
Studienbrief Teil 3

Bildungsmonitoring

Vorstellung der Ergebnisse aus der ersten
Arbeitsphase



Arbeitsauftrag



- ❖ Dieser Studienbrief behandelt das Thema „Bildungsmonitoring auf der Systemebene“
- ❖ Lehrkräfte sollen in die Lage versetzt werden, die Ergebnisse von Large-Scale-Assessments wie TIMSS, PISA und IGLU zu verstehen und angemessen zu interpretieren.

(aus der Projektskizze)

Gliederung des Studienbriefes



Leitkonzept	Inhaltsbezogene Kompetenzen
0. „Einleitung“	Begriffsdefinitionen
1. „Zielsetzung“	Verständnis über die Ziele und Relevanz von Verfahren des Bildungssystemmonitorings
2. „Bewertungskriterien“	Verständnis über die Bedeutung der Kompetenzmessung und individueller/sozialer/kriterialer Bezugsnormen
3. „Testkonstruktion“	Kenntnisse über die Inhalte und die Erhebungsverfahren der internationalen Schulleistungsstudien PISA, TIMSS, IGLU
4. „Anwendungsbereich“	Kenntnisse der wesentlichen Ergebnisse und Verständnis für deren Reichweite und Aussagekraft
5. „Praktische Implikation“	Kenntnis der Leitentscheidungen der Steuerungsebene und ihre Auswirkungen für die Schulpraxis

Einleitung:

Zusammenfassung der Lehr-/Lernziele



- ❖ Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ Bildungsmonitoring und die verschiedenen Ausprägungen zu kennen
 - ❖ Bildungsmonitoring von Evaluation abgrenzen zu können

Einleitung:

Zusammenfassung der Lehr-/Lernziele



- ❖ Bildungsmonitoring und die verschiedenen Ausprägungen zu kennen
- ❖ Übersichtliche Begriffsbestimmungen und Definitionen

Definition: Bildungsmonitoring¶

„Bildungsmonitoring meint dabei die kontinuierliche und systematische, auf wissenschaftliche Methoden gestützte Beobachtung der Bedingungen, Verläufe, Ergebnisse und Wirkungen von Bildungsprozessen in und außerhalb von Institutionen mit dem Ziel, bildungspolitischen Akteuren wissenschaftlich aufbereitetes, empirisches Wissen zur Information für bildungspolitische Entscheidungsprozesse zur Verfügung zu stellen.“ (Wolter, 2009)¶

„Bildungsmonitoring ist die systematische und auf Dauer angelegte Beschaffung und Aufbereitung von Informationen über ein Bildungssystem und dessen Umfeld.“ (Maritzen, 2008)¶

Im Jahr 2006 wurde durch die KMK eine Gesamtstrategie zum Bildungsmonitoring beschlossen. Bildungsprozesse in Deutschland sollen durch vier miteinander verbundene Bereiche beobachtet und weiterentwickelt werden:¶

- regelmäßige Teilnahme an internationalen Schulleistungsuntersuchungen¶
- zentrale Überprüfung der Bildungsstandards im Ländervergleich¶
- Vergleichsarbeiten¶
- gemeinsame Bildungsberichterstattung¶

Einleitung:

Zusammenfassung der Lehr-/Lernziele



- ❖ Bildungsmonitoring von Evaluation abgrenzen zu können
- ❖ Tabellarische Darstellung von Unterschieden

⌘	Bildungsmonitoring⌘	Evaluation⌘
Grundlage/Ziel⌘	Informationen über das Bildungssystem⌘	Erfolg bestimmter Maßnahmen/Institutionen⌘
Untersuchungsgegenstand⌘	Unterschiedliche Bereiche des Bildungssystems⌘	Bestimmte Maßnahme⌘
Untersuchungsart⌘	Umfassend⌘	Eng und detailliert⌘
Ergebnis⌘	Informationen⌘	Konkrete Entscheidung⌘

Tabelle 1: Abgrenzung der Begriffe „Bildungsmonitoring“ und „Evaluation“[¶]

