu praxisbedeutsamen Planungsbereichen können vier inhaltliche Fähigkeitsdimensionen des interdependenten Planungsentscheidens in einem fachübergreifenden und fachdidaktisch anschlussfähigen Rahmenmodell unterschied-

den werden. Interdependenz bedeutet hier, Entscheidungen zu Lehr-Lernaktivitäten im Wechselverhältnis der unterschiedenen Planungsbereiche zu treffen und diese Entscheidungen an die Unterrichts- und Lernvoraussetzungen zu adaptieren:

- *intentionales Entscheiden* (FD 2.1): Unterrichtliche Lehr-Lernaktivitäten können an unterschiedlich abstrakten Lehr- und Lernzielen (Wittwer et al. 2020) für verschieden lange Planungszeiträume (Yinger 1980) ausgerichtet und als Teil eines (normativen) curricularen Handlungsrahmens institutionalisierten Unterrichts (Neumann 2020; Mutton et al. 2011) ausgewiesen werden.
- *thematisches Entscheiden* (FD 2.2): Themen der Lehr-Lernaktivitäten können durch die „themenkonstituierende Funktion von Fragestellungen“ (Klafki 1996, S. 260) begründet ausgewählt, in ihren Elementen konkretisiert und in ihrem Ertrag als exemplarisch für einen übergeordneten thematischen Zusammenhang bestimmt werden (Rawson et al. 2015).
- *methodisches Entscheiden* (FD 2.3): Lehr-Lernaktivitäten können in Form von Handlungsskripten (Borko und Livingston 1989; Stender et al. 2017) in eine sinnvolle ziel- und lernvoraussetzungsbezogene zeitliche Abfolge gebracht werden (König et al. 2021).
- *mediales Entscheiden* (FD 2.4): Lehr-Lernaktivitäten können zielführend kommunikativ durch ausgewählte bzw. selbstgestaltete Medien vermittelt werden (Dijkstra 2004), die an die kognitiven Voraussetzungen der Informationsverarbeitung angepasst sind (Mayer 2014).

3 Forschungsfragen

Auf der Grundlage dieses skizzierten Rahmenmodells zielt die berichtete Studie darauf, einen Test zur Erfassung des theoretisch spezifizierten interdependenten Unterrichtsplanungsentscheidens zu entwickeln. Um dabei Hinweise auf Konstruktvalidität zu finden, soll die angenommene innere Struktur des Konstrukts geprüft werden (Jenßen et al. 2015). Zusätzlich sollen erste Aspekte der inhaltlichen und diskriminanten Validität betrachtet werden. So sollen auch Expert*innenurteile die Operationalisierung tragen und die Testergebnisse z.B. nicht mit den kognitiven Grundfähigkeiten der Teilnehmer*innen zusammenhängen, weil die professionelle Unterrichtsplanung als spezifisches Ergebnis eines berufsbezogenen Lernprozesses erachtet wird (König et al. 2015). Die leitenden Forschungsfragen lauten:

1. Kann das angenommene zweidimensionale (formale und inhaltliche) Rahmenmodell interdependenten Entscheidens faktoriell valide (als Teilaspekt von Konstruktvalidität) in einem standardisierten, raschhomogenen Test abgebildet werden?
2. Lassen sich während der Entwicklung und Pilotierung des Tests Hinweise auf dessen inhaltliche und die diskriminante Validität finden?

4 Methode

4.1 Testkonstruktion

Unter Rückgriff auf die Stärken von vignettengestützten Testverfahren (Friesen und Feige 2020) und bei Beachtung planungstestkritischer Aspekte (Rothland 2021) wurde entschieden, das Rahmenmodell interdependenten Planungsentscheidens mittels eines Vignettentests zu operationalisieren. Obwohl offene Vignetten (s. zur Übersicht Weingarten 2019) für die Unterrichtsplanung die Chance bieten, die Komplexität dieses konstruktiven Problemlöse- und Entscheidungsprozesses samt aktiver Identifikation von Entscheidungsoptionen umfassender zu erfassen, wurde ein geschlossenes Testverfahren (Rutsch et al. 2018) gewählt: Der Test sollte ökonomisch einsetzbar sein und dabei analytisch scharf nur auf das interdependente Entscheiden als Teilprozess der Unterrichtsplanung zielen. Zur Stimulation dieses Teilprozesses sollten bereits identifizierte und auszuwählende Planungsoptionen unter gedanklicher Vorwegnahme ihrer möglichen Folgen auf ihre Stimmigkeit (Kohärenz) hin geprüft werden. Durch die alleinige Abbildung dieses Teilprozesses sollten außerdem die weitergehenden kognitiven Herausforderungen und zusätzlichen Planungsprobleme insbesondere von Noviz*innen bei der komplexen Planungsaufgabe möglichst geringgehalten werden (Gassmann 2013).

Für die Entwicklung dieses Tests wurde einer rationalen Testkonstruktionsstrategie gefolgt (Moosbrugger und Brandt 2020). Hierbei wurden zunächst die Anforderungen an die Struktur doppeltgeladener Items (formale und inhaltliche Dimension) theoretisch aus dem zweidimensionalen Rahmenmodell abgeleitet. Anschließend wurden entsprechende Items in zwei Schritten entwickelt.

Tab. 1 Zuteilung der Vignetten zu den angenommenen Fähigkeitsdimensionen und Einführung der Vignettenkürzel

		Formale Fähigkeitsdimensionen		
		1.1 – konditional interdependentes Entscheiden	1.2 – begrenzt interdependentes Entscheiden	1.3 – (potenziell) unbegrenzt interdependentes Entscheiden
Inhaltliche Fähigkeitsdimensionen	2.1 – intentionales Entscheiden	Ernährung (Sachunterricht) ER 1	Verkehrserziehung (Sachunterricht) VE 2	Ernährung (Sachunterricht) ER 3
	2.2 – thematisches Entscheiden	Wortschatz (Englisch) WO 1	Darstellung statistischer Daten (Mathematik) DA 2	Darstellung statistischer Daten (Mathematik) DA 3
	2.3 – methodisches Entscheiden	Verkehrserziehung (Sachunterricht) VE 1	Ernährung (Sachunterricht) ER 2	Verkehrserziehung (Sachunterricht) VE 3
	2.4 – mediales Entscheiden	Darstellung statistischer Daten (Mathematik) DA 1	Wortschatz (Englisch) WO 2	Wortschatz (Englisch) WO 3

4.2 Schritt 1 – Itemformulierung

Gemeinsam mit einer erfahrenen Grundschullehrkraft (30 Jahre Berufspraxis) wurden realitätsnahe Unterrichtsplanungssituationen und Planungsoptionen als Vignetten formuliert. Diese Vignetten wurden jeweils einer der drei formalen und einer der vier inhaltlichen Fähigkeitsdimensionen zugeordnet (s. Tab. 1). Zur Vermeidung fachlicher Einseitigkeiten sollten die Unterrichtsplanungssituationen ein (fremd-)sprachliches, naturwissenschaftliches, mathematisches und sachunterrichtliches Planungsspektrum abbilden.

