

Peter Furthmüller

Methodenbericht

Systemmonitoring 2012: Stichprobe und Erhebungsverfahren

StEG

STUDIE ZUR ENTWICKLUNG VON GANZTAGSSCHULEN



In Zusammenarbeit mit dem StEG-Konsortium und den Mitarbeiterinnen und Mitarbeitern des StEG-Teams

Gefördert von:



Durchgeführt von:



Die *Studie zur Entwicklung von Ganztagschulen (StEG)* ist ein bundeslandübergreifendes Forschungsprogramm zur empirischen Analyse schulischer Veränderungsprozesse und individueller Entwicklungen von Schülerinnen und Schülern an Ganztagschulen. Die Studie wird von Mitarbeiterinnen und Mitarbeitern aus mehreren Forschungseinrichtungen in umfassender Zusammenarbeit durchgeführt. Den Arbeitsverbund bilden das Deutsche Institut für Internationale Pädagogische Forschung (DIPF), das Institut für Schulentwicklungsforschung (IFS), das Deutsche Jugendinstitut (DJI) sowie die Justus-Liebig-Universität Gießen (JLU). StEG wird durch das Bundesministerium für Bildung und Forschung gefördert.

Die in dieser Dokumentation dargestellten Materialien wurden von den Mitarbeiterinnen und Mitarbeitern der an StEG beteiligten Institute gemeinschaftlich erarbeitet.

© 2016 Studie zur Entwicklung von Ganztagschulen (StEG)

Ansprechpartner:
Peter Furthmüller
Deutsches Jugendinstitut e.V.
Studie zur Entwicklung von Ganztagschulen
Nockherstraße 2, 91541 München
Telefon: +49 (0)89 6 23 06-0
Fax: +40 (0)89 6 23 06-162
Internet: <http://www.dji.de>
E-Mail: furthmueller@dji.de

Deutsches Institut für Internationale Pädagogische Forschung
Schloßstraße 29
60486 Frankfurt am Main
Internet: <http://www.dipf.de>

Institut für Schulentwicklungsforschung der Technischen Universität Dortmund
Vogelpothsweg 78
44227 Dortmund
Internet: <http://www.ifs-dortmund.de>

Justus-Liebig-Universität Gießen
Ludwigstraße 23
35390 Gießen
Internet: <http://www.uni-giessen.de>

Version vom 30. November 2016
Vervielfältigung und Verbreitung mit Quellenangabe gestattet.

Inhaltsverzeichnis

1	Einleitung	6
2	Stichprobendesign	7
2.1	Grundgesamtheit	7
2.2	Stichprobenrahmen	8
2.3	Stichprobenziehung	9
3	Datenerhebungsmethode und Feldprozeduren	12
3.1	Erhebungsinstrument	12
3.2	Feldarbeit	16
4	Realisierte Stichprobe	17
4.1	Ausfallgründe	17
4.2	Ausschöpfung	20
4.3	Systematik der Ausfälle	22
4.4	Bearbeitungsdauer und Fragebogenabbruch	29
5	Gewichtung	32
5.1	Designgewicht	32
5.2	Ausfallgewicht und Gesamtgewicht	34
5.3	Normierung des Gesamtgewichts	35
6	Scientific-Use-File	36
6.1	Datenaufbereitung	36
6.2	Konventionen und Bedienungshilfen	39

Literaturverzeichnis	42
-----------------------------	-----------

Tabellenverzeichnis

1	Stichprobenrahmen StEG-Schulleitungsbefragung 2012	9
2	Bruttostichprobe der StEG-Schulleitungsbefragung 2012	11
3	Gezogene Schulen und verfügbare Ersatzschulen	12
4	Versandtermine der Kontaktierungsschreiben zur Schulleitungsbefragung 2012	17
5	Ausfallgründe und Umfang der Nettostichprobe von Schulen mit Primarstufe	18
6	Ausfallgründe und Umfang der Nettostichprobe von Schulen mit Sekundarstufe I	18
7	Ausfallgründe und Umfang der Nettostichprobe von Gymnasien	19
8	Ausschöpfungsquoten nach Schulgruppen und Bundesländern	20
9	Chi-Quadrat-Test der Schulgruppen auf Ausfall und Teilnahme	22
10	Zusammenfassung der Mittelwertvergleiche von Schulen mit Primarstufe zur bundeslandspezifischen Beteiligung an StEG	23
11	Zusammenfassung der Mittelwertvergleiche von Schulen mit Sekundarstufe I zur bundeslandspezifischen Beteiligung an StEG	24
12	Zusammenfassung der Mittelwertvergleiche von Gymnasien zur bundeslandspezifischen Beteiligung an StEG	25
13	Schulgruppenspezifische Chi-Quadrat-Tests von Organisationsform und Ausfällen	26
14	Schulgruppenspezifische Chi-Quadrat-Tests von siedlungsstrukturellen Kreistypen und Ausfällen	28
15	Auswahlwahrscheinlichkeiten für Schulen im Systemmonitoring	33

Abbildungsverzeichnis

1	Typische Seite aus dem StEG-Onlinefragebogen 2012	14
2	Verbleib in der Online-Befragung nach Fragennummer	30
3	Befragungsdauer vollständig bearbeiteter Fragebogen (mit Befragungsbeginn und -ende am gleichen Tag)	31
4	Normierte Gesamtgewichte im Datensatz	36

1 Einleitung

Die *Studie zur Entwicklung von Ganztagschulen* (StEG) ist ein umfassendes Forschungsprojekt zur Untersuchung der Ganztagschullandschaft in Deutschland, das seit 2004 von einem Konsortium sozialwissenschaftlicher Institute durchgeführt wird. Dieser Forschungsverbund besteht aus Mitarbeiterinnen und Mitarbeitern des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) unter Leitung von Prof. Dr. Eckhard Klieme, des Instituts für Schulentwicklungsforschung (IFS) unter Leitung von Prof. Dr. Heinz Günter Holtappels, des Deutschen Jugendinstituts (DJI) unter Leitung von Prof. Dr. Thomas Rauschenbach und der Justus-Liebig-Universität Gießen (JLU) unter Leitung von Prof. Dr. Ludwig Stecher. StEG wird durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert. Das Projekt wird von einem administrativen Beirat beraten, der sich aus Vertretern der Länder und des Städtetages zusammensetzt. Im wissenschaftlichen Beirat wird StEG durch Experten aus der Bildungs-, Schul- und Jugendforschung sowie der Sozialpädagogik unterstützt.

In seiner zweiten Förderphase von 2012 bis 2015 verfolgte das StEG-Projekt u.a. das Ziel, die deutsche Ganztagschullandschaft im Rahmen eines repräsentativen Systemmonitorings zu beobachten, um sowohl aktuelle Erkenntnisse über Struktur, Ausgestaltung und Inhalt des Ganztagsbetriebs zu gewinnen, als auch Trendentwicklungen beschreiben zu können. In diesem Bericht sind methodische Hintergründe zum ersten Systemmonitoring von 2012 zusammengefasst und für die Verwendung mit dem Scientific-Use-File dokumentiert. Einen Schwerpunkt bildet dabei die Stichprobenkonzeption und das Ziehungsverfahren von Ganztagschulen (Abschnitt 2). Details zu Datenerhebungsmethoden, Online-Fragebogen und Feldprozeduren sind in Abschnitt 3 dargestellt. Im Bericht werden außerdem die realisierte Stichprobe, die Ausschöpfungsquoten und Ausfallgründe betrachtet (Abschnitt 4). In den abschließenden Teilen wird näher auf die Berechnung der Gewichte (Abschnitt 5) und die Beschreibung des Datensatzes eingegangen (Abschnitt 6) .

An dieser Stelle möchte das StEG-Team allen Personen danken, die das Projekt beim Systemmonitoring 2012 unterstützt haben. Wir danken Dr. habil. Ulrich Pötter und Holger Quellenberg vom Deutschen Jugendinstitut (DJI) für die fachliche Beratung bei der Stichprobenkonzeption. Unser Dank gilt Dr. Sabine Meinck vom IEA Data Processing Center (DPC) für die Bereitstellung der Designgewichte und ihre hilfreichen Anmerkungen zum Gewichtungungsverfahren. Elham Müller von der Abteilung Technology Based Assessment (TBA) des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) danken wir für ihren unermüdlichen Einsatz bei der Programmierung und Testung des Fragebogens. Und nicht zuletzt möchten wir uns bei allen Schulleitungen und Ganztagsverantwortlichen bedanken, die an der Studie teilgenommen und somit das Projekt erst ermöglicht haben.

2 Stichprobendesign

2.1 Grundgesamtheit

Ganztagsschulen werden in den einzelnen Bundesländern zum Teil nach unterschiedlichen Kriterien definiert, die im Rahmen des StEG-Systemmonitorings möglichst gleichberechtigt behandelt werden sollten: Jede öffentliche Schule sollte eine Chance haben in die Stichprobe zu gelangen und an der Studie teilzunehmen, sofern es sich um eine Ganztagschule bzw. „Schule mit ganztägigen Angeboten“ nach der jeweiligen Landesdefinition handelte. Berufsbildende Schulen und Förderschulen sollten – nach Absprache mit dem wissenschaftlichen und administrativen Beirat von StEG – nicht berücksichtigt werden, da Förderschulen z.B. eine sehr heterogene Gruppe darstellen und eine andere Ganztagsstradition vorweisen.

Darüber hinaus sollten die erhobenen Daten sowohl unverzerrte Aussagen über einzelne Bundesländer ermöglichen als auch auf die gesamte Ganztagschullandschaft Deutschlands generalisierbar sein. Dieser Anspruch an eine bundesweite und landesspezifische Repräsentativität legte nahe, die Grundgesamtheit nach Bundesländern zu stratifizieren und aus jeder Schicht Ganztagschulen entsprechend ihres proportionalen Anteils in Deutschland zu ziehen.

Da der Ganztag im Primarbereich häufig anders konzipiert ist als im Sekundarbereich, sollten jeweils voneinander unabhängige Stichproben gezogen werden. Eine besondere Bedeutung kommt in der empirischen Bildungsforschung dem Vergleich von Gymnasien mit anderen Schulen der Sekundarstufe I zu, da sich diese Schularten in vielen Aspekten voneinander unterscheiden (z.B. soziale Zusammensetzung der Schülerschaft, Umfang der unterrichtlichen Pflichtstundentafel) und in Studien häufig weitere Unterschiede identifiziert werden. Es erschien daher sinnvoll, neben der Einteilung der Grundgesamtheit in den Primar- und Sekundarbereich auch für Gymnasien und andere Schulen der Sekundarstufe I separate Schichten zu bilden.¹ Um sowohl die Besonderheiten des deutschen Schulsystems als auch wissenschaftliche Erfordernisse sinnvoll berücksichtigen zu können, wurden in Absprache mit dem wissenschaftlichen Beirat für das StEG-Systemmonitoring eine Einteilung der Schulen in drei Schulgruppen beschlossen:

- Schulen mit Primarstufe (Primar),
- Schulen mit Sekundarstufe I ohne Gymnasien (Sek. I) sowie
- Gymnasien (Gym.)

Für die weitere Stichprobenkonzeption wurde diese schulgruppenspezifische Aufteilung der Grundgesamtheit verwendet.

¹Eine Unterteilung nach weiteren Schularten der Sekundarstufe I sollte hingegen nicht umgesetzt werden, da die existierenden Schulformen bundeslandspezifisch unterschiedlich verteilt sind und dadurch die Fallzahlen für die einzelnen Schulformen in der realisierbaren Stichprobe meist sehr klein geworden wären.

2.2 Stichprobenrahmen

Die Stichprobenziehung erfolgte anhand von Bestandslisten, die dem StEG-Forschungsverbund von den Kultusministerien der Länder zur Verfügung gestellt wurden und die alle Ganztagschulen nach den Definitionen der Länder aus dem Schuljahr 2011/2012 enthielten. Für Baden-Württemberg musste auf die Bestandsliste des Schuljahres 2010/2011 zurückgegriffen werden, da noch keine Fassung für 2011/2012 vorlag. Da Sachsen-Anhalt und Mecklenburg-Vorpommern nur über sehr wenige Ganztagschulen im Primarbereich verfügten und Gymnasien aus Berlin 2011/2012 noch nicht umfassend im Ganztage organisiert waren, sollten die entsprechenden Strata unbesetzt bleiben.

Für Baden-Württemberg, Berlin, Bremen, Hessen, Niedersachsen und Rheinland-Pfalz wurden außerdem Ganztagschulen gelistet, die sowohl über einen Primar- als auch über einen Sekundarbereich verfügten. Eine Mehrfachziehung bzw. Mehrfachbefragung dieser Schulen in mehreren Schulgruppen sollte vermieden werden. Daher wurde zu den drei Schulgruppen eine zusätzliche Schicht *Schulen mit Primar- und Sekundarstufe I* gebildet, deren Elemente für eine Zufallsziehung in die Schulgruppe der *Schulen mit Primarstufe* oder der *Schulen mit Sekundarstufe I* vorgemerkt wurden. Nach dieser Präzisierung ergab sich ein Stichprobenrahmen von insgesamt 11.366 auswählbaren Elementen in 51 expliziten Strata (vgl. besetzte Zellen in Tabelle 1).

Die Kultusministerien hatten auf den Listen außerdem die Organisationsformen der Ganztagschulen angegeben. Anhand der bundeslandspezifischen Bezeichnungen ließen sich die Organisationsformen im Stichprobenrahmen einer der Kategorien *gebunden*, *teilgebunden*, *offen*, *verschiedene Betreuungsangebote* oder *keine Angabe zum Betreuungsangebot* zuordnen. Im Anschluss an die Kategorisierung wurde der Stichprobenrahmen separat für jedes Bundesland nach diesem Merkmal sortiert. Durch diese implizite Stratifizierung sollte die Heterogenität der Stichprobe in Hinblick auf die Organisationsform verbessert werden.