¶

Durch die Unterschiede wird deutlich, dass eine Studie nicht gleichzeitig beiden „Anliegen“ gerecht werden kann (Klieme u.a., 2007). Allerdings ist es möglich, im Rahmen des „normalen“ Bildungsmonitorings ergänzende Befragungen mit evaluativem Charakter durchzuführen. Beispielsweise wurden in einigen Studien ergänzende Videoanalysen erstellt, um das Lehrverhalten zu betrachten, andere Studien betrachteten den Zusammenhang zwischen den Ergebnissen und dem sozialen Hintergrund der Schülerinnen und Schüler. Zudem wurden in den meisten Studien Zusatzstudien auf nationaler Ebene durchgeführt, sodass hierdurch untersucht werden konnte, ob beispielsweise Bildungsstandards erreicht werden (Klieme u.a., 2007).[¶]

Zielsetzung:

Zusammenfassung der Lehr-/Lernziele



- ❖ Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ die Charakteristika von Systemmonitoring-Studien zu kennen
 - ❖ zu wissen, warum diese Studien durchgeführt werden
 - ❖ die Studien TIMSS, PISA, IGLU und DESI mit den spezifischen Inhalten zu kennen (optional)

Zielsetzung:

Zusammenfassung der Lehr-/Lernziele



- ❖ die Charakteristika von Systemmonitoring-Studien zu kennen
- ❖ Übersichtliche Darstellung der Charakteristika

Viele Aspekte haben alle Studien des Systemmonitorings gemeinsam. Diese sollen hier zunächst erläutert werden. Im Anschluss daran besteht die Möglichkeit, sich über die einzelnen Studien näher zu informieren.¶

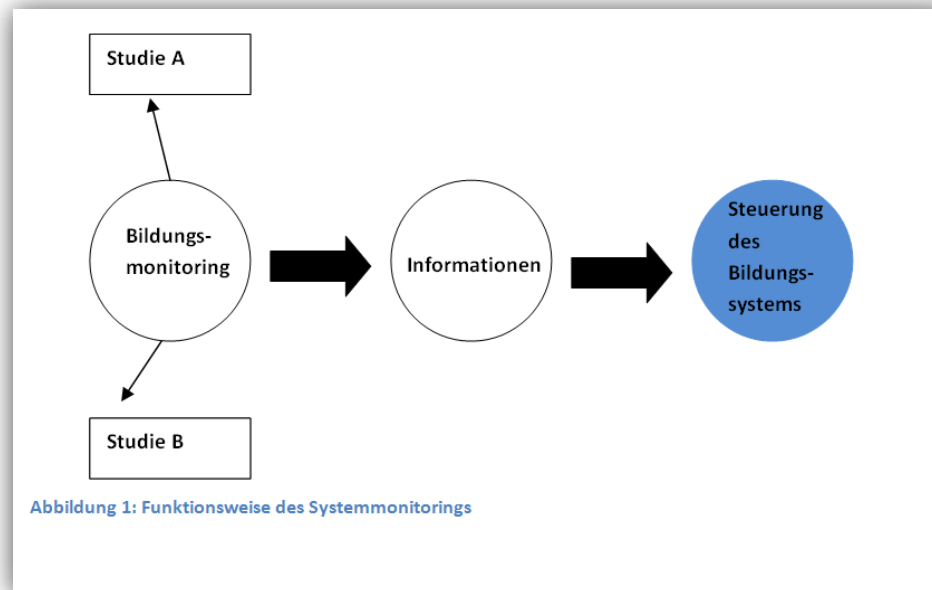
- → Sie testen international eine bestimmte Kompetenz von Schülerinnen und Schülern einer festgelegten Altersgruppe¶
- → Untersuchung findet zeitgleich und mit denselben Aufgaben in allen teilnehmenden Staaten statt¶
- → Federführung liegt bei einer allgemein anerkannten Institution, die die Durchführung und Auswertung der Daten plant und begleitet¶
- → in einigen Teilnehmerstaaten werden dazu noch zusätzliche Kompetenzen innerhalb der Studien getestet um einen intranationalen Vergleich zu ermöglichen¶
- → Federführung dieser Zusatz- oder Ergänzungsstudien liegt im jeweiligen Land selbst¶

Zielsetzung:

Zusammenfassung der Lehr-/Lernziele



- ❖ zu wissen, warum diese Studien durchgeführt werden
- ❖ Verlauf des Bildungsmonitorings übersichtlich mit Hilfe von Grafiken dargestellt