4.3 Schritt 2 – Leitfadengestütztes Expert*innenrating der Items zur Prüfung der Inhaltsvalidität und Itemüberarbeitung

Um Hinweise auf die Inhaltsvalidität der entwickelten Vignetten zu erhalten, wurden die Vignetten von drei Fachdidaktiker*innen für Englisch, Mathematik und Sachunterricht als Expert*innen beurteilt. Definiert und gewählt wurden diese Expert*innen (Bauder 2013) anhand von drei Kriterien: Für die inhaltliche und methodische Einschätzung des theoretischen Hintergrunds der Studie und der Operationalisierung durch die Vignetten sollten die Expert*innen erstens über eine wissenschaftliche und in Publikationen ausgewiesene fachdidaktische Expertise verfügen. Zur Einschätzung der Unterrichtsplanung als Aspekt der professionellen Entwicklung von (angehenden) Lehrkräften sollten die Expert*innen zweitens seit mehreren Jahren in der Lehramtsausbildung tätig sein. Um bei dieser Einschätzung außerdem die Besonderheiten (der Planung) des alltäglichen Unterrichts berücksichtigen zu können, die sich nicht aus der Literatur erschließen lassen und die das Entscheidungsverhalten im Test beeinflussen könnten, sollten die Expert*innen drittens mehrere Jahre lang Erfahrungen mit der Praxis der Unterrichtsplanung als Lehrkräfte in einer Schule gesammelt haben.

Die ausgewählten Expert*innen wurden gebeten, die ihnen fachlich zugeordneten Planungssituationen und Vignetten entlang einer Systematik von fünf Leitfragen schriftlich zu beurteilen. Auf eine Standardisierung dieser Fragen wurde zugunsten einer differenzierten inhaltlich-argumentativen Auseinandersetzung mit den Situationen und Vignetten verzichtet (Döring und Bortz 2016). Die Leitfragen bezogen sich auf 1. die Nähe der Vignetten zum realen Fachunterricht, 2. die fachliche und fachdidaktische Korrektheit der Vignetten, 3. die fachdidaktische Stimmigkeit der Begründungen für die im Vergleich zu den Distraktoren kohärenteste Planungsentscheidung und Unabhängigkeit von eventuell leitenden schulpraktischen oder fachdidaktischen Planungserwägungen, 4. die Disjunktheit der Vignetten und 5. die Angemessenheit der Vignetten als Indikatoren für das Rahmenmodell interdependenten Entscheidens.

Die verschriftlichten Beurteilungen wurden mit den Expert*innen diskutiert, bevor die Änderungsvorschläge in die Vignetten eingearbeitet wurden (z.B. Verzicht auf das Wort „Mathematikbuch“ als möglichem Hinweisreiz auf eine „sozial unerwünschte Antwort“, weil eine alleinig lehrbuchgestützte Unterrichtplanung unter Mathematiklehrkräften als problematisch gelten könnte). Die Überarbeitung wurde den Expert*innen schließlich erneut zur Prüfung vorgelegt.

4.4 Instrument

Insgesamt wurden zu vier Themenbereichen jeweils drei, zusammengenommen also zwölf disjunkte Vignetten entwickelt, in denen interdependente Planungsentschei-

- a) Sie planen eine Unterrichtseinheit zur Verkehrserziehung in einer zweiten Klasse mit 23 Schüler*innen, die aus der Hamburger Innenstadt¹ kommen. Laut Lehrplan sollen Schüler*innen im zweiten Schuljahr im Aufgabengebiet Verkehrserziehung Kompetenzen erwerben, die es ihnen ermöglichen, sich im Straßenverkehr zunehmend selbstständig, sicher und mitverantwortlich zu verhalten, ihren Bewegungsraum zunehmend eigenverantwortlich zu erweitern und sich altersangemessen mit den Bedingungen und Auswirkungen des Verkehrs auseinanderzusetzen. Ihre Einheit bildet gleichzeitig die Grundlage für eine Einheit zum Fahrradfahren im dritten Schuljahr, in der die Schüler*innen ihren „Fahrradführerschein“ machen, um fortan mit dem Fahrrad zur Schule fahren zu dürfen. Die eine Hälfte der Schüler*innen der Klasse geht allein oder in Begleitung zur Schule² und passiert dabei insgesamt bis zu drei Ampeln¹. Die andere Hälfte wird mit dem Auto zur Schule gebracht². Direkt vor Ihrer Schule befindet sich eine Ampel an einer wenig befahrenen Straße¹. Die Klasse ist insgesamt lebhaft und neugierig², insbesondere bei jeglicher Art von Aktivität außerhalb des Klassenraums, wobei sie meist in der Klasse aufgestellte Regeln befolgt².

- b) Gesetzt den Fall, dass diese Planungssituation den Rahmen für Ihre Planungsüberlegungen absteckt: Bitte vollziehen Sie die folgenden Planungsentscheidungen nach und wählen Sie diejenige aus, die die Planungssituation am stimmigsten berücksichtigt.

Vignette VE 1 (FD 1.1, 2.3)	Vignette VE 2 (FD 1.2, 2.1)	Vignette VE 3 (FD 1.3, 2.3)
<p>c) A. Während des Unterrichtseinstiegs erkennen die Schüler*innen mögliche Gefahren im Straßenverkehr, indem sie im Sitzkreis von ihrem heutigen Weg in die Schule erzählen und Gemeinsamkeiten der Situationen sammeln, die ihnen auf diesem Weg gefährlich erscheinen.</p> <p>B. Während des Unterrichtseinstiegs erkennen die Schüler*innen Gefahrenpunkte im Straßenverkehr, indem sie zunächst ein Bild einer gefährlichen Situation im Straßenverkehr malen und daraus anschließend Regeln für das Verhalten im Straßenverkehr ableiten.</p>	<p>A. Am Ende der ersten Stunde der Einheit sollen Ihre Schüler*innen die wichtigsten Verkehrsregeln aus Sicht von Fußgänger*innen (z.B. Verpflichtung, den Gehweg zu nutzen) und für das Mitfahren im Auto (z.B. Anschnallen) benennen können.</p> <p>B. Am Ende der ersten Stunde der Einheit sollen Ihre Schüler*innen die wichtigsten Verkehrsregeln aus Sicht von Fußgänger*innen (z.B. Verpflichtung, den Gehweg zu nutzen) und von Autofahrer*innen (z.B. von links überholen müssen) benennen können.</p>	<p>A. In einer späteren Stunde der Unterrichtseinheit sollen die Schüler*innen optische und akustische Signale des Straßenverkehrs wahrnehmen und angemessen reagieren können. Gemeinsam mit dem Schulsozialarbeiter laufen Sie mit ihrer Klasse zur nächstgelegenen Ampel und üben die Ampelüberquerung zunächst gemeinsam und anschließend in begleiteten Kleingruppen. Dabei bieten Sie bei Bedarf Unterstützung durch Bildtafeln, auf denen die zu beachtenden Schritte abgebildet sind, und greifen ansonsten nur im Notfall ein.</p> <p>B. In einer späteren Stunde der Unterrichtseinheit sollen sich die Schüler*innen im Straßenverkehr sicherheitsbewusst und eigenständig bewegen. Gemeinsam mit dem Schulsozialarbeiter laufen Sie mit Ihrer Klasse zur nächstgelegenen Ampel und lassen sich von den Kindern zunächst die Funktionsweise und Symbolik einer Ampel an einem kleinen Modell, das Sie mitgebracht haben, erklären, bevor Sie die Straße einige Male gemeinsam überqueren.</p>

Anmerkungen: ¹ = situative und ² = individuelle Planungsvoraussetzungen; A = kohärenteste Planungsentscheidung; B = Beispieldistraktor

Abb. 1 Beispelvignetten zur Verkehrserziehung im Sachunterricht der Jahrgangsstufe 2 an einer Grundschule jeweils mit der richtigen Antwort (A) und einem von drei Distraktoren (B) als Beispiel

dungen auf eine Unterrichtssituation zu beziehen sind (s. Tab. 1; Abb. 1). Nach vier Testdurchläufen mit Studierenden wurden 30 min als wahrscheinliche Testbearbeitungsdauer angegeben.

