

## Perspektiven von Lehrkräften über Modelle in den Naturwissenschaften

### Eine generalisierende Replikationsstudie

John Rolf Borrmann, Nadine Reinhardt, Moritz Krell & Dirk Krüger

johnrbo@zedat.fu-berlin.de; moritz.krell@fu-berlin.de

Freie Universität Berlin,

Didaktik der Biologie,

Schwendenerstr. 1, 14195 Berlin

---

#### Zusammenfassung

*Modelle sind ein wichtiges Instrument der naturwissenschaftlichen Praxis und sollten deshalb auch im naturwissenschaftlichen Unterricht prominent thematisiert werden. Van Driel und Verloop untersuchten 1999 in einer Studie das Modellverstehen von niederländischen Lehrkräften. Deren Befunde werden breit rezipiert und wiederholt als Beleg für ein wenig ausgeprägtes Modellverstehen von Lehrkräften angeführt. Diese Studie versucht zu überprüfen, inwieweit sich Van Driels und Verloops (1999) Ergebnisse für eine Stichprobe deutscher Lehrkräfte (N = 226) replizieren lassen. Die Daten wurden, wie auch im Original, deskriptiv und mit einer explorativen Faktorenanalyse ausgewertet. Die Ergebnisse der Originalstudie konnten jedoch nicht repliziert werden. Die vier sich ergebenden Faktoren unterscheiden sich deutlich von den ursprünglichen. Inhaltlich lässt sich feststellen, dass die befragten Lehrkräfte Modelle hauptsächlich als Mittel zur Erklärung und Veranschaulichung verstehen. Einflüsse des Schultyps, der Berufserfahrung und der Fächerkombination konnten nicht nachgewiesen werden. Die Ergebnisse verdeutlichen, dass Replikationsstudien von zentraler Bedeutung sind, empirische Ergebnisse auf eine fundierte Basis zu stellen.*

#### Abstract

*Models are an important instrument in the natural sciences and should hence also be used and discussed at school. In 1999 Van Driel and Verloop conducted a study on how Dutch teachers understand models. Their results are still widely cited as evidence for teachers' rather uninformed understanding of models in science. This study aims to investigate whether Van Driel's and Verloop's results can be reconstructed for a sample of German teachers (N = 226). As in the original study, data was analyzed descriptively and using an*

*exploratory factor analysis. However, we are not able to replicate the results of the original study. The resulting four factors are distinctly different from the original ones. Nevertheless we find that the teachers understand models mainly as explanatory and illustrative tools. We do not find any effects of school type, teaching experience or combination of subjects on teachers' understanding of models. The results illustrate that replication studies are essential to base findings on solid research.*

## 1 Einleitung

Modelle sind von großer Bedeutung für die Veranschaulichung und Kommunikation naturwissenschaftlicher Erkenntnisse sowie für die Erkenntnisgewinnung in den Naturwissenschaften. Sie bilden keine „wissenschaftlichen Wahrheiten“ ab, sondern zeigen vielmehr den aktuellen Stand der Forschung, der sich durch das Falsifizieren von Hypothesen fortwährend ändern kann (vgl. POPPER 1959). Anhand von Modellen kann bestehendes Wissen erklärt, können Hypothesen überprüft, neue Erkenntnisse gewonnen und Vorstellungen wissenschaftlicher Ideen kommuniziert werden (KRELL, UPMEIER ZU BELZEN & KRÜGER 2012; OH & OH 2011). Für den naturwissenschaftlichen Unterricht lässt sich daraus ableiten, dass neben der praktischen Arbeit mit Modellen auch explizit über das Modellieren reflektiert werden sollte. Hierbei sollten Aspekte wie der Zweck der Modellbildung oder die Möglichkeiten, Modelle zu testen und zu ändern, zur Sprache kommen (CRAWFORD 2000; GILBERT 2004; GROSSLIGHT, UNGER, JAY & SMITH 1991; SCHWARZ & WHITE, 2005; UPMEIER ZU BELZEN & KRÜGER 2010). Modellen und dem Modellieren wird in aktuellen Standardbeschreibungen für den Biologie- bzw. naturwissenschaftlichen Unterricht große Bedeutung eingeräumt (z. B. KMK 2005; NGSS LEAD STATES 2013). Konsequenterweise wird ein ausgeprägtes Modellverstehen und das Wissen über adäquate Unterrichtsstrategien, dies bei Lernenden zu entwickeln, als Teil des Professionswissen von Lehrkräften naturwissenschaftlicher Fächer angesehen (TEPNER et al. 2012).

Eine grundlegende Studie über das Modellverstehen von Lehrkräften naturwissenschaftlicher Fächer ist die Untersuchung von Van Driel und Verloop (1999). Diese Fragebogenstudie zeigte, dass die befragten niederländischen Lehrkräfte vorwiegend ein Verständnis von Modellen als eine vereinfachte Abbildung der Realität besitzen und ihr Wissen über Modelle „limited and diverse“ ist (VAN DRIEL & VERLOOP 1999, S. 1141). Die Unterrichtsfächer oder die Erfahrung der Lehrkräfte hatten auf ihr Modellverstehen keinen systematischen Einfluss.