Bezugsnormen und Bewertungskriterien: Zusammenfassung der Lehr-/Lernziele



- ❖ (Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ die Möglichkeiten und Grenzen der Bewertungskriterien kriterial, sozial und individuell einschätzen zu können
 - ❖ den Begriff ‚literacy‘ einordnen zu können
 - ❖ den Unterschied zwischen einer ‚literacy‘-Orientierung und einer curricularen-Orientierung zu kennen

Bezugsnormen und Bewertungskriterien: Zusammenfassung der Lehr-/Lernziele



- ❖ (Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ die Verfahren und die Funktionen von Kompetenzmodellen und Kompetenzniveaus zu kennen
 - ❖ den Zusammenhang zwischen Standards und Kompetenzmodellen/ Kompetenzniveaus zu kennen

Bezugsnormen und Bewertungskriterien: Zusammenfassung der Lehr-/Lernziele



- ❖ die Möglichkeiten und Grenzen der Bewertungskriterien kriterial, sozial und individuell einschätzen zu können
- ❖ Darstellung der sozialen Bezugsnorm am Beispiel IGLU

Vergleichsgruppe 1 ⁸	Vergleichsgruppe 2 ⁹		Vergleichsgruppe 3 ¹⁰
Deutschland	Bulgarien	Niederlande	Argentinien
England	Deutschland	Norwegen	Belize
Frankreich	England	Rumänien	Iran
Griechenland	Frankreich	Schottland	Kolumbien
Italien	Griechenland	Schweden	Kuweit
Niederlande	Island	Slowakei	Marokko
Schottland	Italien	Slowenien	Mazedonien
Schweden	Kanada	Tschechien	Türkei
	Lettland	Ungarn	
	Litauen	USA	
	Neuseeland	Zypern	

Tabelle 7: Ländervergleichsgruppen bei IGLU

Bezugsnormen und Bewertungskriterien: Zusammenfassung der Lehr-/Lernziele



- ❖ den Begriff ‚literacy‘ einordnen zu können
- ❖ Verdeutlichung des Begriffs ‚literacy‘ anhand vom Beispiel PISA

Definition: ‚scientific literacy‘ in PISA ¶

„Naturwissenschaftliches Wissen anzuwenden, naturwissenschaftliche Fragen zu erkennen und aus Belegen Schlussfolgerungen zu ziehen, um Entscheidungen zu verstehen und zu treffen, die die natürliche Welt und die durch menschliches Handeln an ihr vorgenommenen Veränderungen betreffen.“ (PISA-Konsortium, 2001) ¶

Definition: ‚reading literacy‘ in PISA ¶

„(...) die Fähigkeit, geschriebene Texte unterschiedlicher Art in ihren Aussagen, ihren Absichten und ihrer formalen Struktur zu verstehen und sie in einen größeren sinnstiftenden Zusammenhang einzuordnen, sowie in der Lage zu sein, Texte für verschiedene Zwecke sachgerecht zu nutzen.“

Bezugsnormen und Bewertungskriterien: Zusammenfassung der Lehr-/Lernziele



- ❖ den Unterschied zwischen einer ‚literacy‘-Orientierung und einer curricularen-Orientierung zu kennen
- ❖ Unterschied anhand des Beispiel TIMSS dargestellt

Beispiel: curriculare und literacy-Orientierung in TIMSS (1995)

Curriculare Orientierung:

TIMSS II (Sekundarstufe I) und die Untersuchung von voruniversitären Mathematik- und Physikkenntnissen (TIMSS III) lehnten sich eng an die jeweiligen Lehrpläne an. Ziel war es hier, „zentrale Inhalte und Anforderungen des Fachunterrichts“ zu untersuchen. Dies ließ sich vor allem dadurch gut ermöglichen, dass die Curricula international relativ übereinstimmend waren.

literacy-Orientierung:

Die mathematische und naturwissenschaftliche Grundbildung in TIMSS III wurde nach dem Konzept von literacy untersucht. Unter Grundbildung wurde hier „die Kenntnis zentraler Konzepte und Arbeitsprinzipien dieser Fächer“ aber auch „die Fähigkeit, dieses Wissen in alltäglichen Zusammenhängen zu nutzen und zu kommunizieren“ verstanden. Hier wird besonders deutlich, dass literacy „eine Verknüpfung von fachsystematischem Verständnis und Anwendungsorientierung“ beschreibt.