Jede der zwölf Vignetten weist die gleiche Struktur auf (vgl. Abb. 1):

- a) Zunächst wird eine gemeinsame Unterrichtssituation für jeweils drei Vignetten anhand situativer Unterrichtsbedingungen und individueller Lernvoraussetzungen der Schüler*innen beschrieben, für die interdependente Planungsentscheidungen zu treffen sind.
- b) Anschließend folgt die Instruktion, in jeder Vignette die Planungsentscheidung auszuwählen, „die die Planungssituation am stimmigsten berücksichtigt“. Dabei wird angenommen, dass diese Entscheidung nur als Ergebnis vorausgehender eigener interdependenter Planungsüberlegungen zu den vorgegebenen Planungsoptionen auf Planungsstufe 2 gewählt werden kann.
- c) Zur Operationalisierung der drei formalen Fähigkeitsdimensionen werden für jede der vier Planungssituationen drei disjunkte Vignetten à vier Planungsentscheidungen als Antwortmöglichkeiten präsentiert: VE 1 – zwei verknüpfte Planungsoptionen und deren *Folge* in Varianten von Konditionalaussagen als Indikator für konditionales Entscheiden (FD 1.1); VE 2 – Planungsentscheidungen mit *zwei* verknüpften Planungsoptionen als Indikator für begrenzt interdependentes Entscheiden (FD 1.2); VE 3 – Entscheidungen mit *vier* verknüpften Planungsoptionen als Indikator für (potenziell) unbegrenzt interdependentes Entscheiden (FD 1.3). In jeder Vignette ist genau eine – die kohärenteste – Entscheidung zu wählen, bei der die Planungsoptionen im Unterschied zu den Distraktoren stimmig untereinander und mit den situativen Unterrichts- und individuellen Lernbedingungen verknüpft sind.

Die inhaltliche Fähigkeitsdimension wird dadurch operationalisiert, dass jede Vignette schwerpunktmäßig für intentionales (FD 2.1), thematisches (FD 2.2), methodisches (FD 2.3) oder mediales (FD 2.4) Entscheiden steht. Zur Vermeidung von Überschneidungen beziehen sich die jeweiligen Planungsoptionen in den Entscheidungen auf unterschiedliche situative und individuelle Bedingungen der Planungssituation. Um den Einfluss weiterführenden fachlichen und fachdidaktischen Wissens oder eventueller Stärken bzw. Schwächen in einer der inhaltlichen Entscheidungsdimensionen auf das interdependente Planungsentscheiden auszuschließen, sind jeweils nur zwei der drei Vignetten zu einer Situation derselben inhaltlichen Dimension zugeordnet.

4.5 Design der weiteren Validitätsprüfung

Zur Feststellung faktorieller Validität als Teilaspekt der Konstruktvalidität (Jenßen et al. 2015) wurde die innere Struktur des Konstrukts geprüft, das über das vorgestellte Instrument operationalisiert wurde. Dazu wurden die dimensionale Struktur und die Raschhomogenität der Vignetten untersucht. Die notwendigen Daten wurden in einer querschnittlichen Onlineuntersuchung gewonnen, bei der die Teilnehmer*innen alle zwölf Vignetten zu den vier Planungssituationen bearbeiteten.

Die Reihenfolge der vier Planungssituationen, der jeweiligen drei Vignetten pro Planungssituation und der jeweils vier Antwortmöglichkeiten pro Vignette wurde randomisiert. Um außerdem erste Hinweise auf die diskriminante Validität zu erhalten, wurde der Zusammenhang der erhobenen Abiturnote der Teilnehmer*innen als Schätzer der kognitiven Grundfähigkeiten (in einem ähnlichen Zusammenhang, s. z. B. König et al. 2015) mit dem Testscore geprüft.

4.6 Pilotierungsstichprobe

Die Stichprobe sollte mindestens je 100 praktizierende Lehrkräfte (Rothland 2021) und Lehramtsstudierende umfassen, um eine Raschmodellierung mit höherer Schätzgenauigkeit und Modellgüte zu ermöglichen (Rost 2004) und ein breites Spektrum der Itemschwierigkeiten und Personenfähigkeitsparameter abzubilden. Die praktizierenden Lehrkräfte wurden über soziale Medien rekrutiert, die Studierenden über Seminare an verschiedenen Universitätsstandorten und über (digitale) Aushänge an der Universität Vechta. Als Incentive wurden 20 Gutscheine verlost.

Insgesamt haben 941 Personen mindestens eine Testvignette bearbeitet, von denen 369 Personen (39,2 %) den Test komplett abgeschlossen haben (126 Lehramtsstudierende, 172 praktizierende Lehrkräfte, 15 Referendar*innen und 56 Personen mit anderen Tätigkeiten (Nicht-Lehramtsstudierende, Dozierende, Fortbildende)). Bei 178 Personen mit fehlenden Werten, die aber mindestens drei Vignetten vollständig bearbeitet haben, wurden die fehlenden Testwerte bei der Schätzung der Item- und Modellparameter berücksichtigt. Damit bestand die Analysestichprobe aus insgesamt $N=547$ Personen.

Die 369 teilnehmenden Personen ohne fehlende Werte im Test sind durchschnittlich 35 Jahre alt ($M=35,06$, $SD=12,81$; Lehramtsstudierende $M=23,48$, $SD=3,58$; Lehrkräfte $M=42,37$, $SD=10,58$) und kommen aus Niedersachsen (52,6 %), Nordrhein-Westfalen (19,0 %) sowie den anderen Bundesländern mit Ausnahme des Saarlands. 290 (78,6 %) der Teilnehmenden sind weiblich, 71 (19,2 %) männlich und drei (0,8 %) divers. Fünf (1,4 %) haben keine Angabe gemacht. Die 313 Lehramtsstudierenden, Lehrkräfte und Referendar*innen studieren bzw. unterrichten im Durchschnitt zwei Fächer ($M=2,25$, $SD=0,82$): 207 mindestens eine Sprache, 149 Mathematik/Naturwissenschaften, 71 Sachunterricht, 152 Gesellschaftswissenschaften, 35 Musik/Kunst und 26 Sport. Mit 127 ordnen sich die meisten der 313 Lehrkräfte, Lehramtsstudierenden und Referendar*innen der Grundschule zu, 82 dem Gymnasium, 32 der Realschule, 20 der Oberschule, je 16 der Gesamt- und der Berufsschule und 20 anderen Schulen.