Van Driels und Verloops (1999) Untersuchung wurde breit rezipiert und in 204 internationalen Publikationen (*google scholar*, Stand 04/2014) als Beleg

für ein wenig ausgeprägtes Modellverstehen von Lehrkräften weltweit zitiert. Die Ergebnisse der Autoren werden, obwohl sie auf einer relativ kleinen Stichprobe ( $N = 71$ ) ausschließlich niederländischer Lehrkräfte basieren und mittlerweile 15 Jahre zurück liegen, international immer noch als Beleg für das mangelnde Modellverstehen von Lehrkräften und als Grundlage für darauf aufbauende Untersuchungen herangezogen. So führte zum Beispiel Kahn (2011) Analysen zum Modelleinsatz im Unterricht von kanadischen Lehrkräften durch und zog Van Driels und Verloops (1999) Artikel für den Nachweis heran, „that few practicing science teachers systematically engage students in MBT [model based teaching]“ (S. 537). Drei Studien aus den USA zitierten Van Driels und Verloops (1999) Publikation als Beleg dafür, dass Lehrkräfte nur „limited experience and knowledge about scientific modeling or modeling-centered inquiry“ (SCHWARZ 2009, S. 723) besitzen, keine ausgeprägten Fähigkeiten für die Arbeit mit Modellen aufweisen (SCHRAW et al. 2006) oder weder Modelle noch die Modellentwicklung in ausreichendem Maße verstehen (SCHWARZ & GWEKWERERE 2007). Ebenso belegten unter anderem Autoren aus Südamerika (z. B. JUSTI & GILBERT 2003; SILVA 2007) das begrenzte Modellverstehen von Lehrkräften mit den Befunden von Van Driel und Verloop (1999) und entwickelten eigene Untersuchungen auf Grundlage des eingesetzten Fragebogens. Allerdings bleibt trotz der häufigen Verweise die Frage offen, ob die Ergebnisse Van Driels und Verloops (1999) international und über die Zeitspanne hinweg verallgemeinerbar sind. Um die Generalisierbarkeit von empirischen Befunden zu prüfen, sind Replikationsstudien notwendig (AMIR & SHARON 1991; LAMAL 1991; SCHMIDT 2009). Einige Autoren kritisieren, dass aufgrund mangelnder Replikationsstudien viele publizierte Befunde empirischer Forschung *false positives* sind, also nicht als verallgemeinerbar angesehen werden dürfen (z. B. FANELLI 2012; IOANNIDIS 2005; YONG 2012).

Die vorliegende Studie versucht die niederländischen Ergebnisse mit einer Stichprobe deutscher Lehrkräfte zu replizieren. Dazu wurde der Fragebogen von Van Driel und Verloop (1999) unter Zuhilfenahme des niederländischen Originals und der publizierten englischen Version ins Deutsche übersetzt und diese Version dann von deutschen Lehrkräften mit mindestens einem naturwissenschaftlichen Unterrichtsfach beantwortet.

## 2 Theorie

### 2.1 Replikationsstudien

Die Wichtigkeit von Replikationsstudien für wissenschaftliche Erkenntnisse wird an vielen Stellen betont (z. B. AMIR & SHARON 1991; HERSEN & BARLOW 1976; ROSENTHAL 1991). Replikationsstudien gelten als essentiell und als „heart of any science“ (HERSEN & BARLOW 1976, S. 317). Amir und Sharon (1991) warnen vor der falschen Annahme, dass ein einzelnes signifikantes Ergebnis ohne eine Überprüfung als gültig und replizierbar anerkannt wird. Darüber hinaus helfen Replikationsstudien zukünftige explorative Studien und didaktische Konzepte zielführender zu entwickeln. Trotzdem verweist Yong (2012) auf die Problematik, dass zum Beispiel in der Psychologie kaum Replikationsstudien durchgeführt würden und es so zu zahlreichen Betrugsfällen kam, die lange unentdeckt blieben. Schon Hendrick (1991) wies darauf hin, dass der Umgang mit Replikationsstudien sich zwischen den einzelnen Wissenschaftsdisziplinen deutlich unterscheiden würde. Während in der Physik Replikationsstudien als ein „natural and essential part of the business of doing science“ (S. 42) angesehen würden, wäre das Ansehen von diesen in den Sozialwissenschaften und Verhaltenswissenschaften deutlich geringer. Der Grund für die Vernachlässigung von Replikationsstudien wird vor allem in der derzeitigen Veröffentlichungspraxis gesehen (z. B. LAMAL 1991; NEULIEP & CRANDALL 1991, 1993; YONG 2012). So werden hauptsächlich neue, Aufsehen erregende Erkenntnisse veröffentlicht. Eine Forschungsarbeit, die lediglich Ergebnisse repliziert, hat es wegen des fehlenden Neuigkeitswertes schwer, zur Veröffentlichung zu gelangen. Das schreckt viele Forscher/-innen ab, entsprechende Studien überhaupt durchzuführen. Auch kann es passieren, dass die Beweispflicht nach einer nicht erfolgreichen Replikation dem Replizierenden und nicht dem ursprünglichen Autor übertragen wird. Dies führte dazu, dass das Durchführen von Replikationsstudien unattraktiv und in Folge dessen vernachlässigt wird (NEULIEP & CRANDALL 1991, 1993). Daher vermuten einige Autoren, dass allein aufgrund statistischer Logik ein großer Teil der publizierten wissenschaftlichen Befunde *false positives* sind: „It can be proven that most claimed research findings are false“ (IOANNIDIS 2005, S. 696; vgl. FANELLI 2012; YONG 2012).

Einen weiteren Grund für die seltene Durchführung und Publikation von Replikationsstudien sieht Schmidt (2009) darin, dass es kaum Literatur gibt, die einheitliche Anleitungen zur Durchführung von Replikationsstudien bietet. Daher entwirft er ein Konzept zur Replikation, das insbesondere auf der Sys-

tematik von Hendrick (1991) aufbaut. Er unterscheidet zwischen *direkten* (Wiederholung der experimentellen Prozesse der Ausgangsstudie) und *konzeptuellen* (Überprüfung einer Hypothese mit anderen Methoden) Replikationsstudien (Tab. 1). Die vorliegende Studie entspricht einer direkten Replikation (SCHMIDT 2009), es geht also darum, mit den gleichen Instrumenten erneut zu erheben.