Bezugsnormen und Bewertungskriterien: Zusammenfassung der Lehr-/Lernziele



- ❖ die Verfahren und die Funktionen von Kompetenzmodellen und Kompetenzniveaus zu kennen
- ❖ PISA-Kompetenzniveaus als Beispiel verwendet

Beispiel: Kompetenzniveaus in PISA: Mathematik

Kompetenzniveau I: Rechnen auf Grundschulniveau

Personen, die dieses Niveau zugeordnet werden, verfügen lediglich über arithmetisches und geometrisches Wissen auf Grundschulniveau. Sie können dieses Wissen abrufen und unmittelbar anwenden, wenn die Aufgabenstellung von vornherein eine bestimmte Standard-Mathematisierung nahe legt. Begriffliche Modellierungen sind nicht leistbar.

Kompetenzniveau II: Elementare Modellierungen

Auf diesem Niveau werden auch einfachste begriffliche Modellierungen vorgenommen, die in einem außermathematischen Kontext eingebettet sind. Personen auf diesem Kompetenzniveau können unter mehreren möglichen Lösungsansätzen den passenden finden, wenn durch Graphiken, Tabellen, Zeichnungen usw. eine Struktur vorgegeben ist, die das Modellieren erleichtert. Auch auf diesem Niveau sind allerdings nur die Wissensinhalte der Grundschulmathematik sicher verfügbar.

Testkonstruktion:

Zusammenfassung der Lehr-/Lernziele



- ❖ (Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ die Anforderungen an die Testaufgaben zu kennen
 - ❖ den Ablauf der Testdurchführung nachvollziehen zu können
 - ❖ die Aufgabenauswahl und die Verortung in Kompetenzmodellen nachvollziehen zu können
 - ❖ das Verfahren der Stichprobenziehung zu verstehen

Testkonstruktion:

Zusammenfassung der Lehr-/Lernziele



- ❖ die Anforderungen an die Testaufgaben zu kennen
- ❖ Testgütekriterien Validität, Objektivität und Reliabilität beschrieben

Besonders für die internationalen Studien des Systemmonitorings ist es bedeutsam, dass Untersuchungsgegenstände ausgewählt werden, die Auskunft über die Leistungsfähigkeit der jeweiligen Bildungssysteme geben. Außerdem müssen Voraussetzungen, Durchführung und Auswertung der Tests für alle Teilnehmerstaaten gleich sein. Die Betrachtung der Validität soll sicherstellen, dass genau das gemessen wird, was auch gemessen werden soll (Heller/Hany, 2001). Dazu werden die Studien im Rahmen des Systemmonitorings sorgfältig konzipiert und die Aufgaben gewissenhaft ausgesucht. Zudem durchlaufen die Testaufgaben eine Erprobungsphase.

Validität

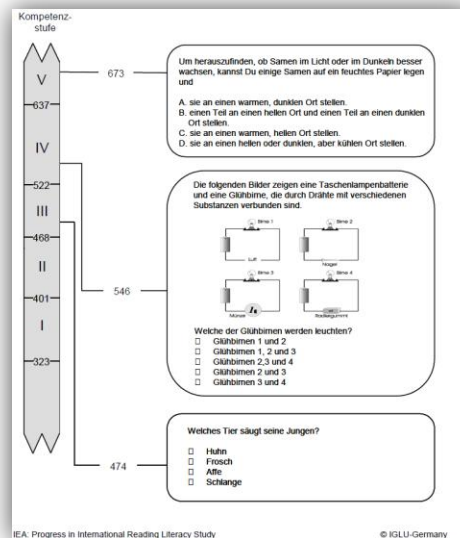
Um dem Kriterium der Objektivität gerecht zu werden, erhalten bei Studien des Bildungsmonitorings alle Testleiter von den Organisationen und Instituten, die die Studie konzipiert haben, genaue Anweisungen zur Durchführung der Studie. Darin ist genau beschrieben, wie die Aufgaben verteilt werden müssen, welche Anweisungen gegeben werden dürfen und welche Hilfestellungen und Hilfsmittel erlaubt sind. Die Einhaltung der Anweisungen wird von unangemeldeten Beobachtern überwacht.