Tabelle 1: Stichprobenrahmen StEG-Schulleitungsbefragung 2012

	Primar	Primar & Sek. I	Sek. I	Gym.	Gesamt
BB	180	14	97	33	324
BE	372	14	104	-	490
BW ¹	135	123	206	99	563
BY	218	-	974	218	1.410
HB	23	-	31	3	57
HE	266	37	275	99	677
HH	73	-	35	59	167
MV	-	-	137	36	173
NI	521	45	504	161	1.231
NW	2.668	-	653	127	3.448
RP	277	30	205	66	578
SH	135	-	186	64	385
SL	151	-	62	28	241
SN	597	-	266	115	978
ST	-	-	68	23	91
TH	431	-	106	16	553
D	6.047	263	3.909	1.147	11.366

¹ Für Baden-Württemberg beziehen sich die Zahlen auf das Schuljahr 2010/2011

Zur Bildung eines Stichprobenrahmens wurden die Bestandslisten der Länder vom IEA Data Processing and Research Center (DPC) in Hamburg mit einer einheitlichen Struktur versehen und geprüft. Einige Schullisten wiesen doppelt genannte Schulen auf, die anhand identischer Schulnamen und -adressen vom DPC identifiziert und soweit wie möglich gelöscht wurden. Aus Zeit- und Kostengründen war allerdings keine vollständige Säuberung aller Duplikate möglich. Das DPC schätzte den Restbestand an doppelten Einträgen auf 4 Prozent des Stichprobenrahmens. Die Qualität der Stichprobe muss vor diesem Hintergrund beurteilt werden.

2.3 Stichprobenziehung

Wie sich am Stichprobenrahmen zeigen lässt, verteilen sich die Ganztagschulen der drei Schulgruppen äußerst unterschiedlich auf die einzelnen Bundesländer. Neben besonders stark besetzten Schichten (z.B. Schulen mit Primarstufe aus Nordrhein-Westfalen) gibt es Bundesländer, in denen für bestimmte Schulgruppen nur sehr wenige Institutionen existieren (z.B. Gymnasien aus Bremen). Ein streng proportional geschichteter Stichprobenplan mit minimalem Umfang hätte dazu geführt, dass viele Bundesländer nur durch sehr wenige Schulen repräsentiert worden und keine

aussagekräftigen Analysen auf Landesebene mehr möglich gewesen wären.² Daher sollte die Allokation in einigen Schichten durch Ziehung zusätzlicher Schulen erhöht und Abweichungen der proportionalen Länderanteile in der finalen Gesamtstichprobe durch Designgewichte kompensiert werden. Die disproportional geschichtete Stichprobe sollte sich aber an den Proportionen in der Grundgesamtheit *orientieren*, um die Unterschiede in den Designgewichten möglichst gering zu halten (vgl. Abschnitt 5).

In einem ersten Schritt wurde pro Schulgruppe die Mindestanzahl an Ganztagschulen ermittelt, die für eine proportional geschichtete Stichprobe unabdingbar war. Da die Schulstichprobe des Systemmonitorings auch die Grundlage für Schülerbefragungen in weiteren Teilprojekten des StEG-Konsortiums darstellen sollte, wurde z.B. auch über Power-Analysen eingeschätzt, wie viele *Primäreinheiten* (Schulen) für eine hinreichende Untersuchung von *Sekundäreinheiten* (Schülerinnen und Schüler) aufgenommen werden sollten. Die Zellbesetzung des Stichprobenplans ist in der Schulgruppe der *Gymnasien* etwas niedriger, da es in Deutschland deutlich weniger Ganztagsgymnasien als ganztägig organisierte *Schulen mit Primarstufe* oder *Schulen mit Sekundarstufe I* gibt.

Aufgrund von Erfahrungen aus der ersten Projektphase wurde ein Ausfall von etwa 25 Prozent der kontaktierten Schulen vermutet und der Stichprobenumfang in jeder Schulgruppe um ein Viertel der Fälle angehoben. Das DPC hat dem Projekt außerdem geraten, die Strata von *Schulen mit Primarstufe*, *Schulen mit Sekundarstufe I* und *Gymnasien* mindestens mit 25 Schulen pro Bundesland zu besetzen, um sichere Aussagen auf Länderebene treffen zu können. In Zuge dessen sollte ein überbordender Gesamtstichprobenumfang vermieden werden, indem zum Ausgleich weniger Schulen aus solchen Bundesländern gezogen werden, die über einen überdurchschnittlich hohen Anteil an Ganztagschulen verfügten. In jenen Ländern hingegen, in denen keine 25 Ganztagschulen zur Verfügung standen, wurde eine Vollerhebung innerhalb der entsprechenden Schulgruppe angestrebt.

Die Stichprobenziehung erfolgte im April 2012 am DPC über eine systematische Zufallsauswahl aus den 51 Strata. Für die Schicht *Schulen mit Primarstufe und Sekundarstufe I* wurde ein zweiter Ziehungsschritt durchgeführt, um die gezogenen Einheiten jeweils zufällig der Schulgruppe *Schulen mit Primarstufe* oder *Schulen mit Sekundarstufe I* zuzuweisen. Kurz nach der Ziehung teilte das Kultusministerium in Bremen dem StEG Forschungsverbund mit, dass auch für Bremer *Schulen mit Sekundarstufe I* eine Vollerhebung erwünscht war. Die Fallzahl konnte in dieser Schicht um sechs zusätzliche Schulen angehoben werden, wodurch sich der Stichprobenplan auf insgesamt 1.885 Schulen erweitert hat. Dieser angestrebte Stichprobenplan wird im Folgenden als *Bruttostichprobe* bezeichnet und u.a. zur Berechnung von Ausschöpfungsquoten herangezogen (vgl. Tabelle 2).

²Das StEG-Team bedankt sich bei Dr. habil. Ulrich Pötter (DJI), Holger Quellenberg (DJI) und Ivo Züchner (DIPF) für die Beratung und die Kalkulation der benötigten Stichprobengrößen.

Tabelle 2: Bruttostichprobe der StEG-Schulleitungsbefragung 2012

	Primar	Sek. I	Gym.	Gesamt
BB	27	30	20	77
BE	44	26	-	70
BW	29	62	25	116
BY	25	160	40	225
HB	23	31 ¹	3	57
HE	34	56	25	115
HH	25	25	25	75
MV	-	26	25	51
NI	54	88	35	177
NW	250	120	28	398
RP	34	42	25	101
SH	25	30	25	80
SL	25	25	25	75
SN	60	45	25	130
ST	-	25	23	48
TH	49	25	16	90
D	704	816	365	1885

¹Die Fallzahl wurde auf Wunsch des Kultusministeriums von 25 auf 31 Schulen angehoben (Vollerhebung)

Um die Ausschöpfung zu verbessern, wurden nach Möglichkeit für jede in die Bruttostichprobe gezogene Schule zwei sogenannte *Ersatzschulen*³ bereitgehalten, die bei einem Teilnahmeausfall der gezogenen Schule kontaktiert werden konnten. Als Auswahlkriterium galt, dass die Ersatzschulen aus der gleichen expliziten und impliziten Schicht stammen sollten, wie die ursprünglich gezogene Schule. Dies wurde technisch durch eine entsprechende Sortierung der Schulliste erreicht: Als 1. Ersatzschule wurde das zur gezogenen Schule jeweils nächste Element, als 2. Ersatzschule das jeweils vorangegangene Element auf der Schulliste ausgewählt. Durch Ersetzen der gezogenen Schule durch eine möglichst „ähnliche“ Schule soll die Chance erhöht werden, eine Ersatzschule mit gleichen Eigenschaften auszuwählen und den Nonresponse-Bias zu reduzieren (vgl. ASSMANN ET AL. 2011: 61; STURGIS ET AL. 2006: 31ff). Insgesamt konnte mit diesem Vorgehen auf 5.232 Adressen zurückgegriffen werden, um im besten Fall eine Stichprobe von 1.885 Fällen zu realisieren (vgl. Tabelle 3).

³Die in StEG als „Ersatzschulen“ bezeichneten Stichprobenelemente sind keine Schulen in privater Trägerschaft gemäß Art. 7 Abs. 4 GG, sondern jene öffentlichen Ganztagschulen, die in die Stichprobe aufgenommen wurden, wenn eine der ursprünglich gezogenen Schulen ausgefallen ist.

Tabelle 3: Gezogene Schulen und verfügbare Ersatzschulen

	Primar			Sek. I			Gym.		
	Gez.	1. Ers.	2. Ers.	Gez.	1. Ers.	2. Ers.	Gez.	1. Ers.	2. Ers.
BB	27	25	25	30	26	22	20	13	-
BE	44	44	44	26	26	25	-	-	-
BW	29	29	29	62	62	62	25	25	25
BY	25	25	25	160	160	160	40	40	40
HB	23	-	-	31	-	-	3	-	-
HE	34	34	34	56	56	56	25	25	25
HH	25	25	22	25	10	-	25	25	9
MV	-	-	-	26	26	26	25	11	-
NI	54	54	54	88	88	87	35	35	35
NW	250	250	250	120	120	120	28	28	28
RP	34	34	34	42	42	42	25	25	15
SH	25	25	25	30	30	30	25	25	14
SL	25	25	25	25	25	12	25	3	-
SN	60	60	60	45	45	45	25	25	25
ST	-	-	-	25	25	18	23	-	-
TH	49	49	49	25	25	25	16	-	-
D	704	679	676	816	766	730	365	280	216

3 Datenerhebungsmethode und Feldprozeduren

3.1 Erhebungsinstrument

Die Daten für das Systemmonitoring 2012 wurden über einen standardisierten Fragebogen erhoben, der passwortgeschützt über das Internet ausgefüllt werden konnte. Alternativ bestand für die Schulen die Möglichkeit, statt der Online-Variante eine gedruckte Version des Fragebogens anzufordern und auszufüllen.⁴

Der Fragebogen enthielt insgesamt 85 Fragen, darunter allgemeine Fragen zur Schule sowie spezifische Fragen zum Ganztagsbetrieb, zu den institutionellen Ressourcen, zur Schulentwicklung und -praxis, zu Lehrkräften, pädagogischem Personal und Kooperationspartnern sowie zu den Ganztagsangeboten. Für die drei Schulgruppen wurden immer nur jene Fragen und Antwortmöglichkeiten angezeigt, die für sie relevant waren. Maximal umfasste der Fragebogen für *Schulen mit Primarstufe* 70 Fragen, für *Schulen mit Sekundarstufe I* 77 Fragen und für *Gymnasien* 68 Fragen.

⁴Diese Möglichkeit wurde allerdings nur von sechs Schulen in Anspruch genommen.

Durch Filterung von Fragen und Items ließ sich der Befragungsaufwand für die Schulen reduzieren. Eine weitere Maßnahme zur Verkürzung der Befragungsdauer bestand darin, dass die Teilnehmerinnen und Teilnehmer die aktuell angezeigte Frage an das Ende des Fragebogens verschieben konnten. Dadurch sollte eine schnellere Bearbeitung des Fragebogens ermöglicht werden (z.B. indem Fragen hinten ange stellt werden konnten, für deren Beantwortung weitere Unterlagen benötigt wurden oder telefoniert werden musste). Ferner bestand für die Schulleitungen die Mög lichkeit, den Fragebogen zu unterbrechen und zu einem späteren Zeitpunkt ab der zuletzt aufgerufenen Position fortzusetzen.

Der Fragebogen wurde für StEG unter dem Einsatz von PHP, HTML und MySQL von der Abteilung Technology Based Assessment (TBA) am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) programmiert und seine Funktionalität mit unterschiedlichen Internetbrowsern ausgiebig getestet. Die Anforderungen an die Hardware- und Softwareausstattung wurden bewusst sehr niedrig angesetzt, um keine Ausfälle wegen technischer Inkompatibilität zu erzeugen.

Als HTML-Formularelemente wurden ausschließlich Radiobuttons, Checkboxes und Eingabefelder eingesetzt, da z.B. Drop-Down-Menüs zu methodischen Problemen führen können (vgl. DILLMAN ET AL. 1998). Lange Itembatterien und komplexe Frageformate (wie z.B. die Präsentation der Antwortmöglichkeiten in umfangreichen Matrizen) wurden soweit wie möglich vermieden, weil sie bei Online-Umfragen die Datenqualität negativ beeinflussen können (vgl. TOEPOEL ET AL. 2009; ELMENOUAR/BLASIUŠ 2005: 75). Eine Plausibilitätsprüfung wurde nicht in den Fragebogen implementiert, da die sinnvolle Umsetzung nur mit JavaScript möglich gewesen wäre und dabei das Risiko bestand, dass die Browser deren Funktionalität aus Sicherheitsgründen verhindern. Die Internetverbindung wurde für alle Teilnehmerinnen und Teilnehmer mit einem SSL-Zertifikat verschlüsselt.

Jede Frage wurde im Fragebogen auf einer separaten Seite angezeigt, um die Anzeige der Filterführung entsprechend anzupassen. Bei Pen-and-Paper-Instrumenten haben die Teilnehmerinnen und Teilnehmer durch die Seitenzahl stets einen Eindruck davon, wie lange es noch dauert, bis der Fragebogen vollständig ausgefüllt ist. In Online-Umfragen wird diese Information häufig über eine Fortschrittsanzeige angeboten, die Unsicherheiten über die verbleibende Befragungsdauer abbauen und so dem Abbruch vorbeugen soll. Die empirischen Befunde zu diesem Effekt sind allerdings ambivalent und bei langen Fragebogen scheint die Verwendung einer Fortschrittsanzeige sogar tendenziell zu einem früheren Abbruch zu führen (vgl. YAN ET AL. 2011: 132ff; vgl. CONRAD ET AL. 2010: 5). Aufgrund der komplexen Filterführung wäre es in StEG zudem problematisch gewesen, den Teilnehmerinnen und Teilnehmern zuverlässige Informationen über die Anzahl der noch ausstehenden Fragen zu geben. Statt einer detaillierten Fortschrittsanzeige mit genauem Umfang der noch zu beantwortenden Fragen, wurde der Fragebogen in thematische Abschnitte unterteilt und der jeweils aktuelle Abschnitt im Kopfbereich jeder Seite hervorgehoben.