4.7 Statistische Auswertung

Für die Untersuchung der dimensional Struktur des Tests (Konstruktvalidität) wurde zunächst ein eindimensionales Raschmodell in *Acer ConQuest 4.0* (Adams et al. 2015) spezifiziert (Modell 1). Anschließend wurden diesem Modell zwei multidimensionale between-item Raschmodelle (Adams et al. 1997) mit den drei formalen (Modell 2) und den vier inhaltlichen (Modell 3) Dimensionen der Planungsfähigkeit gegenübergestellt. Zum Schluss wurden diese beiden Modelle in ein multidimensio-

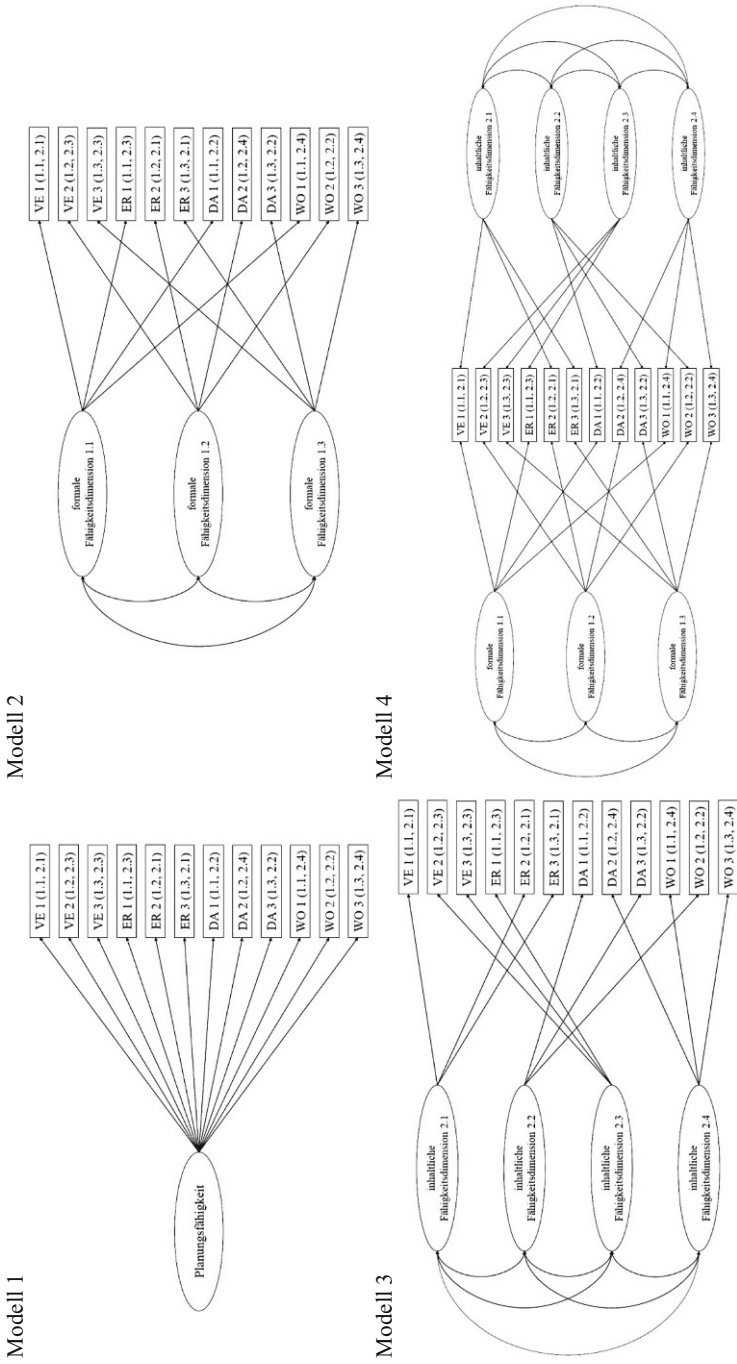


Abb. 2 Schematische Darstellung der vier Vergleichsmodelle

nales within-item Raschmodell (Modell 4; ebd.) überführt (s. Abb. 2 für alle vier Modelle), dessen Güte mittels *Deviance-Test* (Rost 2004) mit den vorherigen drei Modellspezifikationen verglichen wurde. Außerdem wurden die Zusammenhänge zwischen der Abiturnote als Schätzer der kognitiven Grundfähigkeiten und dem Testscore betrachtet (diskriminante Validität).

Zur Prüfung der Raschhomogenität der einzelnen Items wurde der Infit der Items auf Basis der weighted *MNSQ*-Werte betrachtet. Mithilfe der von ConQuest ausgegebenen *T*-Test-Statistik und der Daumenregel von Adams und Khoo (1996) wurde festgestellt, inwieweit die jeweiligen Infit-Werte vom erwarteten Wert 1 abweichen. Im Sinne dieser Daumenregel weichen Items im Intervall $0,75 \leq wMNSQ \leq 1,3$ nicht bedeutsam von dem *wMNSQ* von 1 ab, der unter den Bedingungen des Raschmodells zu erwarten ist. Neben diesen probabilistischen Testkennwerten wurde die Güte des Tests auch anhand von Maßen der klassischen Testtheorie beurteilt. Auf Itemebene wurden die in ConQuest ausgegebene Schwierigkeit und Trennschärfe in Form der Item-Discrimination-Parameter sowie die für die einzelnen Dimensionen in ConQuest ermittelten *EAP*-Reliabilitäten in den Blick genommen.

5 Ergebnisse

5.1 Prüfung der dimensionalen Struktur

Die Ergebnisse der *Deviance-Tests* für die Gegenüberstellung der Modelle sind Tab. 2 zu entnehmen. Das eindimensionale Modell (Modell 1) bildet die Fähigkeitsstruktur signifikant besser ab als das dreidimensionale (Modell 2) und das vierdimensionale (Modell 3) between-item Modell. Diese drei Modelle unterscheiden sich allerdings signifikant von dem multidimensionalen within-item-Modell (Modell 4), das den geringsten Abweichungskoeffizienten aufzeigt und damit die Fähigkeitsstruktur am besten abbildet. Die postulierten formalen und inhaltlichen Fähigkeitsdimensionen können also empirisch abgebildet werden.