Objektivität

Testkonstruktion: Zusammenfassung der Lehr-/Lernziele



- ❖ die Aufgabenauswahl und die Verortung in Kompetenzmodellen nachvollziehen zu können
- ❖ Verortung der Aufgaben auf den Kompetenzniveaus am Beispiel PISA verdeutlicht



Testkonstruktion: Zusammenfassung der Lehr-/Lernziele



- ❖ das Verfahren der Stichprobenziehung zu verstehen
- ❖ Umgang mit Besonderheiten bei der Stichprobenziehung beschrieben

Ein weiteres Problem entsteht, wenn die zu unterscheidenden Schulsysteme unterschiedliche Quoten der Sonderschulen aufweisen und diese nicht mit in die Untersuchung einbezogen werden (Arnold, 2001). Aus diesem Grund werden in den internationalen Studien Quoten festgelegt, nach denen bestimmte Schülergruppen von der Untersuchung ausgeschlossen werden dürfen. Bei IGLU betrug diese Quote beispielsweise 5 Prozent der Population, ansonsten konnte das Land nicht in der internationalen Berichterstattung analysiert werden (Bos u.a., 2003). Ausgeschlossen werden können auch Schülerinnen und Schüler, die die Testsprache nicht als Erstsprache sprechen und erst weniger als ein Jahr in dieser Sprache Unterricht erhalten (PISA-Konsortium, 2007). Werden in einem Land zu viele Schülerinnen und Schüler ausgeschlossen, die durch ihre Einschränkungen tendenziell schlechtere Ergebnisse erreicht hätten, so werden die Ergebnisse überschätzt (Baumert/Lehmann, 1997).

Umgang mit
Sonderschulen

Umgang mit
anderen
Erstsprachen

Anwendungsbereich: Zusammenfassung der Lehr-/Lernziele



- ❖ (Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ Grundlagen für die Interpretation der Ergebnisse kennenzulernen
 - ❖ die Ergebnisse bisher durchgeführter Studien nachvollziehen zu können (optional)
 - ❖ die Anwendung der Ergebnisse im Schulalltag zu betrachten

Anwendungsbereich: Zusammenfassung der Lehr-/Lernziele



- ❖ Grundlagen für die Interpretation der Ergebnisse kennenzulernen
- ❖ Definition wichtiger Begriffe

Wie bereits im vorangegangenen Kapitel beschrieben, werden die Ergebnisse in [Skalierung](#) Studien des Systemmonitorings skaliert. Die Ergebnisse der Skalierung in den Studien werden durch eine Transformation so standardisiert, dass alle Studien einen Mittelwert von 500 Punkten aufweisen und eine Standardabweichung von 100 Punkten¹⁶.

Definition: Standardabweichung

Die Standardabweichung kann als „Maß der durchschnittlichen Abweichung vom Mittelwert“ definiert werden. Aus ihr ist abzulesen, wie weit sich die Werte um den Mittelwert herum verteilen. Ein hoher Wert sagt aus, dass die Werte weit um den Mittelwert herum verteilt sind, ein kleiner Wert, dass die Werte nah am Mittelwert liegen. Zudem sagt die Standardabweichung auch etwas über die Reichweite der Wertverteilung aus (Albert/Koster, 2002).

Anwendungsbereich: Zusammenfassung der Lehr-/Lernziele



- ❖ die Ergebnisse bisher durchgeführter Studien nachvollziehen zu können (optional)
- ❖ Darstellung der Verteilung auf die Kompetenzniveaus am Beispiel TIMSS

	Deutschland	Frankreich	Niederlande	Norwegen	Schweiz
Alltagsbezogene Schlussfolgerungen	15,4	0,0	3,7	0,6	0,8
Anwendung von einfachen Routinen	36,6	34,3	21,5	35,0	29,3
Bildung von Modellen und Verknüpfung von Operationen	34,1	47,6	41,3	40,2	43,0
Mathematisches Argumentieren	13,9	18,1	33,4	24,2	26,9