Abbildung 1: Typische Seite aus dem StEG-Onlinefragebogen 2012



Sie sind hier: | [Allgemeine Angaben](#) | [Fragen zur Ganztagschule](#) |
 | [Fragen zu Ressourcen](#) | [Fragen zur \(Ganztags-\)Schulentwicklung](#) | [Fragen zur Angebotsstruktur im Ganztagsbetrieb](#)

Info

Die nächsten Fragen beziehen sich auf die Zeitorganisation. Wie ist das an Ihrer Schule?

Bitte jeweils ein Kästchen pro Zeile ankreuzen!

	Nein	Ja
Es gibt einen offenen Anfang vor dem eigentlichen Unterrichtsbeginn	<input type="radio"/>	<input type="radio"/>
Es gibt einen offenen Schulschluss	<input type="radio"/>	<input type="radio"/>
Unterrichtsstunden nach Stundenplan finden über den ganzen Schultag verteilt statt	<input type="radio"/>	<input type="radio"/>
Außerunterrichtliche und Freizeit-Angebote finden über den gesamten Schultag verteilt – also sowohl vormittags als auch nachmittags – statt	<input type="radio"/>	<input type="radio"/>
Schüler/-innen können sich einen Teil ihrer Zeit in der Schule selbst einteilen, z. B. anhand eines Wochenplans	<input type="radio"/>	<input type="radio"/>
In der Tagesstruktur wurden größere Zeitblöcke gebildet	<input type="radio"/>	<input type="radio"/>

Möchte am Ende antworten

Angaben zurücksetzen

Pausieren

Weiter

Die Schaltfläche *Eingaben zurücksetzen* diente dazu, die angeklickten Elemente der jeweils aktuellen Seite wieder zu löschen. Mit *Möchte am Ende antworten* konnte die aktuelle Frage an das Ende des Fragebogens verschoben werden. Bei Fragen, die klärungsbedürftige Begriffe enthielten, wurde eine auffällige Schaltfläche mit dem

Titel *Info* eingeblendet, die auf eine separate Seite mit entsprechenden Erläuterungen verwies. Um den Teilnehmerinnen und Teilnehmern zu erkennen zu geben, dass es möglich ist, den Fragebogen zu unterbrechen und zu einem späteren Zeitpunkt fortzusetzen, wird der Einsatz einer entsprechenden Schaltfläche empfohlen (vgl. BEST/KRUEGER 2004: 68). Es besteht allerdings kein Konsens darüber, ob dadurch der Nonresponse reduziert oder ein frühes Aussteigen aus dem Fragebogen gefördert wird, da es keine Garantie dafür gibt, dass die Befragung zu einem späteren Zeitpunkt wieder aufgenommen wird (vgl. ebd.). Bei der StEG-Befragung war es auch ohne Klick auf Pausieren möglich, den Fragebogen an der letzten Position fortzusetzen (z.B. wenn der Browser einfach geschlossen oder der Computer abgeschaltet wurde).

Während des Ausfüllprozesses wurden die Angaben sukzessive gespeichert, wenn ein Teilnehmer bzw. eine Teilnehmerin mit dem Klick auf die *Weiter*-Schaltfläche die nächste Frage aufgerufen hat. Gefilterte Fragen und Items wurden im Datensatz mit einem separaten Code versehen, bei Abbruch des Fragebogens wurden ab der jeweiligen Position des Ausstiegs keine weiteren Daten gespeichert. Datum und Uhrzeit des ersten Zugriffs sowie des letzten Klicks auf *Weiter* wurden als Zusatzinformation für methodische Analysen gespeichert. Technische Daten wie IP-Adressen der Teilnehmerinnen und Teilnehmer, verwendete Browser, Betriebssysteme usw. wurden aus Gründen des Datenschutzes nicht erhoben.

Auf Wunsch sollte die Antwort auf ganze Fragen oder einzelne Items verweigert werden können. Aus unterschiedlichen Methodenstudien ist allerdings bekannt, dass das Anbieten einer separaten Schaltfläche zur Antwortverweigerung den Nonresponse drastisch erhöhen kann (zusammenfassend siehe DEROUVRAY/COUPER 2002: 3ff). Wird hingegen keine separate Schaltfläche angeboten und einfach das Nicht-Bearbeiten einer Frage als Verweigerung akzeptiert, kann dies zu Missing Values aufgrund von versehentlich ausgelassenen Angaben oder unbeabsichtigten Klicks führen. Um zwischen absichtlich und versehentlich ausgelassenen Antworten unterscheiden zu können, wurde im StEG-Fragebogen eine „implizite Ausweichoption“ (vgl. ebd.: 4) implementiert: Wenn die Befragten trotz unvollständiger Angaben auf die *Weiter*-Schaltfläche geklickt haben, erschien diese Frage nochmals und wurde um einen Hinweis in roter Schrift ergänzt, dass die Angaben noch unvollständig sind. Falls dann erneut auf *Weiter* geklickt wurde, ohne dass die Befragten eine Korrektur vorgenommen haben, wurden die fehlenden Angaben als Verweigerung gespeichert und der Fragebogen fortgesetzt. In Online-Umfragen steigt das Risiko des Fragebogenabbruchs, je nachdrücklicher eine Antwort von den Teilnehmerinnen und Teilnehmern verlangt wird, bevor der Fragebogen fortgesetzt werden kann (vgl. DILLMAN ET AL. 1998: 11). Um keine derartigen Abbrüche zu provozieren, wurde die Vollständigkeitsprüfung deshalb bei Fragen deaktiviert, die besonders viele Texteingaben erforderten.

Im Fragebogen gab es keine Möglichkeit für die Befragten, zu einer bereits beantworteten Frage zurück zu navigieren. Da sowohl Fragen als auch einzelne Items anhand der zurückliegenden Antworten gefiltert wurden, hätte eine Rückwärtsnavigation

mit anschließenden Korrekturen dazu geführt, dass die Abfrage von gänzlich neuen Frage- bzw. Itemkonstellationen nötig gewesen wäre. Bereits beantwortete Fragen hätten z.B. nochmals gestellt werden müssen, um die vormals gefilterten Items nachträglich abzufragen. Es wurde vermutet, dass mit höherem Befragungsaufwand auch die Zahl der Abbrüche gestiegen und das Risiko unplausibler Angaben erhöht worden wäre.

3.2 Feldarbeit

Erhebungen an Schulen müssen von den Kultusministerien der Länder genehmigt werden, wobei die Genehmigungsverfahren in jedem Bundesland unterschiedlich lange dauern und schwer abzuschätzen ist, wann die Freigabe erteilt wird. In StEG wurde deshalb beschlossen, nach Bundesländern gestaffelt ins Feld zu gehen, um jeweils vor den Sommerferien die Erhebungen abschließen zu können. Der Feldstart in den Bundesländern erfolgte mit dem Versand der Zugangsdaten zum Online-Fragebogen. Die Schulen konnten vom 2. Mai 2012 bis zum 17. September 2012 auf den Online-Fragebogen zugreifen, für Berliner *Schulen mit Sekundarstufe I* wurde die Frist wegen des späteren Feldstarts bis zum 8. Oktober verlängert.

Zu wenige Kontaktversuche mit den Zielpersonen gelten als ein Hauptgrund für niedrige Teilnahmequoten (vgl. SCHAEFER/DILLMAN 1998: 380; DILLMAN 2007: 149ff), weshalb für StEG ein dreistufiges Kontaktierungsverfahren gewählt wurde: Im ersten Schritt wurden die Schulen postalisch über die Studie informiert und ihnen die Zugangsdaten zum Online-Fragebogen zugeschickt. Rund zwei Wochen nach diesem ersten Anschreiben wurden die Schulen, die noch nicht teilgenommen hatten, per E-Mail an die Studie erinnert. Etwa drei Wochen nach dem ersten Anschreiben erfolgte für alle Schulen, die noch nicht teilgenommen hatten, eine weitere postalische Erinnerung inklusive der Zugangsdaten zum Online-Fragebogen. Falls eine Schule nach rund vier Wochen noch nicht teilgenommen hatte, wurde ihre Ersatzschule kontaktiert.

Die Termine zum Versand der Anschreiben und Erinnerungen wurden z.T. aus logistischen oder bundeslandspezifischen Gründen verschoben (z.B. wegen der Pfingstferien in Baden-Württemberg und Bayern). Mit dem Kultusministerium Berlin wurde vereinbart, die Teilnahmebereitschaft bei einer Informationsveranstaltung für Schulleitungen aus der Bundeshauptstadt vorab zu ermitteln. Die Informationsveranstaltung wurde am 3. Mai 2012 vom IFS durchgeführt und anschließend wurden nur jene Schulen kontaktiert, die ein Interesse an der Teilnahme bekundet hatten. Da in Berlin parallel zu StEG mehrere Untersuchungen an Schulen mit Sekundarstufe I durchgeführt wurden, startete die Befragung für diese Ganztagschulen nach den Sommerferien. Die Berliner Schulleitungen wurden im Anschreiben ausdrücklich darauf hingewiesen, dass sich die Fragen auf das Schuljahr 2011/2012 beziehen. Niedrige Teilnahmequoten führten bereits während der Feldphase zu dem Schluss, dass für alle Bundesländer auf Ersatzschulen zurückgegriffen werden sollte, um eine Ausschöpfungsquote von mindestens 50 Prozent zu erreichen. Gleichzeitig hatte sich

durch besonders zeitintensive Genehmigungsverfahren das Erhebungsfenster bis zu den Sommerferien stark verkleinert. Anfang Juli 2012 wurden aus diesen Gründen alle noch verfügbaren Ersatzschulen parallel angeschrieben, sofern deren zugehörige gezogene Schule (oder die erste Ersatzschule) noch nicht teilgenommen hatte (vgl. Tabelle 4)

Tabelle 4: Versandtermine der Kontaktierungsschreiben zur Schulleitungsbefragung 2012

	Gezogene Schulen			1. Ersatzschulen			2. Ersatzschulen		
	Erstkontakt	Hinweis E-Mail	Hinweis Post	Erstkontakt	Hinweis E-Mail	Hinweis Post	Erstkontakt	Hinweis E-Mail	Hinweis Post
BB	02.05.	22.05.	31.05.	11.06.	19.06.	02.07.	02.07.	10.07.	17.07.
BE	14.05.	30.05.	11.06.	18.06.	26.06.	10.07.	02.07.	10.07.	17.07.
BW	08.05.	22.05.	11.06.	18.06.	26.06.	10.07.	17.07.	24.07.	30.07.
BY	24.05.	12.06.	19.06.	26.06.	10.07.	17.07.	17.07.	24.07.	30.07.
HB	08.05.	22.05.	31.05.	-	-	-	-	-	-
HE	25.05.	12.06.	19.06.	02.07.	10.07.	17.07.	02.07.	10.07.	17.07.
HH	08.05.	22.05.	31.05.	11.06.	19.06.	02.07.	02.07.	10.07.	17.07.
MV	02.05.	22.05.	31.05.	11.06.	19.06.	03.07.	03.07.	10.07.	17.07.
NI	18.05.	31.05.	11.06.	18.06.	26.06.	10.07.	02.07.	10.07.	17.07.
NW	10.05.	24.05.	01.06.	11.06.	19.06.	02.07.	02.07.	10.07.	17.07.
RP	04.06.	18.06.	02.07.	02.07.	10.07.	17.07.	02.07.	10.07.	17.07.
SH	02.05.	22.05.	31.05.	11.06.	19.06.	02.07.	02.07.	10.07.	17.07.
SL	23.05.	12.06.	19.06.	02.07.	10.07.	17.07.	02.07.	10.07.	17.07.
SN	10.05.	24.05.	31.05.	11.06.	19.06.	02.07.	02.07.	10.07.	17.07.
ST	29.05.	12.06.	19.06.	02.07.	10.07.	17.07.	02.07.	10.07.	17.07.
TH	08.05.	22.05.	31.05.	11.06.	19.06.	02.07.	02.07.	10.07.	17.07.

Für allgemeine Fragen zur Studie wurde am DIPF eine telefonische Hotline eingerichtet, die werktags von 9 Uhr bis 16 Uhr besetzt war. Bei technischen Problemen bestand entsprechend die Möglichkeit, eine technische Hotline am DJI anzurufen. Die Löschung bereits eingegebener Daten konnte von den Schulen ohne Nennung von Gründen über die technische Hotline, per E-Mail oder Post veranlasst werden.

4 Realisierte Stichprobe

4.1 Ausfallgründe

Die meisten Ausfälle kamen im StEG Systemmonitoring durch eine passive Teilnehmerweigerung zustande, indem nicht auf die Kontaktversuche reagiert bzw. nicht auf den Online-Fragebogen zugegriffen wurde. Die Verweigerungsgründe ließen sich

nicht systematisch erheben, weil auch dafür ein aktiver Zugriff auf den Fragebogen notwendig gewesen wäre. Konkrete Ausfallgründe konnten nur für jene Fälle ermittelt werden, die selbstständig einen entsprechenden Ausfallgrund rückgemeldet haben (z.B. per E-Mail oder Telefon) oder wenn sich dieser aus den vorliegenden Daten ableiten ließ (z.B. Eintrag in offenes Textfeld).