**Tabelle 1:** Strategien bei Replikationsstudien (nach Schmidt 2009)

		Direkte Replikationsstudie			Konzeptuelle Replikationsstudie	
Wissenschaftlicher Gewinn:		Ergebnisbestätigung			Erweiterung des Wissens	Verstehen
Funktion:		Überprüfung ...				
Klassen		... vom $\alpha$ -Fehler	... von Störvariablen	... von Betrug	Generalisierung	Hypothesenprüfung
1	Studiendesign: Bedingungen der Durchführung und Manipulation der unabhängigen Variable (Konzeption der Untersuchungssituation)	konst.	konst.	konst.	konst.	konst.
	Studiendesign (Operationalisierung und Konkretisierung in den Instrumenten für die Untersuchung)	konst.	konst.	konst.	konst.	var.
2	Kontext der Studie (Geschlecht, Hintergrundwissen, Bereitschaft der Probanden, Einbettung der Studie, Setting am Befragungsort, Experimentator, Gestaltung des Instruments)	konst.	var.	var.	konst.	var.
3	Probandenpopulation	konst.	konst.	konst.	var.	konst.
4	Konstituierung der abhängigen Variablen	konst.	var.	konst.	konst.	var.

konst.: konstant halten; var.: variieren

Direkte Replikationsstudien haben nach Schmidt (2009) die folgenden vier spezifischen Funktionen, wobei die ersten drei der Bestätigung oder Widerlegung der Befunde dienen:

(1) *to control for sampling error ( $\alpha$ -Fehler)*: Es gilt zu überprüfen, ob der beobachtete Effekt zufällig war (*false positive*);

(2) *to control for artifacts (Störgrößen)*: Es geht darum, mögliche Einflüsse von Störgrößen, z. B. durch die Einbettung der Studie oder die Auswahl der Probanden, ausschließen zu können;

(3) *to control for fraud (Betrug)*: Dies verfolgt den Zweck, grundsätzlich Betrug aufzudecken;

(4) *to generalize results to a larger or to a different population (Generalisierung)*: Es wird versucht, die Ergebnisse durch Einbezug weiterer Gruppen verallgemeinern zu können bzw. ihre Spezifität nachzuweisen. Diese Funktion ist somit der Versuch, die Aussagekraft einer Studie zu erweitern. Dies ist Gegenstand der vorliegenden Replikationsstudie.

Schmidt (2009) beschreibt vier Klassen von Merkmalen einer experimentellen Studie, die je nach Funktion einer Replikationsstudie verändert oder konstant gehalten werden sollen (Tab. 1). Demzufolge ist nach der zugrunde liegenden Funktion dieser Studie (Generalisierung von Befunden) die Probandenpopulation zu variieren, während die anderen Merkmale, also das Studiendesign, der Kontext der Studie und die Konstituierung der abhängigen Variablen, möglichst konstant gehalten werden sollten (HENDRICK 1991; SCHMIDT 2009).

## 2.2 Modelle: Van Driel und Verloop (1999)

Der bewusste und theoriegeleitete Umgang mit Modellen stellt einen wichtigen Bestandteil der naturwissenschaftlichen Grundbildung dar (KRELL 2013; SCHWARZ & WHITE 2005; UPMEIER ZU BELZEN & KRÜGER 2010). Modelle spielen eine Schlüsselrolle im Bereich der naturwissenschaftlichen Forschung und dienen dort nicht nur zur Darstellung und Kommunikation des aktuell postulierten Wissenstandes zwischen Forschungsgruppen, sondern mit ihrer Hilfe können auch neue Hypothesen abgeleitet und im Zuge ihrer Überprüfung dann neue Erkenntnisse gewonnen werden (KRELL et al. 2012; OH & OH 2011). Gilbert, Boulter und Rutherford (1998) argumentieren, dass Modelle die Grundlage für alle fünf Bereiche des Erklärens bilden: wissentliche, beschreibende, interpretierende, kausale und voraussagende Erklärungen. Die Vorgänge des Modellierens und des kritischen Reflektierens über Modelle sind somit essentielle Elemente der naturwissenschaftlichen Erkenntnisgewinnung und finden sich verpflichtend in Steuerungsdokumenten für den naturwissenschaftlichen Unterricht in der Schule wieder (z. B. KMK 2005; NGSS LEAD STATES 2013). Dies zeigt den Ansatz eines authentischen, an naturwissenschaftlicher Praxis orientierten Unterrichts (CRAWFORD 2000; GILBERT 2004).

Van Driel und Verloop (1999) beschreiben theoriegeleitet sieben Charakteristika naturwissenschaftlicher Modelle:

1. Ein Modell steht immer in Beziehung zu einem Original, welches ein Objekt, ein Phänomen, ein Prozess oder ein System sein kann.

2. Ein Modell ist ein Werkzeug zum Forschen, um das Original zu beobachten oder an ihm Messungen vorzunehmen. Ein Modell als annähernd exakte Kopie, welches nur in seiner Größe zum Original geändert wurde, ist somit kein naturwissenschaftliches Modell.

3. Ein Modell kann nicht direkt mit dem Original interagieren, wie zum Beispiel ein Foto.

4. Ein Modell besitzt bestimmte Analogien zum Original, welche es erlauben, Hypothesen von dem Modell abzuleiten und zu testen. Dieser Vorgang erlaubt den Gewinn neuen Wissens über das Original.

5. Im Allgemeinen ist ein Modell so einfach wie möglich und abhängig vom Interesse der Forscher. Somit sind bewusst Merkmale vom Original im Modell nicht vorhanden.

6. Die Erstellung eines Modelles ist immer ein Kompromiss zwischen Gemeinsamkeiten und Unterschieden in Bezug zum Original, der von den Forschern in Abhängigkeit ihrer Forschungsfrage ausgehandelt wird.

7. Ein Modell wird durch einen sich wiederholenden Prozess entwickelt. Empirische Daten können dazu führen, dass das Modell überarbeitet werden muss, da sich neue Informationen zu dem Original ergeben haben.