5.2 Itemanalyse

Tab. 3 zeigt die Ergebnisse der probabilistischen und klassischen Itemanalyse für Modell 4 im Überblick. Die Infit-bezogenen *wMNSQ* variieren im Bereich von $0,91 \leq wMNSQ \leq 1,1$ und liegen trotz der im *T*-Wert ersichtlichen Abweichung von Item VE 2 (1.2, 2.3) vom raschbezogenen Erwartungswert des Infits von 1 im Daumenbereich von Adams und Khoo (1996). Insofern ist von einer hinreichenden Homogenität der Items gemäß den Bedingungen des Raschmodells auszugehen. Die Itemschwierigkeiten liegen dabei im Intervall $-0,04 \leq p_i \leq 1,23$ und die einzelnen Schwierigkeitsparameter der Items überwiegend im mittleren Bereich. Auf der Ebene individueller Items zeigt sich deskriptiv keine eindeutige Abstufung in Abhängigkeit von der formalen Dimension. Zwar ist beispielsweise das (mit Abstand) schwerste Item (ER 3) der formalen Fähigkeitsdimension 1.3 zuzuordnen, das zweit-schwerste (ER 1) aber der Dimension 1.1 und das leichteste (VE 2) der Dimensi-

Tab. 2 Ergebnisse der Deviance-Tests

Modell	Abweichung	Anzahl geschätzter Parameter	Differenz Modellvergleich	Abweichung	Parameterdifferenz	<i>p</i>
Modell 1: eindimensional	7791,29	13	–	–	–	–
Modell 2: dreidimensional (formal: 1, 1–1,3)	7822,46	18	Modell 1 vs. Modell 2	31,17	5	<0,01
Modell 3: vierdimensional (inhaltlich: 2, 1–2,4)	7813,72	22	Modell 1 vs. Modell 3	22,43	9	<0,01
Modell 4: 4 × 3-within-item-dimensionales Modell	7747,42	40	Modell 1 vs. Modell 4	43,87	27	<0,05
			Modell 2 vs. Modell 4	75,04	22	<0,01
			Modell 3 vs. Modell 4	66,30	18	<0,01

Tab. 3 Itemschwierigkeitsparameter des within-Item-Modells (Modell 4) sowie Angaben zum Weighted Item Fit (Mean Square, Konfidenzintervall, T-Wert) und zur Item-discrimination

Item/Vignette	Lösungswahrscheinlichkeit	Itemschwierigkeitsparameter	SE	wMNSQ	Konfidenzintervall	T	Discrimination
VE 1	0,40	0,28	0,05	0,96	0,92, 1,08	-1,1	0,57
VE 2	0,52	-0,04	0,05	0,91	0,92, 1,08	-2,2	0,60
VE 3	0,38	0,36	0,05	1,03	0,91, 1,09	0,8	0,50
ER 1	0,22	0,84	0,06	1,09	0,88, 1,12	1,4	0,34
ER 2	0,42	0,24	0,05	1,05	0,92, 1,08	1,1	0,45
ER 3	0,15	1,23	0,07	1,10	0,84, 1,16	1,3	0,29
DA 1	0,41	0,32	0,05	0,99	0,91, 1,09	-0,2	0,56
DA 2	0,33	0,49	0,05	1,02	0,91, 1,09	0,5	0,43
DA 3	0,41	0,33	0,05	1,01	0,90, 1,10	0,1	0,57
WO 1	0,33	0,46	0,05	1,04	0,92, 1,08	1,0	0,36
WO 2	0,37	0,43	0,05	1,07	0,91, 1,09	1,5	0,50
WO 3	0,39	0,33	0,05	0,97	0,92, 1,08	-0,8	0,53

Kürzel vgl. Tab. 1

VE Verkehrserziehung (Sachunterricht), ER Ernährung (Sachunterricht), WO Wortschatz (Englisch), DA Darstellung statistischer Daten (Mathematik)

Tab. 4 Reliabilität und Varianzen der Fähigkeitsdimensionen und messfehlerbereinigte Zusammenhänge

Fähigkeitsdimension	1.1	1.2	1.3	2.1	2.2	2.3	2.4
1.1	–	–	–	–	–	–	–
1.2	0,59	–	–	–	–	–	–
1.3	0,81	0,56	–	–	–	–	–
2.1	0,48	0,56	0,54	–	–	–	–
2.2	0,63	0,58	0,57	0,64	–	–	–
2.3	0,31	0,27	0,16	0,57	0,70	–	–
2.4	–0,01	0,10	–0,17	0,28	0,45	0,57	–
<i>EAP-Reliabilität</i>	0,62	0,62	0,61	0,60	0,67	0,65	0,57
<i>Varianz</i>	0,42	0,60	0,90	0,71	1,11	0,73	0,90

on 1.2. Die Discriminationparameter im Bereich von $0,29 \leq r_{it} \leq 0,60$ zeigen eine akzeptable bis hinreichend gute Trennschärfe der einzelnen Items.

Die Reliabilitäten der sieben Fähigkeitsdimensionen einschließlich ihrer Varianz und interkorrelativen Zusammenhänge sind in Tab. 4 abgetragen. Die *EAP*-Reliabilitäten liegen im Bereich von $0,57 \leq EAP \leq 0,67$ und sind mit Ausnahme der Fähigkeitsdimension 2.4 als hinreichend akzeptabel zu bewerten. Die jeweiligen Interkorrelationen zwischen den Dimensionen variieren ferner im Intervall $-0,17 \leq r \leq 0,81$. Insbesondere die Fähigkeitsdimension 2.4 zeigt sehr kleine und zum Teil negative Korrelationen mit den formalen Fähigkeitsdimensionen, ebenso Fähigkeitsdimension 2.3 mit Fähigkeitsdimension 1.3.

5.3 Zusammenhänge zwischen dem Testscore für die drei formalen Fähigkeitsdimensionen und den kognitiven Grundfähigkeiten

Als Schätzer für die kognitiven Grundfähigkeiten wurden die Abiturnoten der Teilnehmer*innen herangezogen. Für die Gesamtstichprobe zeigt sich ein zwar statistisch signifikanter, aber nicht inhaltlich bedeutsamer Zusammenhang zwischen diesen Noten und der Fähigkeitsdimension 1.1 (konditionales Entscheiden), $r(360) = -0,12$, $p = 0,018$: Eine bessere (niedrigere) Abiturnote geht also mit einem höheren Score (und damit mehr gelösten Items) einher.

Um etwaige Effekte des Lernstandes unterrichtsplanerischer Fähigkeiten genauer zu prüfen, wurden die zwei Gruppen der Lehramtsstudierenden und der Lehrkräfte getrennt voneinander untersucht. Für die Lehramtsstudierenden findet sich dieselbe statistisch signifikante, jedoch praktisch nicht bedeutsame Korrelation zwischen der Abiturnote und Fähigkeitsdimension 1.1, $r(126) = -0,18$, $p = 0,045$. Für die Lehrkräfte können dagegen keine signifikanten Korrelationen zwischen deren Abiturnoten und den drei formalen Fähigkeitsdimensionen festgestellt werden.

6 Diskussion

In einer Linie mit didaktischen Ansätzen und Studien des Decision-Making-Paradigmas wird in aktuellen Unterrichtsplanungskompetenzmodellen zwar die Bedeutung interdependenten Planungsentscheidens betont, allerdings noch nicht in den empiri-

schen Forschungsfokus gerückt. Deshalb wurde in diesem Beitrag ein zweidimensionales Rahmenmodell mit einer formalen (konditionales, begrenzt und potenziell unbegrenzt interdependentes Entscheiden) und einer inhaltlichen Fähigkeitsdimension (intentionales, thematisches, methodisches und mediales Entscheiden) zusammen mit einem vignettenbasierten Test vorgestellt.