Tabelle 5: Ausfallgründe und Umfang der Nettostichprobe von Schulen mit Primarstufe

	Gezogen	1. Ersatz	2. Ersatz	Gesamt
Verfügbare Schuladressen	704	679	676	2.059
- kein Kontaktversuch	5	199	290	494
Schulen mit Kontaktversuch	699	480	386	1.565
- keine Ganztagschule/ Listenfehler	3	2	1	6
- Schule geschlossen	-	-	-	-
- Teilnahmeverweigerung rückgemeldet	3	14	2	19
- kein Kontakt nach drei Versuchen	402	335	302	1.039
- Widerruf der Teilnahme Einwilligung	1	-	-	1
Teilnehmende Schulen	290	129	81	500
- mehrfache Teilnahme von gezogenen Schulen und Ersatzschulen für ein Stichprobenelement	3	10	16	29
Nettostichprobe	287	119	65	471

Tabelle 6: Ausfallgründe und Umfang der Nettostichprobe von Schulen mit Sekundarstufe I

	Gezogen	1. Ersatz	2. Ersatz	Gesamt
Verfügbare Schuladressen	816	766	730	2.312
- kein Kontaktversuch	16	242	336	594
Schulen mit Kontaktversuch	800	524	394	1.718
- keine Ganztagschule/ Listenfehler	5	2	1	8
- Schule geschlossen	5	-	-	5
- Teilnahmeverweigerung rückgemeldet	4	5	1	10
- kein Kontakt nach drei Versuchen	448	349	269	1066
- Widerruf der Teilnahme Einwilligung	1	-	-	1
Teilnehmende Schulen	337	168	123	628
- mehrfache Teilnahme von gezogenen Schulen und Ersatzschulen für ein Stichprobenelement	1	21	32	54
Nettostichprobe	336	147	91	574

Tabelle 7: Ausfallgründe und Umfang der Nettostichprobe von Gymnasien

	Gezogen	1. Ersatz	2. Ersatz	Gesamt
Verfügbare Schuladressen	365	280	216	861
- kein Kontaktversuch	1	93	106	200
Schulen mit Kontaktversuch	364	187	110	661
- keine Ganztagschule/ Listenfehler	4	-	-	4
- Schule geschlossen	1	-	-	1
- Teilnahmeverweigerung rückgemeldet	3	2	-	5
- kein Kontakt nach drei Versuchen	183	114	81	378
- Widerruf der Teilnahmeeinwilligung	-	-	1	1
Teilnehmende Schulen	173	71	28	272
- mehrfache Teilnahme von gezogenen Schulen und Ersatzschulen für ein Stichprobenelement	1	11	13	25
Nettostichprobe	172	60	15	247

Aus Tabelle 5 bis Tabelle 7 ist ersichtlich, dass für einzelne Schulen kein Kontaktversuch unternommen wurde, obwohl diese regulär in die Stichprobe gezogen wurden. Die betreffenden *Schulen mit Primarstufe* und *Schulen mit Sekundarstufe I* stammen aus Berlin, wo die Teilnahmebereitschaft vorab durch eine Informationsveranstaltung erhoben wurde und bei Desinteresse an der Studie kein schriftliches Teilnahmegesuch verschickt wurde. Aus Mecklenburg-Vorpommern hat hingegen ein *Gymnasium* bereits während der Genehmigungsphase auf Nachfrage des dortigen Kultusministeriums eine explizite Absage erteilt. Als teilnehmende Fälle werden in StEG jene Schulen gezählt, die mit der Eingabe von Benutzerkennung und Passwort aktiv auf den Online-Fragebogen zugegriffen und keine nachträgliche Löschung ihrer Daten beantragt haben. Ersatzschulen wurden nur dann kontaktiert, wenn die entsprechende gezogene Schule bzw. eine andere Ersatzschule noch nicht an der Befragung teilgenommen hat.

Bis auf die Verlängerung der Feldphase für Berliner Schulen gab es keine weiteren Zugangsbeschränkungen zum Online-Fragebogen, wodurch es z.B. möglich war, dass sich eine gezogene Schule sehr spät noch an der Umfrage beteiligte, auch wenn bereits deren Ersatzschule teilgenommen hatte. Um Verzerrungen zu vermeiden, wurde die realisierte Stichprobe um die mehrfache Teilnahme unter gezogenen Schulen und deren Ersatzschulen bereinigt. Bei der Auswahl der Schulen für die Nettostichprobe wurde auch die Vollständigkeit der vorhandenen Angaben berücksichtigt: Falls z.B. eine gezogene Schule den Fragebogen abgebrochen aber eine ihrer Ersatzschulen alle Fragen bearbeitet hatte, wurde die Schule mit den meisten Angaben in die Nettostichprobe aufgenommen. In allen anderen Fällen wurden gezogene Schulen bevorzugt und 1. Ersatzschulen vor 2. Ersatzschulen ausgewählt.

4.2 Ausschöpfung

Mit der Ausschöpfungsquote wird in StEG das Verhältnis des realisierten zum angestrebten Stichprobenumfang beschrieben, d.h. das Verhältnis des Nettostichprobenumfangs zur Anzahl der gezogenen Schulen in der Bruttostichprobe. Über alle Schulgruppen ergibt sich demnach eine Ausschöpfungsquote von insgesamt 68,5 Prozent (vgl. Tabelle 8).

Tabelle 8: Ausschöpfungsquoten nach Schulgruppen und Bundesländern

	Primar			Sek. I			Gym.			Gesamt		
	Br.	Net.	Quote	Br.	Net.	Quote	Br.	Net.	Quote	Br.	Net.	Quote
BB	27	18	66,7%	30	24	80,0%	20	10	50,0%	77	52	67,5%
BE	44	25	56,8%	26	11	42,3%	-	-	-	70	36	51,4%
BW	29	23	79,3%	62	41	66,1%	25	19	76,0%	116	83	71,6%
BY	25	21	84,0%	160	109	68,1%	40	29	72,5%	225	159	70,7%
HB	23	18	78,3%	31	16	51,6%	3	3	100%	57	37	64,9%
HE	34	30	88,2%	56	47	83,9%	25	19	76,0%	115	96	83,5%
HH	25	14	56,0%	25	9	36,0%	25	14	56,0%	75	37	49,3%
MV	-	-	-	26	19	73,1%	25	18	72,0%	51	37	72,5%
NI	54	39	72,2%	88	72	81,8%	35	31	88,6%	177	142	80,2%
NW	250	146	58,4%	120	70	58,3%	28	18	64,3%	398	234	58,8%
RP	34	29	85,3%	42	37	88,1%	25	22	88,0%	101	88	87,1%
SH	25	15	60,0%	30	27	90,0%	25	17	68,0%	80	59	73,8%
SL	25	13	52,0%	25	17	68,0%	25	10	40,0%	75	40	53,3%
SN	60	45	75,0%	45	37	82,2%	25	15	60,0%	130	97	74,6%
ST	-	-	-	25	20	80,0%	23	11	47,8%	48	31	64,6%
TH	49	35	71,4%	25	18	72,0%	16	11	68,8%	90	64	71,1%
D	704	471	66,9%	816	574	70,3%	365	247	67,7%	1.885	1.292	68,5%

Wie in Tabelle 8 dargestellt, sind die Ausschöpfungsquoten in den Bundesländern und Schulgruppen relativ heterogen und reichen von 36 Prozent (Hamburger *Schulen mit Sekundarstufe I*) bis zu 100 Prozent (Bremer *Gymnasien*). Dies führt zu einer zusätzlichen Unwucht in der Verteilung von Schulen auf die einzelnen Bundesländer. Für repräsentative Aussagen über Deutschland sollte daher neben dem oben genannten Designgewicht ein Anpassungsgewicht verwendet werden, das die entstandenen Disproportionalitäten kompensiert (vgl. Abschnitt 5.2).

Aufgrund der Verwendung von Ersatzschulen lassen sich die Ausschöpfungsquoten nicht direkt als *Teilnahmebereitschaft* o.ä. interpretieren, da mit insgesamt 3.944 Fällen wesentlich häufiger um Teilnahme an der Studie gebeten wurde, als Schulen in der Bruttostichprobe enthalten sind. Ohne den Einsatz von Ersatzschulen wäre in StEG lediglich eine Ausschöpfungsquote von 42,2 Prozent erreicht worden.

Wie können diese Beteiligungsdaten bewertet werden? Aufgrund der Rückmeldungen einzelner Schulen, die die Teilnahme verweigert haben sowie Informationen aus den Kultusministerien und der Feldarbeit des DPC gibt es Hinweise darauf, dass die Bereitschaft von Schulen allgemein sinkt, sich an wissenschaftlichen Studien zu beteiligen. In Großbritannien ist dieser Trend bereits seit einigen Jahren zu beobachten (vgl. STURGIS ET AL. 2006: 18ff). In der *Hessischen Studie zu ganztägig arbeitenden Schulen* (HeGS) wurde 2011 bei der Online-Schulleitungsbefragung eine Teilnahmequote von knapp 50 Prozent erreicht (vgl. STECHER/HOPF 2011: 3). Das Projekt *Bildungsberichterstattung Ganztagschule NRW* erzielte in seiner Online-Schulleitungsbefragung 2010 eine Rücklaufquote von 20,4 Prozent (vgl. BÖRNER ET AL. 2011: 6f). In beiden Studien wurde eine Vollerhebung von Ganztagschulen des jeweiligen Landes angestrebt, d.h. es war nicht notwendig, mit Ersatzschulen zu arbeiten. Eine Ausschöpfungsquote von insgesamt 68,5 Prozent erscheint vor diesem Hintergrund als eine zufriedenstellende Größe.

Teilnahmequoten werden allerdings häufig als Gütekriterium per se aufgefasst und nicht als Indikator für einen möglichen Nonresponse-Bias (vgl. STURGIS ET AL. 2006: 14). So schreiben die technischen Standards von PISA die Realisierung von bestimmten Beteiligungsquoten des intendierten Schulsamples vor und Erhebungsdaten eines Landes werden nicht veröffentlicht, sofern die Quote insgesamt unter 85 Prozent bzw. ohne Ersatzschulen unter 65 Prozent liegt (vgl. OECD 2012: 7). Mit einer Ausschöpfungsquote von rund 42 Prozent ohne Ersatzschulen hätte das StEG-Systemmonitoring 2012 die in PISA verwendete *cut-off-rate* formal verfehlt. Allerdings würde bei einem solchen Vergleich nicht in Rechnung gestellt, dass die Teilnahme an StEG im Gegensatz zu PISA für die Schulen freiwillig ist. Mindestquoten, wie sie in PISA eingesetzt werden, erfahren auch deshalb kritische Würdigung, weil sie letztlich kein Garant für zuverlässige Schätzer sind: „Although all survey researchers may agree that obtaining a high response rate is generally a ‚good thing‘, the key problem with this kind of criterion based approach is that a survey failing to meet the response rate target may yield less biased estimates than another survey which easily exceeds it“ (STURGIS ET AL. 2006: 16). STURGIS ET AL. empfehlen daher, die Stichprobenqualität auch anhand einzelner Merkmale zu beurteilen, die sowohl für Teilnehmer als auch Ausfälle vorliegen (vgl. ebd.: 16f).

4.3 Systematik der Ausfälle

In diesem Abschnitt soll die Frage erörtert werden, ob Ausfälle im StEG Systemmonitoring zufällig zustande gekommen sind oder ob sich systematische Verzerrungen identifizieren lassen. Im Fokus steht dabei der Vergleich zwischen angestrebter und realisierter Stichprobe, d.h. die Interpretation ist von einer Untersuchung mit Bezug auf alle *kontaktierten* Schulen bzw. alle von Kultusministerien *gelisteten* Schulen zu unterscheiden.⁵

Ausfälle nach Schulgruppen

Das Verhältnis von Brutto- und Nettostichprobenumfang ist für alle drei Schulgruppen relativ ähnlich und reicht von Ausschöpfungsquoten mit rund 67 Prozent bei *Schulen mit Primarstufe*, 68 Prozent bei *Gymnasien* bis zu 70 Prozent bei *Schulen mit Sekundarstufe I*. Ein statistisch bedeutsamer Unterschied in den Beteiligungsraten der Schulgruppen kann allerdings nicht festgestellt werden (vgl. Tabelle 9).

Tabelle 9: Chi-Quadrat-Test der Schulgruppen auf Ausfall und Teilnahme

		Ausfall	Teiln.
Schulen mit Primarstufe	Anzahl	233	471
	in %	33,1	66,9
	Erwartet	222	483
	Adj. Std. Resid.	1,2	-1,2
Schulen mit Sek. I	Anzahl	242	574
	in %	29,7	70,3
	Erwartet	257	559
	Adj. Std. Resid.	-1,5	1,5
Gymnasien	Anzahl	118	247
	in %	32,3	67,7
	Erwartet	115	250
	Adj. Std. Resid.	0,4	-0,4
Gesamt	n	1.885	
	χ^2 (df)	2,23(2)	
	Cramér's V	0,03	

* p < 0,05; ** p < 0,01; *** p < 0,001

⁵Bei einem Vergleich der teilnehmenden Schulen mit allen *kontaktierten* Schulen läge der Schwerpunkt der Interpretation stärker auf der Teilnahmebereitschaft. Unterschiede zwischen *Grundgesamtheit* und realisierter Stichprobe wären hingegen auch durch das Stichprobendesign erklärbar.

Ausfälle nach Bundesländern

Insbesondere die heterogene Beteiligung in den Bundesländern legt die Frage nahe, ob diese Unterschiede lediglich durch zufällige Schwankungen bedingt sind. Um Bundesländer zu identifizieren, in denen die Ausfälle mit hoher Wahrscheinlichkeit auf systematische Ursachen zurückzuführen sind, werden im Folgenden für jede Schulgruppe die bundeslandspezifischen Ausschöpfungsquoten⁶ mit der Beteiligungsrate in den jeweils übrigen Bundesländern verglichen.