Auf Basis dieser sieben Charakteristika und orientiert an Grosslight et al. (1991) haben Van Driel und Verloop (1999) 32 Items zu Modellen und dem Modellieren in den Naturwissenschaften entwickelt und mit einer vierstufigen Ratingskala versehen (*never – sometimes – usually – always*). Niederländische Lehrkräfte ( $N = 71$ ), die an einer Fortbildung im Rahmen der Einführung eines neuen naturwissenschaftlichen Unterrichtsfaches teilnahmen, haben die Items beantwortet (teilweise zu Hause). Die Probanden waren zu 80 % männlich, hatten eine durchschnittliche Unterrichtserfahrung von 17.5 Jahren und unterrichteten Biologie (35 %), Chemie (37 %) oder Physik (27 %). Van Driel und Verloop (1999) konnten die Items zu drei Faktoren bündeln (explorative Faktorenanalyse): 1) *relating models and targets* (11 Items), 2) *physical appearance of models* (8 Items) und 3) *social context of models* (8 Items; 5 Items luden auf keinem Faktor ausreichend stark). Die Zustimmung zu Faktor 1 ( $M = 2.99$ ) war relativ hoch, was ein Verständnis der Lehrkräfte von Modellen als vereinfachte Abbildung der Realität anzeigte. Eine nur etwas schwächere Zustimmung wur-

de gegenüber der Idee ausgedrückt, „that models are the products of human thought, creativity and communication“ (S. 1148; Faktor 3;  $M = 2.76$ ). Der Möglichkeit, dass Modelle unterschiedlichen Repräsentationsformen aufweisen können, wurde weniger stark zugestimmt (Faktor 2;  $M = 2.21$ ). Van Driel und Verloop (1999) schlussfolgerten trotzdem, „that the respondents allotted a large variety of representational modes to scientific models“ (S. 1148). Schließlich konnten die Autoren (fast) keine signifikanten Unterschiede im Antwortverhalten der Lehrkräfte in Abhängigkeit ihrer Unterrichtserfahrung oder -fächer beschreiben. Lediglich zwei Items aus Faktor 1 stimmten Chemie-Lehrkräfte signifikant stärker zu als Physik-Lehrkräfte.

Die Erfassung, Beschreibung und gegebenenfalls darauf aufbauende Förderung des Modellverstehens von Lehrkräften ist bedeutsam, da Lehrkräfte für die unterrichtspraktische Umsetzung von in Steuerungsdokumenten formulierten Vorgaben verantwortlich sind (BYBEE 2014). Das Modellverstehen sowie ein Wissen über adäquate Unterrichtsstrategien, dies bei Schülerinnen und Schülern zu entwickeln, wird daher als Teil des Professionswissens naturwissenschaftlicher Lehrkräfte betrachtet (TEPNER et al. 2012). Neben dem Aspekt, die Ergebnisse von Van Driel und Verloop (1999) durch eine Replikation zu generalisieren, skizziert diese Studie auch, welche Vorstellungen die befragten deutschen Lehrkräfte über Modelle besitzen.

### 3 Fragestellungen und Hypothesen

Der vorliegenden Studie liegen folgende Fragestellungen zu Grunde:

1. Inwieweit lassen sich die Befunde von Van Driel und Verloop (1999) bei einer Befragung deutscher Lehrkräften replizieren?

*Hypothese 1:* Es wird erwartet, dass sich die Ergebnisse von Van Driel und Verloop (1999) replizieren lassen, da die Studie in vielen bedeutenden internationalen Quellen als grundlegend aufgefasst wird. Mit dem Untersuchungsdesign, bei dem die wesentliche Variation in der Auswahl der Probandenpopulation liegt (SCHMIDT 2009; Tab. 1), soll eine Übertragung der Ergebnisse auf deutsche Lehrkräfte gelingen.

2. Inwiefern haben (a) die Unterrichtserfahrung, (b) der Schultyp und (c) die Fächerkombination der befragten deutschen Lehrkräfte einen Einfluss auf die Beantwortung des Fragebogens?



*Hypothese 2.1:* Sollten die Befunde von Van Driel und Verloop (1999) auf die deutschen Lehrkräfte übertragbar sein, ist zu erwarten, dass weder die Unterrichtserfahrung noch die Fächerkombination einen Einfluss auf die Beantwortung des Fragebogens haben.

*Hypothese 2.2:* Da die Arbeit mit Modellen sowohl Bestandteil der Rahmenlehrpläne für die Primar- (Senatsverwaltung 2004a, 2004b) als auch der Sekundarstufe (Senatsverwaltung 2006a, 2006b) ist, wird vermutet, dass die befragten Lehrkräfte beider Schultypen nicht signifikant unterschiedlich antworten.

## 4 Methodik

### 4.1 Testinstrument

Im Rahmen dieser Replikationsstudie wurden die 32 Items von Van Driel und Verloop (1999) aus dem Englischen und Niederländischen ungekürzt in das Deutsche übersetzt. Die deutsche Übersetzung wurde dem Erstautor der Ursprungsstudie zur Prüfung vorgelegt, der Korrekturen einzelner Formulierungen vorschlug, die umgesetzt wurden. Jedem Item kann auf einer vierstufigen Ratingskala unterschiedliche Zustimmung gegeben werden (*nie – manchmal – meistens – immer*). Der übersetzte Fragebogen wurde durch Lehramtsstudierende der Biologie ( $N = 46$ ) pilotiert. In einer anschließenden diskursiven Klärung wurde unter Hinzuziehung der Originaltexte die Formulierung einzelner Items auf Begriffsebene selektiv optimiert.