Hinsichtlich der ersten Forschungsfrage nach der Konstruktvalidität ist positiv an der Pilotierung des Tests zu werten, dass die postulierten Fähigkeitsdimensionen empirisch in einem mehrdimensionalen Raschmodell abbildbar sind. Die erfolgreiche Prüfung der faktoriellen Validität des Tests liefert also einen Hinweis auf die Konstruktvalidität. Damit liegt erstmals ein komplexer, aber ökonomisch einsetzbarer Fähigkeitstest zur Erweiterung vorhandener Instrumente vor, mit dem unterschiedliche Fähigkeitsdimensionen interdependenten Unterscheidens untersuchbar werden. Dabei kann die empiriegestützte Spezifikation des Planungsentscheidens als Chance für die Verknüpfung der zum Teil sehr verschiedenen Ansätze verstanden werden (Rothland 2021), die die Interdependenzannahme teilen.

Auch eine zurückhaltend bejahende Antwort der zweiten Forschungsfrage nach der inhaltlichen und diskriminanten Validität wird von den Ergebnissen getragen. So wurden mithilfe von Expert*innenurteilen Anzeichen für die Inhaltsvalidität des Tests gefunden; und die Betrachtung des Zusammenhangs zwischen dem interdependenten Planungsentscheiden und der Abiturnote als Schätzer für kognitive Grundfähigkeiten in den beiden Teilstichproben der Lehramtsstudierenden und der Lehrkräfte zeigt, dass (angehende) Lehrkräfte kaum von ausgeprägteren Grundfähigkeiten zu profitieren scheinen. Dieser Befund spricht für eine gewisse Unabhängigkeit der Testbearbeitung von diesen Fähigkeiten und damit für die diskriminante Validität als weiterem stützenden Aspekt der Konstruktvalidität. Möglicherweise wird mit dem Test also ein Teil einer professionellen Kompetenz erfasst, die es in einem spezifischen berufsbiografischen Entwicklungsprozess zu lernen gilt. Ein Nebenertrag der Einbeziehung von Lehrkräften in diese Validitätsprüfung ist außerdem, dass die Befunde auf die Gültigkeit von Annahmen zur Unterrichtsplanungskompetenz auch in der beruflichen Praxis hindeuten. Eventuell sind diese Annahmen deshalb kein alleiniges „Phänomen der Lehrer*innenbildung“ (o. S.), wie es Rothland (2021) bezogen auf bisherige Testentwicklungen nur mit Studierenden fragend anmerkt.

Diese Ergebnisse, die auf einer nicht repräsentativen Stichprobe basieren, unterliegen allerdings einigen Limitationen. So erhärten zwar die Hinweise auf faktorielle und diskriminante Validität die Annahme von Konstruktvalidität. Dennoch sollten künftig zusätzlich die konvergente Validität z.B. im Abgleich mit Ergebnissen zu offenen und domänenspezifischen Testvignetten (z.B. Kirsch 2020) oder die prognostische Validität z.B. im Vergleich mit späteren Unterrichtsentwürfen (König et al. 2015) überprüft werden. Zudem wurden die Hinweise auf Inhaltsvalidität zwar mithilfe kriteriengeleitet gewählter Expert*innen systematisch gewonnen. Gemäß jüngeren Empfehlungen sollten diese Hinweise aber mit einer standardisierten Befragung einer größeren Expert*innengruppe abgesichert werden (Jenßen et al. 2015).

Auch weitere messtheoretische und inhaltliche Limitationen schränken die Aussagekraft der Ergebnisse ein. Die Lösungswahrscheinlichkeiten und Itemschwierigkeiten zeigen deskriptiv keine Tendenz zu einer Abstufung der formalen Fähigkeitsdimensionen, die jedoch bei der zunehmenden Komplexität der Dimensionen

(zumindest hinsichtlich FD 1.2 und 1.3) angenommen werden könnte. Möglicherweise stehen das Rahmenmodell und der Test deshalb lediglich für die dimensionale Struktur interdependenten Entscheidens, nicht aber für eine potenzielle Stufung. Die Reliabilitäten sind nur akzeptabel, was der Vorgabe von Planungsentscheidungen für einen eigentlich offenen und kreativen Problemlöseprozess (Vogelsang und Riese 2017) und der Tatsache geschuldet sein könnte, dass der Test nicht an der initialen Planungsidee ansetzt. Diese Planungsidee samt der aktiven Identifikation und Elaboration von Entscheidungsoptionen könnte aber für die wechselseitige Verknüpfung der Planungsentscheidungen im weiteren Planungsverlauf wichtig sein. Selbst bei der Fokussierung des Teilprozesses interdependenten Entscheidens, dessen vignettenbasierte Abbildung die analytische Stärke des Tests ausmachen soll, wird die Komplexität der wahrzunehmenden und zu interpretierenden Planungsoptionen der Unterrichtssituation nur reduziert und für ausgewählte Schulfächer dargestellt. Deshalb sind lediglich Grob-, aber noch keine Detailplanungsentscheidungen möglich, die darüber hinaus unbeeinflusst von schulischen, fachlichen oder fachdidaktischen Faktoren getroffen werden können. Bei derartigen Entscheidungen können die Teilnehmenden allerdings kein differenzierteres fachliches oder fachdidaktisches Planungswissen bzw. keine elaborierten fachbezogenen Planungsroutinen abrufen. In der Mathematikvignette zeigt sich diese Einschränkung z. B. daran, dass keine Lernaufgaben als zentralem Bezugspunkt mathematikdidaktischer Modelle (z. B. Heckmann und Padberg 2014) geplant werden können.

Trotz dieser Limitationen können das Rahmenmodell und der entwickelte Test zur Suche nach Antworten auf eine der empirisch überraschend offenen Grundsatzfrage zur Unterrichtsplanung beitragen. Denn bisher ist ungeklärt, „wie *fachübergreifend* eine ‚gute‘ Unterrichtsplanung“ und „wie eine qualitätsvolle Leistungsperformanz *empirisch* zu identifizieren und bestimmen ist“ (Rothland 2021, o. S.). Die berichteten Erträge machen nun eine empirische Untersuchung der Wirkungen der Unterrichtsplanung (und damit eine Evidenzbasierung von Prinzipien für eine ‚gute‘ Unterrichtsplanung) denkbar: Hängt das (allgemein-)didaktische Wissen über Interdependenz (Schüle et al. 2017) mit der situationsspezifischen Fähigkeit interdependenten Entscheidens (der vorgestellte Test) zusammen und lässt sich darüber vermittelt auch die Kohärenz eines schriftlichen Unterrichtsentwurfs (König et al. 2015) und in der Folge eine veränderte Performanzqualität des Unterrichts erklären?

Förderung D. Scholl, S. Küth und C. Schüle geben an, keine Förderung erhalten zu haben.

Funding Open Access funding enabled and organized by Projekt DEAL.