Tabelle 10: Zusammenfassung der Mittelwertvergleiche von Schulen mit Primarstufe zur bundeslandspezifischen Beteiligung an StEG

	<i>n</i>	<i>M</i>	<i>SD</i>	95%-KI von <i>M_{diff}</i>			<i>t</i> -Wert
				<i>KI_{min}</i>	<i>M_{diff}</i>	<i>KI_{max}</i>	
BB	27	0,67	0,48	-0,18	-0,02	0,18	-0,27
BE	44	0,57	0,50	-0,26	-0,11	0,05	-1,38
BW	29	0,79	0,41	-0,03	0,13	0,29	1,65
BY	25	0,84	0,37	0,08	0,18	0,34	2,30 *
HB	23	0,78	0,42	-0,07	0,12	0,30	1,31
HE	34	0,88	0,33	0,11	0,22	0,34	3,79 ***
HH	25	0,56	0,51	-0,30	-0,11	0,10	-1,18
MV	-	-	-	-	-	-	-
NI	54	0,72	0,45	-0,07	0,06	0,19	0,90
NW	250	0,58	0,49	-0,21	-0,13	-0,06	-3,49 **
RP	34	0,85	0,36	0,06	0,19	0,32	3,01 **
SH	25	0,60	0,50	-0,26	-0,07	0,12	-0,75
SL	25	0,52	0,51	-0,34	-0,16	0,03	-1,61
SN	60	0,75	0,44	-0,03	0,09	0,21	1,49
ST	-	-	-	-	-	-	-
TH	49	0,71	0,46	-0,09	0,05	0,19	0,70

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Abk.: *M*=Mittelwert (mit 1=teilgenommen, 0=nicht teilgenommen);

SD=Standardabweichung; *KI*=Konfidenzintervall; *M_{diff}*=Mittelwertdifferenz (aus jeweiligem Landesmittelwert – Mittelwert der übrigen Länder);

KI_{min}/*KI_{max}*=Unter/Obergrenze des Konfidenzintervalls

Für *Schulen mit Primarstufe* kann ein signifikanter Zusammenhang zwischen der Bundeslandzugehörigkeit und der Beteiligung an StEG festgestellt werden ($\chi^2(13)=36,33$, $p<0,01$). Die Stärke dieses Zusammenhangs ist insgesamt gering (Cramér's $V=0,23$, $p<0,01$), wobei einige Bundesländer besonders hervorzuheben sind (vgl. Tabelle 10):

⁶Ausgedrückt als Landesmittelwert einer binär codierten Teilnahmevariable.

So konnte nur für Nordrhein-Westfalen festgestellt werden, dass die Ausschöpfung signifikant unterhalb des Durchschnitts der übrigen Bundesländer liegt. Bayern, Hessen und Rheinland-Pfalz weisen hingegen eine überdurchschnittlich hohe Ausschöpfung auf.

Tabelle 11: Zusammenfassung der Mittelwertvergleiche von Schulen mit Sekundarstufe I zur bundeslandspezifischen Beteiligung an StEG

	<i>n</i>	<i>M</i>	<i>SD</i>	95%- <i>KI</i> von <i>M_{diff}</i>			<i>t</i> -Wert
				<i>KI_{min}</i>	<i>M_{diff}</i>	<i>KI_{max}</i>	
BB	30	0,80	0,41	-0,06	0,10	0,26	1,32
BE	26	0,42	0,50	-0,50	-0,29	-0,08	-2,89 **
BW	62	0,66	0,48	-0,16	-0,05	0,07	-0,76
BY	160	0,68	0,47	-0,11	-0,03	0,05	-0,68
HB	31	0,52	0,51	-0,38	-0,20	-0,01	-2,10 *
HE	56	0,84	0,37	0,04	0,15	0,25	2,79 **
HH	25	0,36	0,49	-0,54	-0,35	-0,17	-3,85 ***
MV	26	0,73	0,45	-0,15	0,03	0,21	0,31
NI	88	0,82	0,39	0,04	0,13	0,22	2,87 **
NW	120	0,58	0,50	-0,24	-0,14	-0,05	-2,92 **
RP	42	0,88	0,33	0,08	0,19	0,29	3,52 **
SH	30	0,90	0,31	0,09	0,20	0,32	3,51 **
SL	25	0,68	0,48	-0,21	-0,02	0,16	-0,26
SN	45	0,82	0,39	0,01	0,13	0,25	2,10 *
ST	25	0,80	0,41	-0,07	0,10	0,27	1,20
TH	25	0,72	0,46	-0,17	0,02	0,20	0,18

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Abk.: *M*=Mittelwert (mit 1=teilgenommen, 0=nicht teilgenommen);

SD=Standardabweichung; *KI*=Konfidenzintervall; *M_{diff}*=Mittelwertdifferenz

(aus jeweiligem Landesmittelwert – Mittelwert der übrigen Länder);

KI_{min}/*KI_{max}*=Unter/Obergrenze des Konfidenzintervalls

Auch für *Schulen mit Sekundarstufe I* ist der Zusammenhang zwischen den Bundesländern und der Teilnahme an StEG auf hohem Niveau statistisch bedeutsam ($\chi^2(15)=66,45$; $p<0,001$). Die Beziehung zwischen den beiden Merkmalen ist stärker als bei *Schulen mit Primarstufe* (Cramér's $V=0,29$, $p<0,001$). Auffällig ist bei Betrachtung der einzelnen Landesquoten (vgl. Tabelle 11), dass neben Nordrhein-Westfalen alle Stadtstaaten eine signifikant niedrigere Ausschöpfung vorweisen als die jeweils verbleibenden Bundesländer. Schulen aus Hessen, Niedersachsen, Rheinland-Pfalz, Schleswig-Holstein und Sachsen zeichnen sich durch überdurchschnittlich hohe Ausschöpfungsquoten aus.

Tabelle 12: Zusammenfassung der Mittelwertvergleiche von Gymnasien zur bundeslandspezifischen Beteiligung an StEG

	<i>n</i>	<i>M</i>	<i>SD</i>	95%- <i>KI</i> von <i>M_{diff}</i>			<i>t</i> -Wert
				<i>KI_{min}</i>	<i>M_{diff}</i>	<i>KI_{max}</i>	
BB	20	0,50	0,51	-0,40	-0,19	0,02	-1,74
BE	-	-	-	-	-	-	-
BW	25	0,76	0,44	-0,10	0,09	0,28	0,92
BY	40	0,73	0,45	-0,10	0,05	0,21	0,69
HB ¹	3	1,00	0,00	-	-	-	-
HE	25	0,76	0,44	-0,10	0,09	-0,10	0,98
HH	25	0,56	0,51	-0,32	-0,13	0,07	-1,29
MV	25	0,72	0,46	-0,15	0,05	0,24	0,49
NI	35	0,89	0,32	0,11	0,23	0,35	3,82 ***
NW	28	0,64	0,49	-0,29	-0,04	0,16	-0,40
RP	25	0,88	0,33	0,07	0,22	0,36	3,07 **
SH	25	0,68	0,48	-0,19	0,00	0,20	0,04
SL	25	0,40	0,50	-0,49	-0,30	-0,11	-3,10 **
SN	25	0,60	0,50	-0,27	-0,08	0,11	-0,85
ST	23	0,48	0,51	-0,41	-0,21	-0,01	-2,11 *
TH	16	0,69	0,48	-0,23	0,01	0,25	0,09

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Abk.: *M*=Mittelwert (mit 1=teilgenommen, 0=nicht teilgenommen);

SD=Standardabweichung; *KI*=Konfidenzintervall; *M_{diff}*=Mittelwertdifferenz

(aus jeweiligem Landesmittelwert – Mittelwert der übrigen Länder);

KI_{min}/*KI_{max}*=Unter/Obergrenze des Konfidenzintervalls

¹ Konfidenzintervalle und *t*-Werte sind für Bremer Gymnasien aufgrund der niedrigen Fallzahl nicht ausgewiesen

Für Gymnasien lässt sich ebenfalls ein Unterschied zwischen angestrebter und realisierter Stichprobe feststellen, der auf die Bundeslandzugehörigkeit der Schulen zurückgeführt werden kann ($\chi^2(14)=33,50$; $p<0,01$). Der Zusammenhang weist mittlere Stärke auf (Cramér's $V=0,30$; $p<0,01$). Im Mittelwertvergleich verfügen Gymnasien aus Niedersachsen und Rheinland-Pfalz über signifikant höhere, Schulen aus Sachsen-Anhalt und dem Saarland hingegen über niedrigere Ausschöpfungsquoten (vgl. Tabelle 12).

Organisationsform

Für die implizite Stratifizierung der Stichprobe wurden fünf Kategorien⁷ mit Organisationsformen des Ganztagsbetriebs gebildet, die sich für die Ausfallanalyse nutzen lassen, da sowohl teilnehmende als auch ausgefallene Schulen in diesem Merkmal über eine Ausprägung verfügen (vgl. Abschnitt 2.2). Dabei sollte allerdings berücksichtigt werden, dass die Zuordnung zu den einzelnen Kategorien mit dem Ziel erfolgte, Schulen mit ähnlicher Betreuungsform innerhalb des gleichen Bundeslandes zu bündeln. Die hier dargestellten Organisationsformen stimmen (trotz der identischen Bezeichnungen) inhaltlich nicht immer mit den Definitionen der KMK überein und können daher in verschiedenen Bundesländern für unterschiedliche Konzepte stehen. Insofern ist das Differenzierungsmerkmal der Organisationsform als Proxy-Variable zu betrachten, die lediglich Hinweise auf eine systematische Verzerrung liefern kann.

Tabelle 13: Schulgruppenspezifische Chi-Quadrat-Tests von Organisationsform und Ausfällen

	Primar		Sek. I		Gym.		
	Ausfall	Teiln.	Ausfall	Teiln.	Ausfall	Teiln.	
offen	Anzahl	190	358	102	307	80	166
	in %	34,7	65,3	24,9	75,1	32,5	67,5
	Erwartet	181	367	121	288	80	167
	Adj. Std. Resid.	1,7	-1,7	-3,0	3,0	0,1	-0,1
teilgebunden	Anzahl	17	40	54	102	7	19
	in %	29,8	70,2	34,6	65,4	26,9	73,1
	Erwartet	19	38	46	110	8	18
	Adj. Std. Resid.	-0,5	0,5	1,5	-1,5	-0,6	0,6
gebunden	Anzahl	7	34	73	132	26	49
	in %	17,1	82,9	35,6	64,4	34,7	65,3
	Erwartet	14	27	61	144	24	51
	Adj. Std. Resid.	-2,2	2,2	2,2	-2,2	0,5	-0,5
sonstige	Anzahl	19	39	13	33	5	13
	in %	32,8	66,9	28,3	71,7	27,8	72,2
	Erwartet	19	39	14	32	6	12
	Adj. Std. Resid.	-0,1	0,1	-0,2	0,2	-0,4	0,4
Gesamt	n	704		816		365	
	χ^2 (df)	5,65(3)		9,73(3)*		0,71(3)	
	Cramér's V	0,09		0,11*		0,04	

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

⁷Die Kategorien „Schule mit verschiedenen Betreuungsformen“ und „keine Angabe“ wurden im Folgenden zusammengefasst zu „sonstige“.

Für *Schulen mit Primarstufe* und *Gymnasien* lässt sich keine statistisch bedeutsame Systematik der Ausfälle in Hinblick auf die Organisationsform nachweisen (vgl. Tabelle 13). Teilnehmende und ausgefallene *Schulen mit Sekundarstufe I* unterscheiden sich jedoch nach diesem Merkmal in ihrer Zusammensetzung, wobei die Beziehung nur schwach ausgeprägt ist. Bei näherer Betrachtung haben Schulen mit offener Ganztagsbetreuung häufiger teilgenommen als erwartet. Schulen der Kategorie „gebunden“ tragen hingegen mit einer höheren Ausfallhäufigkeit zum beobachteten Zusammenhang bei.

Siedlungsstruktureller Kreistyp

Für alle in StEG verwendeten Ganztagschulen liegen die amtlichen Gemeindecodenziffern vor, die eine Verknüpfung von Erhebungsdaten und Daten der amtlichen Statistik ermöglichen. Diese Informationen können wiederum für Ausfallanalysen verwendet werden. Tabelle 14 zeigt einen Chi-Quadrat-Test zwischen den teilnehmenden und nicht-teilnehmenden Schulen hinsichtlich des siedlungsstrukturellen Landkreistyps⁸, in dem die Schulen verortet sind.

Für *Schulen mit Primarstufe* und *Schulen mit Sekundarstufe I* existiert ein schwacher, aber signifikanter Zusammenhang zwischen dem regionalen Kontext der Schulen und der Häufigkeit von Ausfällen. Als besonders bedeutsam zur Erklärung dieses Zusammenhangs im Sekundarbereich erweisen sich *Schulen mit Sekundarstufe I* aus kreisfreien Großstädten, die den Fragebogen auffallend häufig nicht ausgefüllt haben. Dieses Ergebnis ist angesichts der geringeren Beteiligung in Stadtstaaten plausibel (vgl. Abschnitt 4.3), allerdings bleibt der Zusammenhang auch ohne Schulen aus Berlin, Bremen oder Hamburg bestehen ($n=734$; $\chi^2(3)=13,29$; Cramér's $V=0,14$; $p<0,01$). Schulen aus ländlichen Kreisen mit Verdichtungsansätzen haben sowohl unter den *Schulen mit Primarstufe* als auch den *Schulen mit Sekundarstufe I* häufiger teilgenommen als erwartet.