### 4.2 Stichprobe

Lehrkräfte von deutschen Oberschulen wurden um Teilnahme an der Studie gebeten, wobei der regionale Schwerpunkt der Datenerfassung in den Ländern Berlin (51 Schulen) und Brandenburg (11 Schulen) lag. Vereinzelt nahmen Lehrkräfte aus Schulen anderer Bundesländer teil, wenn ein direkter Kontakt hergestellt werden konnte (Hamburg, Niedersachsen und Schleswig-Holstein: je 2 Schulen; Baden-Württemberg, Bremen, Sachsen, Sachsen-Anhalt, Thüringen: je 1 Schule). Insgesamt bearbeiteten  $N = 226$  Lehrkräfte den Fragebogen (Gelegenheitsstichprobe). Die Befragten füllten den Fragebogen individuell und ohne Beaufsichtigung aus. Sie mussten mindestens ein MINT-Fach unterrichten. Außerdem wurde bei der Verteilung darauf geachtet, hauptsächlich Lehrende zu befragen, welche aktiv das Fach Biologie unterrichten (87,6 %). Es lassen sich drei großen Gruppen in Bezug auf den Erfahrungsstand der Lehrkräfte erkennen: 56 Lehrkräfte weisen eine geringe ( $< 6$  Jahre), 89 eine

mittlere (6–25 Jahre) und 78 eine hohe Berufserfahrung (> 25 Jahre) auf (Tab. 2).

**Tabelle 2:** Absolute (und prozentuale) Angaben zur Stichprobe

<b>Erfahrung</b>	bis 5 Jahre	6 bis 25 Jahre	mehr als 25 Jahre
Anzahl	56 (24.8 %)	89 (39.4 %)	78 (34.5 %)
<b>Geschlecht</b>	weiblich		männlich
Anzahl	134 (59.3 %)		91 (40.3 %)
<b>Schultyp</b>	Gymnasium		sonstige
Anzahl	156 (69.0 %)		70 (31.0 %)
<b>Fächerkombination</b>	nur MINT		nicht nur MINT
Anzahl	124 (54.9 %)		102 (45.1 %)

*Hinweis:* Vereinzelt haben Lehrkräfte keine (auswertbaren) Angaben gemacht.

### 4.3 Datenanalyse

Die Daten wurden zunächst deskriptiv analysiert und konfirmatorisch auf Passung mit der von Van Driel und Verloop (1999) berichteten Datenstruktur geprüft. Außerdem wurde wie in der zu replizierenden Studie (VAN DRIEL & VERLOOP 1999) eine explorative Faktorenanalyse mit allen 32 Items durchgeführt (paarweiser Fallausschluss). Um die interne Konsistenz der entstandenen Faktoren zu testen, wurde eine Reliabilitätsanalyse (Cronbachs  $\alpha$ ) durchgeführt.

## 5 Ergebnisse

Die konfirmatorische Prüfung der dreidimensionalen Datenstruktur nach Van Driel und Verloop (1999) ergab auf der Basis gängiger Indizes eine ungenügende Modellpassung (z. B. RMSEA = 0.089; CFI = 0.675; SCHERMELLEH-ENGEL, MOOSBRUGGER & MÜLLER, 2003). Das Rotationsverfahren (varimax) der explorativen Faktorenanalyse konvergierte nach 27 Iterationen. Es ergeben sich 12 Faktoren mit teilweise nicht mehr als 2 Items je Faktor. Dies entspricht nicht den Befunden von Van Driel und Verloop (1999), die eine dreifaktorielle Lösung mit elf sowie zweimal acht Items pro Faktor berichten. Der KMO-Test zeigt, dass der Datensatz für eine Faktorenanalyse geeignet ist (KMO = .69).

Alle MSA-Werte weisen darauf hin, dass die Einzelitems „ausreichend“ brauchbar sind ( $MSA > .50$ ; FIELD 2009). Der Bartlett-Test zeigt, dass die Korrelationen zwischen den Items hoch genug sind ( $p < .001$ ). Es konnten 23 Items in fünf Faktoren mit einer internen Konsistenz von  $\alpha > .50$  beschrieben werden (Tab. 3). Diese erklären zusammengenommen etwa 37 % der Varianz.

**Tabelle 3:** Die fünf Faktoren und die dazugehörigen Items

		Faktoren				
	Items	1	2	3	4	5
Kommunikation und Erkenntnisgewinnung	Ein Modell dient dazu, Hypothesen über das zu untersuchende Phänomen zu formulieren. (2)	.69				
	<i>Ein Modell dient der Kommunikation zwischen Wissenschaftler/-innen. (2)</i>	.62				
	Ein Modell ist ein Hilfsmittel für die Untersuchung von Phänomenen, die nicht direkt beobachtbar sind. (2)	.60				
	Die Entwicklung von Modellen ist ein Prozess, bei dem Hypothesen formuliert und überprüft werden, um dann das ursprüngliche Modell anzupassen. (1)	.57				
	<i>Ein Modell bezieht sich auf ein zu untersuchendes Phänomen. (2)</i>	.45				
	<i>Die Fragen der Forschenden steuern die Entwicklung eines Modells. (3)</i>	.34				
Repräsentation	<i>Ein Modell erscheint in Form eines Textes. (-)</i>		.74			
	<i>Ein Modell erscheint in Form einer oder mehrerer Formeln. (-)</i>		.64			
	<i>Ein Modell erscheint in Form einer Zeichnung. (2)</i>		.64			
	<i>Ein Modell erscheint in Form eines Fotos. (2)</i>		.63			
Veranschaulichung und Erklärung	Ein Modell dient dazu, einen Überblick über komplexe Phänomene zu bekommen. (3)			.71		
	Ein Modell dient dazu, ein Phänomen zu erklären. (1)			.71		
	Bei der Entwicklung von Modellen versucht man, möglichst viele irrelevante Aspekte des zu untersuchenden Phänomens unberücksichtigt zu lassen. (1)			.50		
	Ein Modell dient dazu, eine Vorstellung von einem abstrakten Konzept zu geben. (3)			.38		
	Ein Modell bildet die Ideen von Wissenschaftler/-innen ab. (3)			.35		
	Ein Modell ist eine vereinfachte Wiedergabe der Realität. (1)			.35		
Optimierungsprozess	Man versucht, Modelle so einfach wie möglich zu halten. (1)				.71	
	Bei Entwicklung von Modellen wird ein Optimum zw. Übereinstimmungen u. Unterschieden mit dem zu untersuchenden Phänomen gesucht. (-)				.69	
	In einem Modell nähert man sich der Realität aus einer logisch-rationalen Perspektive. (1)				.56	
	Im Laufe seiner Entwicklung weist ein Modell immer mehr Übereinstimmungen mit dem zu untersuchenden Phänomen auf. (1)				.36	
Gültigkeit	Ein Modell ist nur innerhalb eines begrenzten Zeitraums gültig. (3)					.73
	Ein Modell ist lediglich innerhalb eines bestimmten Fachgebiets oder Spezialgebiets gültig. (-)					.62
	Wissenschaftler/-innen verwenden die modernsten Modelle. (1)					.41
erklärte Varianz (%)		13.86	7.43	5.85	5.44	4.70
Median		2.83	1.75	3.00	2.75	2.33
Konsistenz		$\alpha = .68$	$\alpha = .63$	$\alpha = .59$	$\alpha = .59$	$\alpha = .53$