Interessenkonflikt D. Scholl, S. Küth und C. Schüle geben an, dass kein Interessenkonflikt besteht.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung

nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Adams, R.J., & Khoo, S.-T. (1996). *Quest. The interactive test analysis system*. Camberwell: Australian Council for Educational Research.
- Adams, R.J., Wilson, M., & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1–23. <https://doi.org/10.1177/0146621697211001>.
- Adams, R.J., Wu, M.L., & Wilson, M.R. (2015). *ACER ConQuest: Generalised item response modeling software [Computer software]. Version 4.0*. Camberwell: Australian Council for Educational Research.
- Bauder, T. (2013). Mit Expertenrating zur Einschätzungsnorm: Unterrichtshandeln beobachten und diskutieren. Qualität feststellen und begründen. In F. K. Oser, T. Bauder, P. Salzmann & S. Heinzer (Hrsg.), *Ohne Kompetenz keine Qualität: Entwickeln und Einschätzen von Kompetenzprofilen bei Lehrpersonen und Berufsbildungsverantwortlichen* (S. 191–214). Bad Heilbrunn: Verlag Julius Klinkhardt.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>.
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., & Müller, P. (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrwissens*. Münster: Waxmann.
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond Dichotomies. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>.
- Borko, H., & Livingston, C. (1989). Cognition and improvisation: differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, 26(4), 473–498. <https://doi.org/10.3102/00028312026004473>.
- Bromme, R. (1981). *Das Denken von Lehrern bei der Unterrichtsvorbereitung. Eine Untersuchung von kognitiven Prozessen von Mathematiklehrern*. Weinheim: Beltz.
- Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (3. Aufl., S. 255–296). New York: Macmillan.
- Conway, P. F. (2001). Anticipatory reflection while learning to teach: from a temporally truncated to a temporally distributed model of reflection in teacher education. *Teaching and Teacher Education*, 17(1), 89–106. [https://doi.org/10.1016/S0742-051X\(00\)00040-8](https://doi.org/10.1016/S0742-051X(00)00040-8).
- De Jong, T., & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105–113. https://doi.org/10.1207/s15326985ep3102_2.
- Dijkstra, S. (2004). The integration of curriculum design, instructional design, and media choice. In N. M. Seel & S. Dijkstra (Hrsg.), *Curriculum, plans, and processes in instructional design* (S. 145–170). Mahwah: Erlbaum.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Berlin: Springer.
- Fox, J., Cooper, R. P., & Glasspool, D. W. (2013). A canonical theory of dynamic decision-making. *Frontiers in Psychology*, 4, 150. <https://doi.org/10.3389/fpsyg.2013.00150>.
- Friesen, N. (2010). Lesson planning. Anglo-American perspectives. *Bildung und Erziehung*, 63(4), 417–430. <https://doi.org/10.7788/bue.2010.63.4.417>.
- Friesen, M. E., & Feige, E. M. (2020). Konstruktion und Einsatz von Vignetten und Concept Cartoons in der Lehrerbildung. In M. E. Friesen, J. Benz, T. Billion-Kramer, C. Heuer, H. Lohse-Bossenz, M. Resch & J. Rutsch (Hrsg.), *Vignettenbasiertes Lernen in der Lehrerbildung. Fachdidaktische und pädagogische Perspektiven* (S. 28–52). Weinheim: Beltz Juventa.
- Funke, J., & Fritz, A. (1995). Über Planen, Problemlösen und Handeln. In J. Funke & A. Fritz (Hrsg.), *Neue Konzepte und Instrumente zur Planungsdiagnostik* (S. 1–45). Bonn: Deutscher Psychologen Verlag.
- Gassmann, C. (2013). *Erlebte Aufgabenschwierigkeit bei der Unterrichtsplanung. Eine qualitativ-inhaltsanalytische Studie zu den Praktikumsphasen der universitären Lehrerbildung*. Wiesbaden: Springer VS.

- Grunwald, A. (2000). *Handeln und Planen*. München: Fink.
- Haas, A. (1998). *Unterrichtsplanung im Alltag. Eine empirische Untersuchung zum Planungshandeln von Hauptschul-, Realschul- und Gymnasiallehrern*. Regensburg: Roderer.
- Heckmann, K., & Padberg, F. (2014). *Unterrichtsentwürfe Mathematik Primarstufe* (Bd. 2). Berlin: Springer.
- Jenßen, L., Dunekacke, S., & Blömeke, S. (2015). Qualitätssicherung in der Kompetenzforschung. Empfehlungen für den Nachweis von Validität in Testentwicklung und Veröffentlichungspraxis. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (Zeitschrift für Pädagogik: Sonderheft 61, S. 11–31). Weinheim: Beltz Juventa.
- John, P.D. (1991). Course, curricular, and classroom influences on the development of student teachers' lesson planning perspectives. *Teaching and Teacher Education*, 7(4), 359–372. [https://doi.org/10.1016/0742-051X\(91\)90005-A](https://doi.org/10.1016/0742-051X(91)90005-A).
- Kirsch, A. (2020). Entwicklung und psychometrische Qualität eines Instrumentes zur Messung von Planungskompetenzen angehender Lehrkräfte im Fach Sachunterricht. *Zeitschrift für Pädagogische Psychologie*. <https://doi.org/10.1024/1010-0652/a000277>.
- Klafki, W. (1996). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik* (5. Aufl.). Weinheim: Beltz.
- Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876–903. <https://doi.org/10.25656/01:4493>.
- Koepfen, K.E. (1998). The experience of a secondary social studies student teacher: seeking security by planning for self. *Teaching and Teacher Education*, 14(4), 401–411. [https://doi.org/10.1016/S0742-051X\(97\)00047-4](https://doi.org/10.1016/S0742-051X(97)00047-4).
- König, J., Buchholtz, C., & Dohmen, D. (2015). Analyse von schriftlichen Unterrichtsplanungen: Empirische Befunde zur didaktischen Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte. *Zeitschrift für Erziehungswissenschaft*, 18(2), 375–404. <https://doi.org/10.1007/s11618-015-0625-7>.
- König, J., Bremerich-Vos, A., Buchholtz, C., Lammerding, S., Strauß, S., & Fladung, I. (2017). Die Bedeutung des Professionswissens von Referendarinnen und Referendaren mit Fach Deutsch für ihre Planungskompetenz (PlanvoLL-D). In S. Wernke & K. Zierer (Hrsg.), *Die Unterrichtsplanung: Ein in Vergessenheit geratener Kompetenzbereich?!* (S. 121–133). Bad Heilbrunn: Verlag Julius Klinkhardt.
- König, J., Bremerich-Vos, A., Buchholtz, C., Fladung, I., & Glutsch, N. (2020). Pre-service teachers' generic and subject-specific lesson-planning skills: On learning adaptive teaching during initial teacher education. *European Journal of Teacher Education*, 43(2), 131–150. <https://doi.org/10.1080/02619768.2019.1679115>.
- König, J., Krepf, M., Bremerich-Vos, A., & Buchholtz, C. (2021). Meeting cognitive demands of lesson planning: Introducing the CODE-PLAN model to describe and analyze teachers' planning competence. *The Teacher Educator*. <https://doi.org/10.1080/08878730.2021.1938324>.
- Küth, S., Scholl, D., Schüle, C., & Rheinländer, K. (2021). Die Optimierung von Medienentscheidungen in der Unterrichtsplanung – eine Utopie? *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 42, 217–236. <https://doi.org/10.21240/mpaed/42/2021.06.20.X>.
- Lui, A. M., & Bonner, S. M. (2016). Preservice and inservice teachers' knowledge, beliefs, and instructional planning in primary school mathematics. *Teaching and Teacher Education*, 56, 1–13. <https://doi.org/10.1016/j.tate.2016.01.015>.
- Mayer, R.E. (Hrsg.). (2014). *The Cambridge handbook of multimedia learning* (2. Aufl.). New York: Cambridge University Press.
- Moosbrugger, H., & Brandt, H. (2020). Itemkonstruktion und Antwortverhalten. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3. Aufl., S. 67–89). Berlin: Springer. https://doi.org/10.1007/978-3-662-61532-4_4.
- Mumford, M.D., Schultz, R.A., & van Doorn, J.R. (2001). Performance in planning: processes, requirements, and errors. *Review of General Psychology*, 5(3), 213–240. <https://doi.org/10.1037/1089-2680.5.3.213>.
- Munthe, E., & Conway, P.F. (2017). Evolution of research on teachers' planning: implications for teacher education. In D.J. Clandinin & J. Husu (Hrsg.), *SAGE handbook of research on teacher education* (S. 836–852). Thousand Oaks: SAGE.
- Mutton, T., Hagger, H., & Burn, K. (2011). Learning to plan, planning to learn: the developing expertise of beginning teachers. *Teachers and Teaching*, 17(4), 399–416. <https://doi.org/10.1080/13540602.2011.580516>.