⁸Weitere Informationen zu siedlungsstrukturellen Kreistypen und den Raumabgrenzungen des Bundesinstituts für Bau-, Stadt- und Raumforschung unter www.bbsr.bund.de

Tabelle 14: Schulgruppenspezifische Chi-Quadrat-Tests von siedlungsstrukturellen Kreistypen und Ausfällen

		Primar		Sek. I		Gym.	
		Ausfall	Teiln.	Ausfall	Teiln.	Ausfall	Teiln.
Kreisfreie	Anzahl	87	144	91	112	43	64
Großstädte	in %	37,7	62,3	44,8	55,2	40,2	59,8
	Erwartet	76	155	60	143	35	72
	Adj. Std. Resid.	1,8	-1,8	5,5	-5,5	2,1	-2,1
Städtische	Anzahl	96	181	85	211	36	78
Kreise	in %	34,7	65,3	28,7	71,3	31,6	68,4
	Erwartet	92	185	88	208	37	77
	Adj. Std. Resid.	0,7	-0,7	-0,4	0,4	-0,2	0,2
Ländliche	Anzahl	22	80	30	124	19	45
Kreise mit	in %	21,6	78,4	19,5	80,5	29,7	70,3
Verdichtungs-	Erwartet	34	68	46	108	21	43
ansätzen	Adj. Std. Resid.	-2,7	2,7	-3,1	3,1	-0,5	0,5
Dünn	Anzahl	28	66	36	127	20	60
besiedelte	in %	29,8	70,2	22,1	77,9	25,0	75,0
ländliche Kreise	Erwartet	31	63	48	115	26	54
	Adj. Std. Resid.	-0,7	0,7	-2,4	2,4	-1,6	1,6
Gesamt	n	704		816		365	
	χ^2 (df)	9,07(3)*		34,64(3)***		5,22(3)	
	Cramér's V	0,11*		0,21***		0,12	

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Eine Systematik der Ausfälle lässt sich in Bezug auf Organisationsform und siedlungsstrukturellen Kreistypen nur für einzelne Schulgruppen nachweisen. Die Bundeslandzugehörigkeit hingegen erweist sich sowohl bei *Schulen mit Primarstufe*, *Schulen mit Sekundarstufe I* als auch *Gymnasien* als relevantes Merkmal. Bestimmte Bundesländer sind in der (ungewichteten) Nettostichprobe systematisch seltener vertreten bzw. überrepräsentiert. Freilich sollte die Bundeslandzugehörigkeit zur Erklärung des Zusammenhangs als Proxy-Variable interpretiert werden, denn mit einer geographischen Verortung lässt sich das höhere bzw. niedrigere Ausfallrisiko nur unzureichend begründen. Viel ausschlaggebender als der Standort dürften die unterschiedlichen Rahmenbedingungen sein, unter denen die Erhebung in den einzelnen Bundesländern stattgefunden hat: die Qualität der von den zuständigen Behörden bereitgestellten Schullisten, Ferienzeiten, Verfügbarkeit von Ersatzschulen für die Ziehung, unterschiedliche Phasen hoher Arbeitsbelastung z.B. durch parallel stattfindende Evaluationen oder Abschlussprüfungen sowie administrative Vorgaben für die Durchführung wissenschaftlicher Untersuchungen. Die Analyseergebnisse zum

siedlungsstrukturellen Kreistyp legen allerdings die Vermutung nahe, dass auch der nähere regionale Kontext einen bedeutenden Einfluss auf die Beteiligung an wissenschaftlichen Befragungen haben kann. Eine mögliche, aufgrund unzureichender Daten aber hypothetische Erklärung wäre, dass Schulen aus urbanen Regionen häufiger um die Teilnahme an wissenschaftlichen Studien gebeten werden und deren Bereitschaft inzwischen niedriger ausfällt, einen Fragebogen auszufüllen.

Letztlich lassen sich die genannten Rahmenbedingungen während des Erhebungsprozesses in einer auf Freiwilligkeit basierenden, bundesweit durchgeführten Schulbefragung kaum kontrollieren. Die Verteilung der Schulen auf die Bundesländer wird in StEG in der Regel mit Gewichten ausgeglichen, da sie in erster Linie durch das Design bedingt ist (vgl. Abschnitt 5).

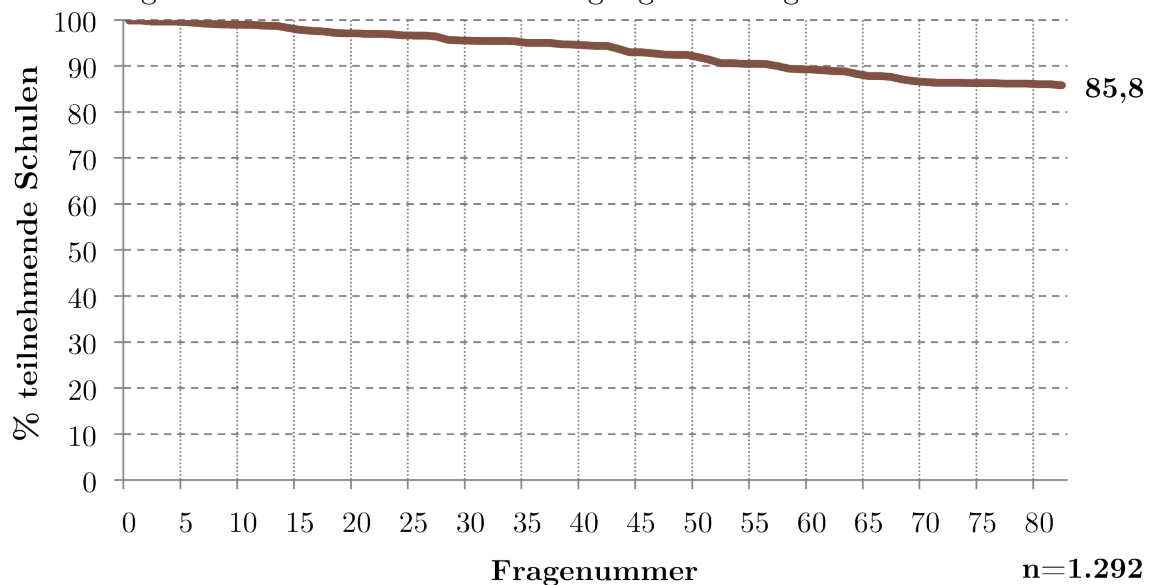
4.4 Bearbeitungsdauer und Fragebogenabbruch

Neben der Totalverweigerung ist bei Online-Untersuchungen der vorzeitige Abbruch der Befragung ein spezielles Problem. Im Gegensatz zu face-to-face oder telefonischen Befragungen gibt es dort in der Regel keine Interviewerinnen und Interviewer, die eine Zielperson zur Vollendung des Fragebogens motivieren können. Häufig sind technische Gründe (z.B. lange Ladezeiten) oder die verwendeten Frageformate (z.B. umfangreiche Matrizen) Ursache für einen vorzeitigen Abbruch (vgl. ELMENOUAR/BLASIUS 2005: 72ff). Zwar wurde bei der Konstruktion des Online-Fragebogens darauf geachtet, typische Abbruchursachen möglichst zu vermeiden, jedoch können sie als Ausfallgründe nicht völlig ausgeschlossen werden, da Fragebogeneigenschaften intersubjektiv unterschiedlich bewertet und toleriert werden (vgl. Abschnitt 3.1).

In Abbildung 2 ist der Anteil der Schulen dargestellt, die sich in den Online-Fragebogen eingeloggt und ihn noch sukzessive bearbeitet haben. Von den insgesamt 1.292 Schulen der Nettostichprobe haben rund 86 Prozent ($n=1.109$) den Fragebogen vollständig⁹ ausgefüllt. Mit zunehmender Länge des Fragebogens steigt auch der Anteil der Schulen kontinuierlich an, die die Befragung vorzeitig beendet haben. Zwar ist die Abbruchquote bei bestimmten Fragen geringfügig höher, wenn diese besonders aufwändig zu bearbeiten waren (z.B. viele Texteingaben oder längere Itembatterien), drastische Einbrüche sind allerdings nicht zu verzeichnen. Dies spricht dafür, dass die Abbrüche in erster Linie nicht durch einzelne Fragen provoziert wurden, sondern v.a. die Bearbeitungsdauer ausschlaggebend war.

⁹Die Vollständigkeit bezieht sich hier auf die Fragen 1 bis 84, da es sich bei Frage 85 um eine unverbindliche offene Abschlussfrage zur Erhebung handelte („Gibt es von Ihrer Seite noch Hinweise oder Anregungen, die Sie uns zu dem Fragebogen oder einzelnen Fragen geben möchten?“). Item-Nonresponse wurde bei der Beurteilung der Vollständigkeit nicht berücksichtigt, sondern lediglich, ob der Fragebogen (auch nach Verweigerung einzelner Angaben) fortgesetzt wurde.

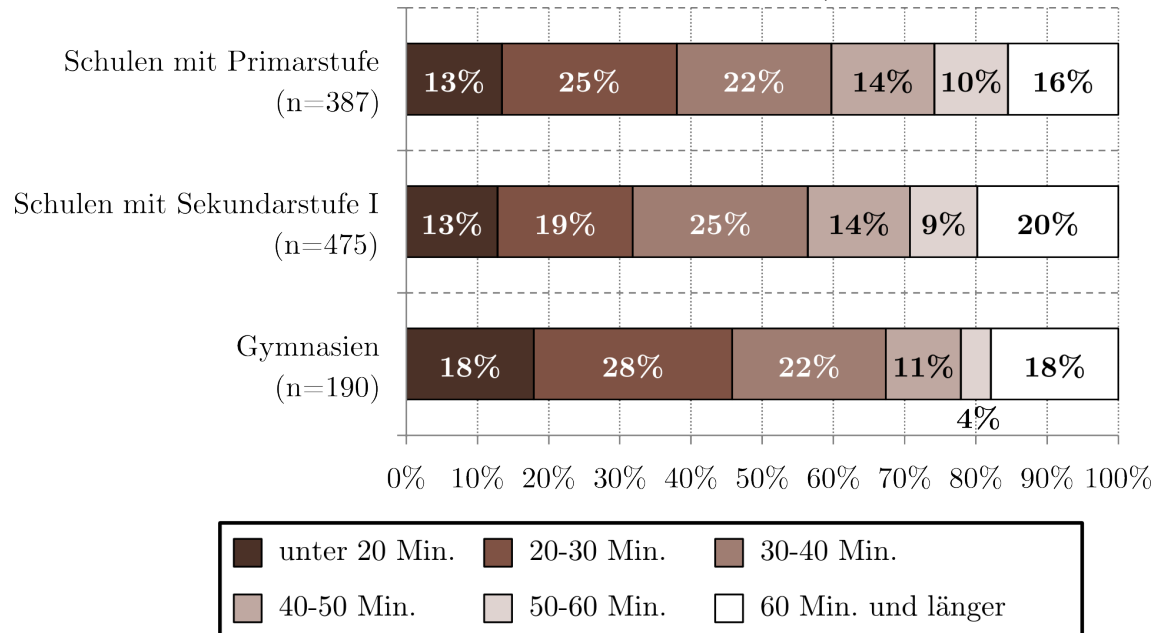
Abbildung 2: Verbleib in der Online-Befragung nach Fragenummer



Die Bearbeitungsdauer stellt insbesondere dann eine Ursache für den Fragebogenabbruch dar, wenn die von den Teilnehmerinnen und Teilnehmern subjektiv wahrgenommene Befragungszeit von der angekündigten Länge des Fragebogens abweicht (vgl. YAN ET AL. 2011). Das Frageprogramm beinhaltete eine komplexe Filterführung und die angezeigten Fragen und Items hingen stark von der Schulgruppe sowie den jeweils gegebenen Antworten ab. Im Anschreiben wurden den StEG-Schulen daher eine geschätzte Bearbeitungsdauer von 40 bis 50 Minuten angekündigt. Von 1.292 Schulen der Nettostichprobe haben 235 Fälle den Fragebogen an unterschiedlichen Tagen begonnen und beendet; für sie kann die tatsächliche Bearbeitungszeit leider nicht abgeleitet werden. Für den überwiegenden Teil der übrigen Schulen, die einen vollständigen Online-Fragebogen ausgefüllt haben, hat sich die Schätzung allerdings als zutreffend erwiesen (vgl. Abbildung 3).

Etwa drei Viertel der Schulen haben den Online-Fragebogen in weniger als 50 Minuten ausgefüllt. An 16 bis 20 Prozent der Schulen wurden für das Ausfüllen mindestens 60 Minuten benötigt. Darunter fallen auch Fälle mit extrem langer Bearbeitungszeit, die den Fragebogen zwar am gleichen Tag begonnen und beendet, dazwischen aber längere Pausen gehalten haben. Trotz verschiedener Filterführung unterscheidet sich die mittlere Bearbeitungszeit der Schulgruppen nicht signifikant voneinander (ohne Abbildung): *Schulen mit Sekundarstufe I* haben zum Ausfüllen mit durchschnittlich rund 49 Minuten am meisten Zeit benötigt ($M=49,4$; $SD=50,4$; $n=475$; $p=0,151$), hatten aber auch das umfangreichste Instrument (vgl. Abschnitt 3.1). *Schulen mit Primarstufe* konnten den Fragebogen im Schnitt nach etwa 46 Minuten abschließen ($M=46,4$; $SD=51,3$; $n=387$; $p=0,754$), *Gymnasien* beendeten die Befragung bereits nach zirka 42 Minuten mit vollständigen Angaben ($M=42,3$; $SD=39,6$; $n=190$; $p=0,143$).

Abbildung 3: Befragungsdauer vollständig bearbeiteter Fragebogen (mit Befragungsbeginn und -ende am gleichen Tag)



5 Gewichtung

5.1 Designgewicht

Wie in Abschnitt 2.3 dargestellt, ist die Stichprobe des StEG-Systemmonitorings nicht streng proportional nach Bundesländern stratifiziert. Daher verteilen sich die Schulen jeder Schulgruppe in der Stichprobe anders auf die Bundesländer als in der Grundgesamtheit. Bei Analysen, die sich auf ganz Deutschland beziehen sollen, werden die Größenverhältnisse daher häufig anhand von *Designgewichten* rekonstruiert.¹⁰ In StEG werden die Designgewichte über die inverse Auswahlwahrscheinlichkeit der Schulen berechnet, indem Kennwerte der Bruttostichprobe und des Stichprobenrahmens herangezogen werden.

$$u_i = \frac{1}{\pi_i}$$

mit

$$\pi_i = \frac{m_h}{N_h}$$

Das Designgewicht wird hier als u bezeichnet, die Auswahlwahrscheinlichkeit üblicherweise mit π notiert. m repräsentiert die Anzahl der aus Stratum h in die Bruttostichprobe gezogenen Schulen und N den Umfang der Grundgesamtheit zur korrespondierenden Schicht. Innerhalb der Strata h sind die Auswahlwahrscheinlichkeiten für jede Schule i gleich groß und somit auch deren Designgewichte identisch. Es sollte allerdings beachtet werden, dass ein Teil der Schulen bei der Stichprobenziehung mehreren Schichten hätte zugeordnet werden können und über einen zusätzlichen Auswahlschritt verhindert wurde, dass diese Schulen doppelt befragt werden. Es handelt sich dabei um Schulen, die sowohl über eine Primarstufe als auch eine Sekundarstufe I verfügen (vgl. Abschnitt 2.2) und für die differenzierte Auswahlwahrscheinlichkeiten vorliegen, je nachdem, ob sie in die Schulgruppe der *Schulen mit Primarstufe* oder *Schulen mit Sekundarstufe I* gezogen wurden (vgl. Tabelle 15).