*Hinweis:* Die Zahlen in Klammern geben an, welchem Faktor die Items bei Van Driel und Verloop (1999) zugeordnet wurden. Ein Minus zeigt an, dass ein Item keinem Faktor zugeordnet wurde. Faktorladungen, deren absoluter Wert unter .30 lag, sind nicht aufgeführt.

Die fünf Faktoren können folgendermaßen benannt werden:

*Kommunikation und Erkenntnisgewinnung* (6 Items): Modelle dienen der Kommunikation zwischen Wissenschaftler/-innen und sind gleichzeitig Forschungswerkzeuge. Mit Modellen können Hypothesen über Phänomene formuliert und überprüft werden.

*Repräsentation* (4 Items): Modelle können in Form von Texten, Formeln, Zeichnungen oder Fotos erscheinen.

*Veranschaulichung und Erklärung* (6 Items): Modelle sind vereinfachte Darstellungen von etwas. Mit Modellen können Phänomene verstanden und erklärt werden.

*Optimierungsprozess* (4 Items): Die Entwicklung von Modellen ist ein rationaler Prozess, in dem das Verhältnis von Gemeinsamkeiten und Unterschieden zwischen Modell und Original optimiert wird.

*Gültigkeit* (3 Items): Modelle sind jeweils nur für eine bestimmte Zeit und innerhalb einzelner Disziplinen gültig. In der Wissenschaft werden die aktuellsten Modelle genutzt.

Die meisten Faktoren werden durchschnittlich zustimmend bewertet, mit Ausnahme der möglichen *Repräsentation* von Modellen (Median = 1.75). Der ablehnende und zustimmende Charakter der einzelnen Vorstellungen über Modelle variiert kaum, wenn Parameter wie z. B. der Schultyp der Lehrkräfte oder deren Fächerkombinationen berücksichtigt werden (Tab. 4).

**Tabelle 4:** Aufschlüsselung des Antwortverhaltens bei den unterschiedlichen Faktoren in Abhängigkeit des Schultyps, der Erfahrung und der Fachkombination der Lehrkräfte.

		Kommunikation und Erkenntnisgewinnung	Repräsentation	Veranschaulichung und Erklärung	Optimierungsprozess	Gültigkeit
ST	Gymnasium	2.83	1.75	3.00	2.75	2.33
	sonstige	3.00	2.00	3.00	3.00	2.33
Erfahrung	niedrig	2.67	2.00	3.08	2.75	2.30
	mittel	2.83	1.75	3.00	2.75	2.33
	hoch	3.00	1.75	3.00	3.00	2.33
Fächer	nur MINT	3.00	1.75	3.00	2.75	2.33
	nicht nur MINT	2.67	1.75	3.00	2.75	2.33

Anmerkung: Angegeben ist jeweils der Median. ST: Schultyp

Lediglich den Faktoren *Veranschaulichung und Erklärung* sowie *Optimierungsprozess* stimmen Lehrende, die an einem Gymnasium unterrichten, signi-

fikant schwächer zu, als die übrigen Lehrkräfte (in beiden Fällen:  $p < .05$ ; Mann-Whitney-Test). Außerdem ergeben sich signifikante Unterschiede im Antwortverhalten bei den Faktoren *Repräsentation* ( $p < .01$ ) und *Optimierungsprozess* ( $p < .05$ ) zwischen den Erfahrungsgruppen (Kruskal-Wallis-Test). Auch diese Befunde entsprechen nicht den Ergebnissen der Ursprungsstudie (VAN DRIEL & VERLOOP 1999). So berichten Van Driel und Verloop (1999), dass keine Unterschiede im Modellverstehen bei Lehrkräften mit verschiedener Unterrichtserfahrung bestehen. Auch zwischen Lehrkräften mit verschiedenen Unterrichtsfächern ergaben sich dort keine systematischen Unterschiede im Modellverstehen. Lediglich bei zwei Items konnten von Van Driel und Verloop (1999) signifikante Unterschiede im Zustimmungsverhalten zwischen Chemie- und Physiklehrern beschrieben werden.

## 6 Diskussion

Zusammenfassend kann festgehalten werden, dass sich die Ergebnisse von Van Driel und Verloop (1999) nicht replizieren ließen. Zum einen sind die Anzahl der Faktoren und ebenfalls deren Zusammensetzung aus den einzelnen Items verschieden. Erklären lässt sich dies vermutlich durch die unterschiedlichen Stichproben der beiden Studien. Allein der Umfang von 226 Befragten bei dieser Replikationsstudie gegenüber den 71 von Van Driel und Verloop (1999) kann maßgeblich für den Unterschied verantwortlich sein. Hinzu kommen weitere Einflussgrößen, wie die weitaus größere Spannweite der Berufserfahrung und die Unterrichtsfächer der Lehrkräfte. Außerdem wurde die ursprüngliche Studie vor mehr als zehn Jahren in den Niederlanden durchgeführt.