- Neumann, K. (2020). Die Bedeutung instruktionaler Kohärenz für eine systematische Kompetenzentwicklung. *Unterrichtswissenschaft*, 48(1), 1–10. <https://doi.org/10.1007/s42010-020-00068-6>.
- Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review*, 27(3), 483–504.
- Rey, T., Lohse-Bossenz, H., Wacker, A., & Heyl, V. (2018). Adaptive Planungskompetenz bei angehenden Lehrkräften in der zweiten Phase der Lehrerbildung. *heiEDUCATION*, 1(2), 127–150. <https://doi.org/10.17885/heiup.heied.2018.1-2.23829>.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. Aufl.). Bern: Huber.
- Rothland, M. (2021). Anmerkungen zur Modellierung und Operationalisierung (allgemeindidaktischer) Unterrichtsplanungskompetenz. *Unterrichtswissenschaft*. <https://doi.org/10.1007/s42010-021-00111-0>.
- Rutsch, J., Vogel, M., Rehm, M., & Dörfler, T. (2018). Modellierung der Testletstruktur bei vignetten-basierten Testverfahren mit geschlossenem Antwortformat. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (S. 27–46). Wiesbaden: Springer.
- Sardo-Brown, D. (1996). A longitudinal study of novice secondary teachers' planning: year two. *Teaching and Teacher Education*, 12(5), 519–530. [https://doi.org/10.1016/0742-051X\(95\)00056-P](https://doi.org/10.1016/0742-051X(95)00056-P).
- Scholl, D. (2018). *Metatheorie der Allgemeinen Didaktik. Ein systemtheoretisch begründeter Vorschlag*. Bad Heilbrunn: Klinkhardt.
- Schrader, J., & Schöb, S. (2016). Die Planung von Lehr-Lern-Einheiten mit digitalen Medien: Konzepte und Befunde. *Zeitschrift für Weiterbildungsforschung*, 39(3), 331–347. <https://doi.org/10.1007/s40955-016-0078-5>.
- Schröder, J., Riese, J., Vogelsang, C., Borowski, A., Buschhüter, D., & Enkrott, P. (2020). Die Messung der Fähigkeit zur Unterrichtsplanung im Fach Physik mit Hilfe eines standardisierten Performanztests. *Zeitschrift für Didaktik der Naturwissenschaften*, 26(1), 103–122. <https://doi.org/10.1007/s40573-020-00115-w>.
- Schüle, C., Besa, K.-S., & Arnold, K.-H. (2017). Entwicklung und Validierung eines Testinstruments zur Erfassung allgemeindidaktischer Kompetenz. In S. Wernke & K. Zierer (Hrsg.), *Die Unterrichtsplanung: Ein in Vergessenheit geratener Kompetenzbereich?!* (S. 17–31). Bad Heilbrunn: Klinkhardt.
- Schulz, W. (1972). Unterricht – Analyse und Planung. In P. Heimann, G. Otto & W. Schulz (Hrsg.), *Unterricht. Analyse und Planung* (6. Aufl., S. 13–47). Hannover: Hermann Schroedel.
- Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research*, 51(4), 455–498.
- Shulman, L. (1987). Knowledge and teaching. Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23.
- Stender, A. (2014). *Unterrichtsplanung: Vom Wissen zum Handeln. Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung*. Berlin: Logos.
- Stender, A., Brückmann, M., & Neumann, K. (2017). Transformation of topic-specific professional knowledge into personal pedagogical content knowledge through lesson planning. *International Journal of Science Education*, 39(12), 1690–1714.
- Tang, T. K. W., & Watkins, D. (1994). Towards a taxonomy of teacher thinking. *Educational Research Journal*, 9(1), 38–43.
- Tillema, H. (1984). Categories in teacher planning. In R. Halkes & J. K. Olson (Hrsg.), *Teacher thinking: a new perspective on persisting problems in education* (S. 176–186). Lisse: Swets & Zeitlinger.
- Vogelsang, C., & Riese, J. (2017). Wann ist eine Unterrichtsplanung ‚gut‘? Planungsperformanz in Praxisratgebern zur Unterrichtsplanung. In S. Wernke & K. Zierer (Hrsg.), *Die Unterrichtsplanung: Ein in Vergessenheit geratener Kompetenzbereich?!* (S. 47–61). Bad Heilbrunn: Klinkhardt.
- Weingarten, J. (2019). *Wie planen angehende Lehrkräfte ihren Unterricht? Empirische Analysen zur kompetenzorientierten Gestaltung von Lernangeboten*. Münster: Waxmann.
- Wengert, H. G. (1989). *Untersuchungen zur alltäglichen Unterrichtsplanung von Mathematiklehrern: Eine kognitionspsychologische Studie*. Frankfurt a. M.: Peter Lang.
- Westerman, D. (1991). Expert and novice teacher decision making. *Journal of Teacher Education*, 42(4), 292–305. <https://doi.org/10.1177/002248719104200407>.
- Wing-mui So, W. (1997). A study of teacher cognition in planning elementary science lessons. *Research in Science Education*, 27(1), 71–86. <https://doi.org/10.1007/BF02463033>.
- Wittwer, J., Kratschmayr, L., & Voss, T. (2020). Wie gut erkennen Lehrkräfte typische Fehler in der Formulierung von Lernzielen? *Unterrichtswissenschaft*, 48(1), 113–128. <https://doi.org/10.1007/s42010-019-00056-5>.
- von Wright, G. H. (1974). *Erklären und Verstehen*. Frankfurt a. M.: Athenäum-Fischer.

- Yinger, R. (1980). A study of teacher planning. *The Elementary School Journal*, 80(3), 107–127.
- Yinger, R. J., & Hendricks-Lee, M. S. (1995). Teacher planning. In L. W. Anderson (Hrsg.), *International encyclopedia of teaching and teach education* (2. Aufl., Bd. 2, S. 188–192). Oxford: Pergamon.
- Zahorik, J. A. (1975). Teachers' planning models. *Educational Leadership*, 33(2), 134–139.
- Zierer, K., & Seel, N. M. (2012). General Didactics and Instructional Design: eyes like twins A transatlantic dialogue about similarities and differences, about the past and the future of two sciences of learning and teaching. *SpringerPlus*, 1, 15. <https://doi.org/10.1186/2193-1801-1-15>.