¹⁰Das StEG-Team dankt Dr. Sabine Meinck (DPC) für die Bereitstellung der Gewichte und die freundliche Beratung.

Tabelle 15: Auswahlwahrscheinlichkeiten für Schulen im Systemmonitoring

	Primar			Primar und Sek. I				Sek. I			Gym.			
	N	m	π	Gezogen zu Prim.		Gezogen zu Sek. I		N	m	π	N	m	π	
				N	m	π	m							π
BB	180	25	0,14	14	2	0,14	8	0,57	97	22	0,23	33	20	0,61
BE	372	42	0,11	14	2	0,14	3	0,21	104	23	0,22	-	-	-
BW	135	15	0,11	123	14	0,11	22	0,18	206	40	0,19	99	25	0,25
BY	218	25	0,11	-	-	-	-	-	974	160	0,16	218	40	0,18
HB	23	23	1,00	-	-	-	-	-	31	31	1,00	3	3	1,00
HE	266	30	0,11	37	4	0,11	8	0,22	275	48	0,17	99	25	0,25
HH	73	25	0,34	-	-	-	-	-	35	25	0,71	59	25	0,42
MV	-	-	-	-	-	-	-	-	137	26	0,19	36	25	0,69
NI	521	50	0,10	45	4	0,09	8	0,18	504	80	0,16	161	35	0,22
NW	2.668	250	0,09	-	-	-	-	-	653	120	0,18	127	28	0,22
RP	277	31	0,11	30	3	0,10	4	0,13	205	38	0,19	66	25	0,38
SH	135	25	0,19	-	-	-	-	-	186	30	0,16	64	25	0,39
SL	151	25	0,17	-	-	-	-	-	62	25	0,40	28	25	0,89
SN	597	60	0,10	-	-	-	-	-	266	45	0,17	115	25	0,22
ST	-	-	-	-	-	-	-	-	68	25	0,37	23	23	1,00
TH	431	49	0,11	-	-	-	-	-	106	25	0,24	16	16	1,00
D	6.047	675		263	29		53		3.909	763		1.147	365	

Sofern eine Gewichtung zu separaten Inferenzpopulationen *Schule mit Primarstufe* bzw. *Schulen mit Sekundarstufe I* vorgesehen ist, können diese Auswahlwahrscheinlichkeiten direkt für die Berechnung verwendet werden. Bei Gewichtung auf eine Inferenzpopulation, in der *Schulen mit Primarstufe* und *Schulen mit Sekundarstufe* zusammengefasst sind, sollten die so berechneten Gewichte vor einer gemeinsamen Analyse der Schulen dahingehend korrigiert werden, dass jede Schule mit Primar- und Sekundarstufe nur einmal in die Stichprobe gezogen werden konnte. Für Schichten, in denen die Bruttostichprobe eine gerade Anzahl dieser Schulen enthält, lässt sich diese Korrektur durch eine einfache Multiplikation des errechneten Gewichts mit dem Faktor 0,5 erreichen. Für Schichten mit einer ungeraden Zahl an gezogenen Schulen (Berlin und Rheinland-Pfalz) berechnet sich der Faktor hingegen aus:

$$adj(x) = \frac{m_{yh}}{(m_{xh} + m_{yh})}$$

$$adj(y) = \frac{m_{xh}}{(m_{xh} + m_{yh})}$$

$adj(x)$ ist der Adjustierungsfaktor für Schulen die in die Gruppe *Schulen mit Primarstufe* gezogen werden, $adj(y)$ der Faktor für Schulen, die in die Gruppe *Schulen mit Sekundarstufe I* gezogen werden. m_{xh} bzw. m_{yh} bezeichnet die Anzahl der in die jeweilige Gruppe gezogenen Schulen innerhalb der Schicht h der Bruttostichprobe. Anhand der Fälle, die für Rheinland-Pfalz aus der Schicht der Schulen mit Primar- und Sekundarstufe in die Schulgruppe *Schulen mit Sekundarstufe I* gezogen wurde, ergäbe sich daraus beispielsweise folgendes Designgewicht:

$$u_i = \frac{1}{\pi_i} \cdot adj(y)$$

Im Datensatz sind die Designgewichte für diese beiden Varianten der Gewichtung auf separate Inferenzpopulationen bzw. auf eine gemeinsame Inferenzpopulation bereits enthalten.

5.2 Ausfallgewicht und Gesamtgewicht

Eine weitere Abweichung von der Verteilung der Schulen auf die Bundesländer ist während der Erhebung durch Ausfälle entstanden, die je nach Bundesland unterschiedlich hoch waren (vgl. Abschnitt 4). Um den daraus entstandenen Mangel an Schulen bundeslandspezifisch zu kompensieren, werden in StEG *Ausfallgewichte* eingesetzt, die sich aus dem Verhältnis von Brutto- und Nettostichprobe ergeben:

$$v_{hi} = \frac{m_h}{n_h}$$

Dabei bezeichnet m den Umfang der angestrebten Stichprobe im Stratum h und n die Anzahl der teilnehmenden Schulen in der Nettostichprobe der korrespondierenden Schicht.¹¹ Für komplexe Stichproben wie beim StEG Systemmonitoring sollten Ausfallgewichte allerdings nicht ohne die zugehörigen Designgewichte verwendet werden (vgl. GABLER/GANNINGER 2010: 144). Ein Gesamtgewicht lässt sich für jede Schule durch einfache Multiplikation von Design- und Ausfallgewicht bilden:

$$w_i = u_i \cdot v_i$$

Da im StEG Systemmonitoring je nach Inferenzpopulation unterschiedliche Designgewichte verwendet werden, unterscheiden sich auch die Gesamtgewichte in Hinblick auf die Reichweite der möglichen Aussagen.

¹¹Wie die Designgewichte werden auch die Ausfallgewichte für Schulen mit Primar- und Sekundarstufe in StEG separat berechnet, je nachdem, in welche Schulgruppe die Schulen gezogen wurden (vgl. Abschnitt 5.1).

5.3 Normierung des Gesamtgewichts

Manche Statistikpakete wie z.B. SPSS Base unterstützen bei der Gewichtung lediglich das Verfahren des *frequency weighting*.¹² Jeder Fall wird dabei als Repräsentant einer dem Gewicht entsprechenden Anzahl an Fällen in die Analyse einbezogen. Als Stichprobenumfang wird in manchen Statistikpaketen allerdings nicht die Anzahl an Fällen im Datensatz angenommen, sondern die Summe der Gewichte. Ohne eine Normierung der Gewichte entspricht deren Summe nicht dem tatsächlichen Stichprobenumfang, sondern der Fallzahl aller Schulen in der Inferenzpopulation, d.h. bei einer Gewichtung für *Gymnasien* läge den Berechnungen nicht $n=365$ zugrunde, sondern $N=1.147$. „Die Folge ist, dass die Stichprobenfehler unterschätzt werden, d.h. sie fallen in den Berechnungen tendenziell zu klein aus. Damit geraten die Konfidenzintervalle bei komplexen Zufallsstichproben ebenfalls tendenziell zu klein.“ (DIAZ-BONE 2006: 157). Zur Normierung wird jedes schulspezifische Gesamtgewicht w lediglich durch den Mittelwert aller Gesamtgewichte in der Stichprobe geteilt:

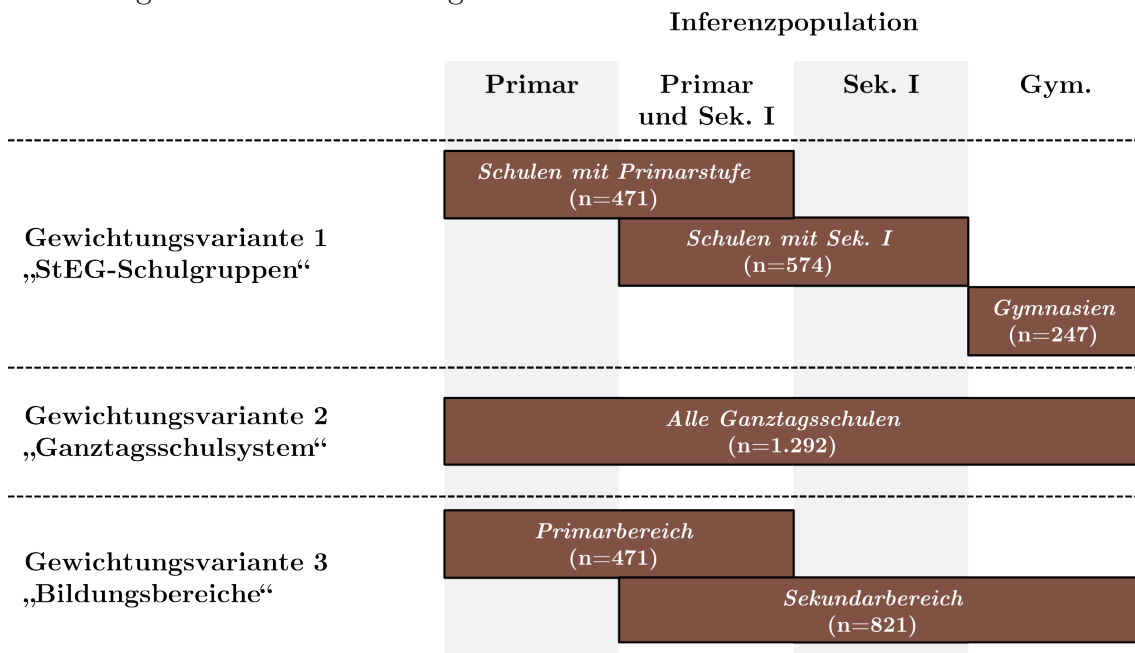
$$\tilde{w}_i = \frac{w_i}{\frac{1}{n} \cdot \sum w_i}$$

$\sum w_i$ entspricht der Anzahl aller Schulen in der Inferenzpopulation, d.h. es sind unterschiedliche Varianten an normierten Gewichten denkbar, je nachdem, auf welche Grundgesamtheit sie sich beziehen sollen (vgl. Abbildung 4).

Variante 1 in Abbildung 4 stellt eine Gewichtung passend für die drei StEG-Schulgruppen *Schulen mit Primarstufe*, *Schulen mit Sekundarstufe I* und *Gymnasien* dar. Damit lassen sich z.B. gewichtete Verteilungen für die einzelnen Schulgruppen ausgeben, die auf den jeweiligen Stichprobenumfang genormt sind. Gewichtungsvariante 2 wird hingegen empfohlen, wenn Auswertungen quer über alle Schulgruppen durchgeführt werden sollen. Die Gewichtung rekonstruiert dabei die Bundeslandanteile in Hinblick auf alle Ganztagschulen in Deutschland und ist normiert auf einen Stichprobenumfang mit 1.292 Fällen. Variante 3 eignet sich besonders dafür, die *Schulen mit Primarstufe* mit allen anderen Schulen des Sekundarbereichs vergleichend zu analysieren. Die Normierung des Gewichts basiert auf den zusammengelegten Fällen von *Schulen mit Sekundarstufe I* und *Gymnasien*.

¹²Das SPSS-Modul *Complex Samples* erlaubt die Verwendung weiterer Gewichtungsverfahren.

Abbildung 4: Normierte Gesamtgewichte im Datensatz



6 Scientific-Use-File

Das Scientific-Use-File stellt eine aufbereitete Version des in StEG verwendeten Analysedatensatzes dar, der für die wissenschaftliche Sekundärnutzung vorgesehen ist. Der Datensatz unterscheidet sich nur geringfügig von der Version, die im Projekt verwendet wurde: Die meisten Änderungen wurden aus datenschutzrechtlichen Gründen, zur Steigerung der Kompatibilität mit Statistik-Software und zur Verbesserung der Handhabung vorgenommen.

Im Scientific-Use-File zum Systemmonitoring sind sowohl die Angaben von allen in StEG befragten Institutionen enthalten als auch generierte Variablen, die im Projekt häufig verwendet wurden (z. B. Stichprobenvariablen und Rekodierungen). Der Datensatz wurde mit der Statistiksoftware SPSS erstellt und ist im *wide format* strukturiert: Für jede StEG-Schule ist in der Datenmatrix genau eine Zeile vorgesehen. Alle Fallattribute sind in Spalten mit eindeutigen Variablenbezeichnungen gespeichert.

6.1 Datenaufbereitung

Für das StEG Systemmonitoring wurden die Erhebungsdaten an zentraler Stelle für die am Projekt beteiligten Forschungsinstitute aufbereitet. Die Datenaufbereitung beinhaltete folgende Schritte:

1. Datenmigration von MySQL nach SPSS

Mit dem Online-Fragebogen wurden alle Angaben in einer MySQL-Datenbank gespeichert. Nach der Erhebung wurden die Daten in einem ersten Verarbeitungsschritt in das von SPSS verwendete Format übertragen.

2. Berichtigung von Fragebogenabbrüchen

Aufgrund der Prüfregeln im Online-Programm wurde bei Abbruch des Fragebogens bei manchen Fragen automatisiert ein Missing-Code (-9) für Items eingetragen, die von den Teilnehmerinnen und Teilnehmern überhaupt nicht aufgerufen worden waren. Fragebogenabbrüche wurden dahingehend vereinheitlicht, dass ab der zuletzt aufgerufenen Frage alle weiteren Daten gelöscht wurden.

3. Integration von Online- und Print-Daten

Insgesamt wurden bei der Schulleitungsbefragung sechs gedruckte Fragebogen ausgefüllt. Diese Angaben wurden manuell in einen SPSS-Datensatz übertragen, den Daten aus der Online-Befragung angefügt und in einer neuen Variable als Print-Daten gekennzeichnet.