Des Weiteren könnte das Scheitern der Replikation neben inhaltlichen Gründen auch auf die Konzeption der Replikation zurückzuführen sein. Nach Schmidt (2009) sind leichte Veränderungen in den Klassen 2 und 4 (Kontext der Studie, Konstituierung der abhängigen Variablen; Tab. 1) bei einer generalisierenden Replikationsstudie möglich. Allerdings könnten diese in der hier vorliegenden Studie zu groß gewesen sein, wodurch sich das Konzept so stark verändert, dass die Ergebnisse von Van Driel und Verloop (1999) nicht übertragen werden konnten. Am wahrscheinlichsten ist aber dennoch, dass die im Gegensatz zu der ursprünglichen Studie stark differente Probandenpopulation der Hauptgrund für die nicht mögliche Übertragbarkeit der Ergebnisse ist.

Auch wenn sich die Befunde nicht replizieren und somit generalisieren ließen, können Aussagen über die befragten deutschen Lehrkräfte und ihr Verständnis über Modellen in der Wissenschaft getroffen werden. Die fünf Fakto-

ren, welche hier differenzierter analysiert wurden, zeigen eine vergleichsweise niedrige interne Konsistenz. Dennoch lassen sich Tendenzen erkennen, z. B. dass Lehrkräfte die Pluralität möglicher Repräsentationsformen von Modellen eher selten anerkennen. Im Gegensatz dazu weißt die tendenziell zustimmende Haltung bei dem Faktor *Veranschaulichung und Erklärung* darauf hin, dass das Verständnis von Modellen und die Vorstellung von ihrem Zweck eher medial sind und damit den bei Upmeier zu Belzen und Krüger (2010) beschriebenen basalen Niveaus entspricht.

Andere Faktoren wie der Schultyp der Lehrkräfte haben erwartungsgemäß keine bedeutsame und systematische Auswirkung auf das Modellverstehen. Die Unterschiede in den Faktoren *Repräsentation* und *Optimierungsprozess* zeigen, dass Lehrkräfte mit eher geringer Berufserfahrung die Pluralität möglicher Repräsentationsformen von Modellen tendenziell anerkennen und gleichzeitig Modelle weniger als abbildende Medien verstehen. Beides deutet auf ein vergleichsweise elaboriertes Modellverstehen hin (UPMEIER ZU BELZEN & KRÜGER 2010). Ursächlich könnte hier die weniger weit zurück liegende universitäre Ausbildungsphase sein, in der eventuell bereits aktuelle Sichtweisen zur Bedeutung von Modellen für den naturwissenschaftlichen Unterricht implementiert worden sind. Befunde von Krell und Krüger (2013) legen allerdings nahe, dass Lehrkräfte mit großer Berufserfahrung Modelle vergleichsweise häufig in ihrem Biologieunterricht einsetzen. Es ist daher offen, inwieweit ein systematischer Zusammenhang zwischen dem Modellverstehen von Lehrerinnen und Lehrern und dem Einsatz von Modellen in ihrem (Biologie-) Unterricht besteht (vgl. VAN DRIEL & VERLOOP 2002).

## 7 Fazit

Diese Studie zeigt, dass zumindest auf die ausgewählte Stichprobe deutscher Lehrkräfte die Ergebnisse von Van Driel und Verloop (1999) nicht übertragbar sind. Deren Befunde sind somit nicht geeignet, als ein Beleg für ein wenig ausgeprägtes Modellverstehen von Lehrkräften weltweit zitiert zu werden (z. B. JUSTI & GILBERT 2003; KHAN 2011; SCHWARZ 2009). Insgesamt scheint es trotz aller Probleme, die Replikationsstudien in der fachdidaktischen Forschung mit sich bringen, sehr wichtig, dieses Überprüfungselement zukünftig weit häufiger als bisher üblich in Bezug auf empirische Studien durchzuführen. Aus den Ergebnissen lassen sich trotzdem Rückschlüsse auf ein eher wenig elaboriertes Modellverstehen der befragten Lehrkräfte ziehen. Deutsche Lehrkräfte verstehen Modelle offenbar eher als Mittel der Veranschaulichung und Erklärung und

stehen ihren mannigfaltigen Repräsentationsformen tendenziell ablehnend gegenüber. Abschließend sollte bedacht werden, dass eine gescheiterte Replikation nicht ausreicht, um eine vorhandene Studie zu falsifizieren.