Tabelle 11: Verteilung der Schülerinnen und Schüler aus Deutschland, Frankreich, Niederlande, Norwegen und der Schweiz auf die vier Kompetenzniveaus des Bereichs Mathematik

Anwendungsbereich: Zusammenfassung der Lehr-/Lernziele



- ❖ die Anwendung der Ergebnisse im Schulalltag zu betrachten
 - ❖ Beschreibung der Verwendung der Ergebnisse

Festzuhalten ist, dass die Studien nicht zu einem rein wissenschaftlichen Zweck durchgeführt werden. Ihr Ziel ist es (vgl. Kapitel 1.2), „Professionelles pädagogisches Handeln in der Unterrichtspraxis, in Schuladministration und Bildungsverwaltung zu unterstützen“. Daher werden die Ergebnisse so aufbereitet, dass sie von einer breiten Mehrheit wahrgenommen und kritisch hinterfragt werden können. Nur wenn die Ergebnisse verständlich sind, können sie für das eigene Handeln berücksichtigt und übernommen werden (Klieme/Baumert, 2001). Veränderungen im Unterrichtsalltag ergeben sich dann einerseits aus den reflektierten Handlungen der Lehrkräfte und andererseits aus den Impulsen der Bildungspolitik und –administration.

Verwendung
der Ergebnisse

Praktische Implikationen: Zusammenfassung der Lehr-/Lernziele



- ❖ (Zukünftige) Lehrerinnen und Lehrer dazu befähigen,
 - ❖ unterschiedliche Rezeptionsarten der Ergebnisse kennenzulernen
 - ❖ politische Entscheidungen mit den Ergebnissen in Verbindung zu setzen
 - ❖ Veränderungen im Schulalltag mit den Ergebnissen in Verbindung zu setzen

Praktische Implikationen: Zusammenfassung der Lehr-/Lernziele



- ❖ unterschiedliche Rezeptionsarten der Ergebnisse kennenzulernen
- ❖ Verschiedene Arten der Rezeption anhand einer Studie von Kohler beschrieben

Kohler (2005) hat eine Studie durchgeführt, um die Rezeption der unterschiedlichen Akteure auf die Ergebnisse von TIMSS zu untersuchen. Dabei wurden die folgenden drei Personengruppen befragt:

- Lehrerinnen und Lehrer
- Eltern
- Schuladministration

Einige besonders interessante und hilfreiche Ergebnisse dieser Studie sollen hier kurz vorgestellt werden. Informationen zu Anlage und Durchführung dieser Studie können der Literatur entnommen werden.

Praktische Implikationen: Zusammenfassung der Lehr-/Lernziele



- ❖ politische Entscheidungen mit den Ergebnissen in Verbindung zu setzen
- ❖ Handlungsfelder und Schwerpunkte der KMK beschrieben

3.5.2.2 Neue Schwerpunkte der KMK

Am 02.06.2006 hat die KMK eine Gesamtstrategie zum Bildungsmonitoring beschlossen. Innerhalb dieser Strategie werden vier miteinander verbundene Bereiche beschrieben, mit deren Hilfe Bildungsprozesse in Deutschland beobachtet und weiterentwickelt werden sollen. Dazu gehören:

- regelmäßige Teilnahme an internationalen Schulleistungsuntersuchungen
- zentrale Überprüfung der Bildungsstandards im Ländervergleich
- Vergleichsarbeiten (vgl. Studienbrief Teil 2)
- gemeinsame Bildungsberichterstattung

Praktische Implikationen: Zusammenfassung der Lehr-/Lernziele



- ❖ Veränderungen im Schulalltag mit den Ergebnissen in Verbindung zu setzen
- ❖ Verdeutlichung des Einflusses des Bildungsmonitorings auf den Schulalltag an Beispielen

3.5.3.1	Standardentwicklung.....	92
3.5.3.2	Einführung von Kernlehrplänen	94
3.5.3.3	Stärkung der Sprachkompetenz	95
3.5.3.4	Ganztagsschulen.....	97
3.5.3.5	SINUS / SINUS transfer	98
3.5.3.6	For.mat	99
3.5.3.7	UDiKom.....	99
3.5.3.8	Bildungsberichterstattung.....	99



❖ Vielen Dank für Ihre Aufmerksamkeit!