4. Umsetzung von Korrekturaufträgen

In drei Fällen wurde über die Hotline von den Schulleitungen um eine Korrektur der eingegebenen bzw. eine Ergänzung noch unvollständiger Daten gebeten. Die gewünschten Änderungen wurden manuell eingearbeitet.

5. Integration von Stamm- und Erhebungsdaten

In der MySQL-Datenbank wurden keine Stammdaten sondern lediglich Daten aus der Befragung gespeichert. Weitere für die Analyse notwendigen Informationen zu Stichprobe und Feldphase wurden den Erhebungsdaten in diesem Schritt daher nachträglich hinzugefügt.

6. Kontrolle der Teilnahme pro Stichprobenelement

Ein Stichprobenelement besteht im StEG-Systemmonitoring entweder aus einer gezogenen Schule, einer 1. Ersatzschule oder einer 2. Ersatzschule. Da keine Zugangsbeschränkungen für den Online-Fragebogen existierten, war es möglich, dass für ein Stichprobenelement über die gesamte Feldzeit bis zu drei Schulen teilgenommen haben. Daher wurde der Teilnahmestatus von gezogenen Schulen und ihren zugehörigen Ersatzschulen ausgewertet und für einen späteren Zuschnitt des Datensatzes gespeichert (siehe auch Aufbereitungsschritt 14).

7. Neuzuweisung von Fallbezeichnern (Schul-ID)

Zur Verbesserung des Datenschutzes wurden für alle Schulen neue Fallbezeichner generiert, die eine zufällige Komponente beinhalten.

8. Anreicherung mit Metadaten

SPSS verfügt über weitreichende Möglichkeiten Metadaten wie Variablenlabels, Wertelabels, Skalenniveaus usw. für jede Variable zu definieren. Zur Verbesserung der Benutzerfreundlichkeit wurde der Datensatz mit diesen Zusatzinformationen angereichert. Zudem wurden dem Datensatz StEG-spezifische Metadaten als benutzerdefinierte Variablenattribute hinzugefügt.

9. Ergänzen von Regionaldaten

Von den Kultusministerien der Länder wurden dem Projekt für alle Schulen des Stichprobenrahmens die zugehörigen Gemeindekennziffern mitgeteilt. Den Erhebungsdaten konnten auf diese Weise Regionaldaten aus der amtlichen Statistik zugespielt werden, um z.B. Schulen aus ländlichen und städtischen Räumen zu unterscheiden.

10. Plausibilitätsprüfung und -korrektur

Die inhaltliche Plausibilität der Angaben wurde überprüft und nach Rücksprache im Forschungsverbund korrigiert bzw. unplausible Daten kenntlich gemacht. Dabei wurden v. a. Ausreißer und fehlerhafte Angaben (z.B. Tippfehler) als Missing definiert und nicht vorhandene Informationen ergänzt, die aus offenen oder anderen Fragen abgeleitet werden konnten.

11. Filterprüfung und -korrektur

Der Online-Fragebogen sah das Überspringen von Fragen und Items anhand von vorab programmierten Regeln und den bereits getätigten Angaben vor. Die Programmroutinen konnten allerdings nicht greifen, sofern von den Teilnehmerinnen und Teilnehmern eine Frage an das Ende des Fragebogens verschoben wurde, die für das Auslösen eines Filters notwendig war. Alle Angaben wurden daher nochmals mit den Filterregeln abgeglichen und ggf. eine Plausibilitätskorrektur vorgenommen.

12. Genese zusätzlicher Analysevariablen

Aus den Erhebungsdaten wurden neue Variablen abgeleitet, die einheitlich in den Analysen der Projektstandorte verwendet werden sollten. Dazu zählen beispielsweise Klassifizierungen, Umrechnungen oder Kodierungen von offenen Angaben.

13. Anfügen von Gewichten

Die Designgewichte für die Inferenzpopulationen nach Schulgruppen wurden vom DPC bereitgestellt und den Erhebungsdaten in diesem Schritt zugespielt. Im Anschluss wurden weitere Gewichte berechnet, z.B. Ausfallgewichte und normierte Gesamtgewichte (vgl. Abschnitt 5).

14. Zuschnitt des Analysedatensatzes

Jedes Stichprobenelement sollte im Analysedatensatz des Systemmonitorings nur durch eine Schule vertreten werden, d.h. entweder durch eine gezogene Schule, die 1. Ersatzschule oder 2. Ersatzschule (siehe auch Datenaufbereitungsschritt 6). Redundante Fälle wurden deshalb aus dem Analysedatensatz entfernt und jeweils die am frühesten kontaktierte Schule im Datensatz belassen. Sofern eine Ersatzschule über vollständigere Daten verfügte als alle anderen zuvor kontaktierten Schulen, wurde dieser Fall übernommen und die unvollständigen Daten aus dem Analysedatensatz gelöscht.

6.2 Konventionen und Bedienungshilfen

Fallbezeichner

Die Fallbezeichner (*Schul-ID*) der Scientific-Use-Files stellen keine beliebige Ziffernfolge dar, sondern wurden systematisch aus codierten Stichprobeninformationen generiert. Die Bedeutung der einzelnen ID-Komponenten wird hier an einem fiktiven Beispiel dargestellt:

Beispiel	Erläuterung
<u>1</u> 12150	Kennziffer StEG Systemmonitoring 1 = StEG Systemmonitoring 2012
1 <u>1</u> 2150	Ersatzschulstatus, mit 0 = gezogene Schule 1 = 1. Ersatzschule 2 = 2. Ersatzschule
11 <u>2</u> 150	Schulgruppe, mit 1 = Schulen mit Primarstufe 2 = Schulen mit Sekundarstufe I 3 = Gymnasien
112 <u>150</u>	Laufende Nummer (innerhalb von Ersatzschulstatus und Schulgruppe)

Variablenamen

Die Daten des StEG-Systemmonitorings sind für Trendanalysen vorgesehen, d.h. für eine Anreicherung mit Datenbeständen aus anderen Erhebungen. Variablenamen ermöglichen daher Rückschlüsse auf die Datenquelle bis hin zur genauen Position eines Items im Fragebogen. Die einzelnen Codes wurden anhand des vollständigen Fragebogens zugewiesen, d.h. bei der Nummerierung von Fragen und Items werden individuell gefilterte Elemente ebenfalls mitgezählt.

Beispiel	Erläuterung
<u>dc80d_3</u>	Studienkennung oder Datenquelle, z.B. dc = StEG Systemmonitoring 2012 rn = Regionaldaten
dc <u>80d_3</u>	Nummer der zugehörigen Frage im Fragebogen. Sofern es sich nicht um ein Fragebogendatum handelt, sondern z.B. um eine Stichprobeninformation, wird als Platzhalter meist ein doppelter Unterstrich verwendet.
dc80 <u>d_3</u>	Anhand dieses Codes lässt sich die Listenposition von Items in Aufzählungen und Itembatterien identifizieren, z.B. a = 1. Zeile der Itembatterie b = 2. Zeile der Itembatterie c = 3. Zeile der Itembatterie
dc80d_ <u>3</u>	Ein einfacher Unterstrich wird in StEG als Trennzeichen vor einem zusätzlichen Variablensuffix verwendet.
dc80d_ <u>3</u>	Bei Variablen, die ihren Ursprung im Fragebogen haben, weist eine Ziffer im Variablensuffix auf die horizontale Position in einer Itemmatrix hin z. B. 1 = 1. Spalte der Itemmatrix 2 = 2. Spalte der Itemmatrix 3 = 3. Spalte der Itemmatrix Für abgeleitete Variablen und Daten aus anderen Quellen werden weitere Suffixe verwendet, z. B. r1 = Rekodierungsvariante 1 r2 = Rekodierungsvariante 2

Diese Konventionen wurden nach Möglichkeit auf alle Variablen angewandt, um die Bedienung zu erleichtern. Einzelne Variablen können aber dennoch einer abweichenden Systematik folgen.

Variablenlabels

Je nach Typ sind die Variablen außerdem mit einer Kennzeichnung im Label versehen, um die Verwendung der Datensätze zu erleichtern. Variablen ohne eine der folgenden Kennzeichnungen stammen hingegen direkt aus der Befragung:

- [**\$**] = System- oder Stichprobenvariable
- [**A**] = Abgeleitete Variable
- [**H**] = Hilfsvariable
- [**O**] = Regionaldaten
- [**R**] = Variable aus Rekodierung
- [**S**] = Variable aus Skalenberechnung
- [**W**] = Gewicht

Fehlende Werte

Im Datensatz werden sowohl *user defined missing values* als auch *system missing values* verwendet: Während bestimmte Codes bzw. Wertebereiche bei der Dateneingabe und -aufbereitung als fehlende Werte definiert wurden und in Auswertungen unberücksichtigt bleiben, stellen *system missing values* fehlende Daten im eigentlichen Sinne dar (Nullwerte). *System missing values* treten im Datensatz v. a. bei Items auf, die von den Teilnehmerinnen und Teilnehmern nach vorzeitigem Abbruch der Befragung nicht mehr aufgerufen werden konnten. *User defined missing values* sind als negative Werte codiert, wobei folgende standardisierte Definitionen verwendet wurden:

- 9/-99/-999 = nicht angezeigt (z.B. aufgrund von Filterführung)
- 8/-98/-998 = nicht bearbeitet (z.B. aufgrund von Verweigerung)
- 7/-97/-997 = nicht valide (z.B. aufgrund von Eingabefehlern)
- 6/-96/-996 = kein Item der Batterie bearbeitet

Literaturverzeichnis

- ASSMANN, C./ STEINHAUER, H.W./ KIESL, H./ KOCH, S./ SCHÖNBERGER, B./ MÜLLER-KULLER, A./ ROHWER, G./ RÄSSLER, S./ BLOSSFELD, H.P. (2011): Sampling designs of the National Educational Panel Study: challenges and solutions. Sonderheft der Zeitschrift für Erziehungswissenschaft, Bd. 14, Nr. 2, S. 51–65.
- BEST, S.J./ KRUEGER, B.S. (2004): Internet data collection, Sage university papers. Quantitative applications in the social sciences, Bd. 141. Thousand Oaks: Sage Publications.
- BÖRNER, N./ EBERITZSCH, S./ GROTHUES, R./ WILK, A. (2011): Bildungsbericht Ganztagschule NRW. Dortmund: Eigenverlag Forschungsverbund DJI/TU.
- CONRAD, FREDERICK G./ COUPER, MICK P./ TOURANGEAU, ROGER/ PEYTCHEV, ANDY (2010): The impact of progress indicators on task completion. Interacting with Computers, Bd. 22, Nr. 5, S. 417–427.
- DEROUVRAY, C./ COUPER, M. P. (2002): Designing a Strategy for Reducing „No Opinion“ Responses in Web-Based Surveys. Social Science Computer Review, Bd. 20, Nr. 1, S. 3–9.
- DIAZ-BONE, R. (2006): Statistik für Soziologen. Konstanz: UKV-Verlagsgesellschaft.
- DILLMAN, D.A. (2007): Mail and internet surveys: The tailored design method. Hoboken/New York: John Wiley & Sons.
- DILLMAN, D.A./ TORTORA, R.D./ BOWKER, D. (1998): Principles for constructing web surveys. Download von: <http://survey.sesrc.wsu.edu/dillman/papers/1998/principlesforconstructingwebsurveys.pdf> (Zugriff am 6.11.2012).
- EL-MENOUAR, Y./ BLASIUS, J. (2005): Abbrüche bei Online-Befragungen: Ergebnisse einer Befragung von Medizinerinnen. ZA-Information, Bd. 56, S. 70–92.
- GABLER, S./ GANNINGER, M. (2010): Gewichtung. In: WOLF, C./ BEST, H. (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse, Wiesbaden: VS-Verlag, S. 143–164.

- OECD (2012): PISA 2009 Technical Report. OECD Publishing. Download von: <http://www.oecd.org/pisa/pisaproducts/pisa2009/50036771.pdf>, Download am 7.11.2012.
- SCHAEFER, D.R./ DILLMAN, D.A. (1998): Development of a Standard E-Mail Methodology: Results of an Experiment. *Public Opinion Quarterly*, Bd. 62, Nr. 3, S. 378–397.
- STECHER, L./ HOPF, ANDREA (2011): Hessische Studie zu ganztägig arbeitenden Schulen (HeGS) - Ergebnisse der Schulleitungsbefragung. Download von: <http://www.uni-giessen.de/fbz/fb03/institute/ifezw/prof/empi/Forschung/Projekte/hegs/Monitoring> (Zugriff am 27.9.2016).
- STURGIS, P./ SMITH, P./ HUGHES, G./ BRITAIN, G. (2006): A study of suitable methods for raising response rates in school surveys. DfES Publications. Download von: <https://www.education.gov.uk/publications/eOrderingDownload/RR721.pdf> (Zugriff am 7.11.2012).
- TOEPOEL, V./ DAS, M./ VAN SOEST, A. (2009): Design of Web Questionnaires: The Effects of the Number of Items per Screen. *Field Methods*, Bd. 21, Nr. 2, S. 200–213.
- YAN, T./ CONRAD, F.G./ TOURANGEAU, R./ COUPER, M.P. (2011): Should I Stay or Should I go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys. *International Journal of Public Opinion Research*, Bd. 23, Nr. 2, S. 131–147.

Die *Studie zur Entwicklung von Ganztagschulen* (StEG) ist ein bundeslandübergreifendes Forschungsprogramm zur empirischen Analyse schulischer Veränderungsprozesse und individueller Entwicklungen von Schülerinnen und Schülern an Ganztagschulen. Die Studie wird von Mitarbeiterinnen und Mitarbeitern aus mehreren Forschungseinrichtungen in umfassender Zusammenarbeit durchgeführt. Den Arbeitsverbund bilden das Deutsche Institut für Internationale Pädagogische Forschung (DIPF), das Institut für Schulentwicklungsforschung (IFS), das Deutsche Jugendinstitut (DJI) sowie die Justus-Liebig-Universität Gießen (JLU). StEG wird durch das Bundesministerium für Bildung und Forschung gefördert.