## Zitierte Literatur

- AMIR, Y., & SHARON, I. (1991). Replication research: A "must" for the scientific advancement of psychology. In J. Neuliep (Hrsg.), *Replication research in the social sciences* (S. 51–69). Newbury Park, CA: Sage.
- BYBEE, R. (2014). NGSS and the next generation of science teachers. *Journal of Science Teacher Education*, 25, 211–221.
- CHAPLIN, J. P. & KRAWIEC, T. S. (1979). *Systems and theories of psychology*. New York, NY: Holt, Rinehart & Watson.
- CRAWFORD, B. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science*, 37, 916–937.
- FANELLI, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90, 891–904.
- FIELD, A. (2009). *Discovering statistics using SPSS*. Los Angeles, CA: Sage.
- GILBERT, J., BOULTER, C. & RUTHERFORD, M. (1998). Models in explanations, Part 1: Horses for courses? *International Journal of Science Education*, 20, 83–97.
- GILBERT, J. (2004). Models and modelling: Routes to more authentic science education. *International Journal of Science and Mathematics Education*, 2, 115–130.
- GROSSLIGHT, L., UNGER, C., JAY, E., & SMITH, C. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28, 799–822.
- HENDRICK, C. (1991). Replications, strict replications, and conceptual replications: Are they important? In J. Neuliep (Hrsg.), *Replication research in the social sciences* (S. 41–49). Newbury Park, CA: Sage.
- IOANNIDIS, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2, 696–701.
- JUSTI, R., & GILBERT, J. (2003). Teachers' views on the nature of models. *International Journal of science education*, 25, 1369–1386.
- KHAN, S. (2011). What's missing in model-based teaching. *Journal of Science Teacher Education*, 22, 535–560.
- KRELL, M. (2013). *Wie Schülerinnen und Schüler biologische Modelle verstehen*. Berlin: Logos.
- KRELL, M., & KRÜGER, D. (2013). Wie werden Modelle im Biologieunterricht eingesetzt? Ergebnisse einer Fragebogenstudie. *Erkenntnisweg Biologiedidaktik*, 12, 9–26.
- KRELL, M., UPMEIER ZU BELZEN, A., & KRÜGER, D. (2012). Students' understanding of the purpose of models in different biological contexts. *International Journal of Biology Education*, 2, 1–34. Retrieved from [http://www.ijobed.com/2\\_2/Moritz-2012.pdf](http://www.ijobed.com/2_2/Moritz-2012.pdf)
- LAMAL, P. A. (1991). On the importance of replication. In J. Neuliep (Hrsg.), *Replication research in the social sciences* (S. 31–35). Newbury Park, CA: Sage.
- NEULIEP, J., & CRANDALL, R. (1991). Editorial bias against replication research. In J. Neuliep (Hrsg.), *Replication research in the social sciences* (S. 85–90). Newbury Park, CA: Sage.
- NEULIEP, J., & CRANDALL, R. (1993). Reviewer bias against replication research. *Journal of Social Behavior and Personality*, 8, 21–29.
- NGSS LEAD STATES (Ed.). (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.
- OH, P., & OH, S. (2011). What teachers of science need to know about models: An overview. *International Journal of Science Education*, 33, 1109–1130.

- POPPER, K. (1959). *The logic of scientific discovery*. New York, NY: Basic Books, Inc.
- ROSENTHAL, R. (1991). Replication in behavioral research. In J. Neuliep (Hrsg.), *Replication research in the social sciences* (S. 1–30). Newbury Park, CA: Sage.
- SCHERMELLEH-ENGEL, K., MOOSBRUGGER, H., & MÜLLER, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, 23–74.
- SCHMIDT, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13, 90-100.
- SCHRAW, G., CRIPPEN, K. J., & HARTLEY, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36, 111-139.
- SCHWARZ, C. (2009). Developing preservice elementary teachers' knowledge and practices through modeling-centered scientific inquiry. *Science Education*, 93, 720-744.
- SCHWARZ, C., & GWEKWERERE, Y. N. (2007). Using a guided inquiry and modeling instructional framework (EIMA) to support preservice K-8 science teaching. *Science Education*, 91, 158-186.
- SCHWARZ, C., & WHITE, B. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction*, 23, 165–205.
- SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BRD (KMK) (Hrsg.) (2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. München & Neuwied: Wolters Kluwer.
- SENATSV ERWALTUNG FÜR BILDUNG, JUGEND UND SPORT BERLIN (2004a). *Rahmenlehrplan Grundschule. Sachunterricht*. Berlin: Wissenschaft und Technik Verlag. Online verfügbar unter [https://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/gr\\_sach\\_1\\_4.pdf?start&ts=1394618125&file=gr\\_sach\\_1\\_4.pdf](https://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/gr_sach_1_4.pdf?start&ts=1394618125&file=gr_sach_1_4.pdf)
- SENATSV ERWALTUNG FÜR BILDUNG, JUGEND UND SPORT BERLIN (2004b). *Rahmenlehrplan Grundschule. Naturwissenschaften*. Berlin: Wissenschaft und Technik Verlag. Online verfügbar unter [http://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/gr\\_natur.pdf?start&ts=1394618125&file=gr\\_natur.pdf](http://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/gr_natur.pdf?start&ts=1394618125&file=gr_natur.pdf)
- SENATSV ERWALTUNG FÜR BILDUNG, JUGEND UND SPORT BERLIN (2006a). *Rahmenlehrplan für die Sekundarstufe I. Biologie*. Berlin: Oktoberdruck. Online verfügbar unter [http://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/sek1\\_biologie.pdf?start&ts=1394618125&file=sek1\\_biologie.pdf](http://www.berlin.de/imperia/md/content/sen-bildung/schulorganisation/lehrplaene/sek1_biologie.pdf?start&ts=1394618125&file=sek1_biologie.pdf)
- SENATSV ERWALTUNG FÜR BILDUNG, JUGEND UND SPORT BERLIN (2006b). *Rahmenlehrplan für die Sekundarstufe II. Biologie*. Berlin: Oktoberdruck. Online verfügbar unter [https://www.berlin.de/imperia/md/content/sen-bildung/unterricht/lehrplaene/sek2\\_biologie.pdf](https://www.berlin.de/imperia/md/content/sen-bildung/unterricht/lehrplaene/sek2_biologie.pdf)
- SILVA, C. C. (2007). The role of models and analogies in the electromagnetic theory: A historical case study. *Science & Education*, 16, 835-848.
- TEPNER, O., BOROWSKI, A., DOLLNY, S., FISCHER, H., JÜTTNER, M., KIRSCHNER, S., ... WIRTH, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- UPMEIER ZU BELZEN, A., & KRÜGER, D. (2010). Modellkompetenz im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 41–57.
- VAN DRIEL, J., & VERLOOP, N. (1999). Teachers' knowledge of models and modelling in science. *International Journal of Science Education*, 21, 1141–1153.
- VAN DRIEL, J., & VERLOOP, N. (2002). Experienced teacher's knowledge of teaching and learning of models and modeling in science education. *International Journal of Science Education*, 24, 1255–1277.
- YONG, E. (2012). In the wake of high profile controversies, psychologists are facing up to problems with replication. *Nature*, 483, 298-300.