

*Monitoring des Umsetzungsstandes der im Nationalen Aktionsplan  
Wirtschaft und Menschenrechte 2016–2020 beschriebenen  
menschenrechtlichen Sorgfaltspflicht von Unternehmen*

# Abschlussbericht



## Präambel

Mit der Erhebung zur Überprüfung des Umsetzungsstandes der im Nationalen Aktionsplan Wirtschaft und Menschenrechte 2016–2020 (NAP) beschriebenen menschenrechtlichen Sorgfaltspflicht von Unternehmen (Monitoring) hat das Auswärtige Amt die Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft (EY) in Zusammenarbeit mit ihren Konsortialpartnern adelphi consult GmbH, Sustain Consulting GmbH und focusright GmbH beauftragt.

Der vorliegende Abschlussbericht verfolgt die folgenden, in der Leistungsbeschreibung definierten Anliegen:

- Analyse und Bewertung der Gesamtergebnisse (2018–2020)
- Darstellung und Diskussion des methodischen Vorgehens
- Diskussion, inwiefern sich das Monitoring des NAP bewährt hat und inwiefern es effektiv und effizient ist, dies ggf. in weiteren NAPs bzw. gesetzlichen Vorgaben fortzusetzen.

Der Entwurf des Abschlussberichts wurde dem IMA am 29. September 2020 zur Kommentierung vorgelegt. Die Abnahme des Berichts erfolgte am 8. Oktober 2020 durch den IMA.

Ihr Ansprechpartner aus dem Referat Wirtschaft und Menschenrechte im Auswärtigen Amt bei Rückfragen zum Abschlussbericht:

Holger Dreiseitl  
Auswärtiges Amt  
Referat Wirtschaft und Menschenrechte (401)  
Telefon: +49 1817 2270 (Sekretariat)  
401-R@diplo.de

Ihre Ansprechpartnerin aus dem Konsortium bei Rückfragen zum Abschlussbericht:

Nicole Richter  
Partner  
Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft  
Climate Change & Sustainability Services  
Telefon: +49 89 14331 19332  
nap.monitoring@de.ey.com

## Inhalt

Abbildungs- und Tabellenverzeichnis.....	IV
1 Einleitung.....	1
1.1 Ausgangslage .....	1
1.2 Projektziel und -überblick.....	1
1.3 Inhalte und Funktion des Abschlussberichts.....	3
2 Analyse und Bewertung der Gesamtergebnisse .....	5
2.1 Zusammenfassung der Ergebnisse der Erhebungsphasen 2018–2020.....	5
2.2 Fokus Risikoanalyse und Maßnahmen .....	8
2.3 Vertiefte Analyse der Erhebungsdaten 2019 und 2020 .....	10
3 Reflexion des Monitorings .....	12
3.1 Reflexion des Monitoring-Prozesses .....	12
3.2 Reflexion der Erhebungsmethodik.....	14
3.2.1 Definition der Grundgesamtheit .....	14
3.2.2 Bewertungskategorien .....	15
3.2.3 Fragebogen und Antwortmöglichkeiten .....	15
3.2.4 Anforderungsrahmen .....	16
3.2.5 Stufenverfahren.....	16
3.2.6 Prüfung der „Comply or explain“- und Freitext-Antworten .....	17
3.3 Reflexion der NAP-Anforderungen und des Anforderungsrahmens.....	17
3.4 Limitationen der Methodik .....	19
Anhang: Methodik der vertieften Analyse .....	21

## Abbildungs- und Tabellenverzeichnis

Abbildung 1: Projektablauf im Überblick .....	3
Abbildung 2: Ergebnisse der repräsentativen Erhebung 2020.....	5
Abbildung 3: Ergebnisse der repräsentativen Erhebung 2019.....	6
Abbildung 4: Ergebnisse der repräsentativen Erhebungen 2019 und 2020 im Vergleich .....	7
Tabelle 1: Menschenrechtsbezogene Maßnahmen in der Wertschöpfungskette.....	10

# 1 Einleitung

## 1.1 Ausgangslage

Am 21. Dezember 2016 wurde der „Nationale Aktionsplan – Umsetzung der VN-Leitprinzipien für Wirtschaft und Menschenrechte (2016–2020)“ (NAP) vom Bundeskabinett verabschiedet. Darin formuliert die Bundesregierung ihre Erwartung an Unternehmen bezüglich der Umsetzung der menschenrechtlichen Sorgfaltspflicht und der Achtung der Menschenrechte entlang unternehmerischer Liefer- und Wertschöpfungsketten. Die Definition der menschenrechtlichen Sorgfaltspflicht orientiert sich dabei an den Leitprinzipien für Wirtschaft und Menschenrechte der Vereinten Nationen (VN-Leitprinzipien), die sich wiederum auf die international anerkannten Menschenrechte beziehen. Damit bilden sie einen allgemein anerkannten Referenzrahmen für menschenrechtliche Schutzpflichten von Staaten und für die Verantwortung von Unternehmen, die Menschenrechte zu achten.

In Kapitel III des NAP hat die Bundesregierung ihre Erwartung an Unternehmen formuliert, den im NAP beschriebenen Prozess der unternehmerischen Sorgfalt in einer ihrer Größe, Branche und Position in der Liefer- und Wertschöpfungskette adäquaten Weise einzuführen. Dieser Prozess umfasst fünf Kernelemente menschenrechtlicher Sorgfaltspflicht:

1. eine Grundsatzerklärung zur Achtung der Menschenrechte
2. ein Verfahren zur Ermittlung tatsächlich und potenziell nachteiliger Auswirkungen auf die Menschenrechte
3. Maßnahmen zur Abwendung negativer Auswirkungen und zur Überprüfung der Wirksamkeit dieser Maßnahmen
4. die Berichterstattung
5. einen Beschwerdemechanismus

## 1.2 Projektziel und -überblick

Hinsichtlich des Standes der Umsetzung der fünf Kernelemente in den Unternehmen hat das Auswärtige Amt 2018 ein Monitoring gestartet, das mit der letzten Erhebungsphase 2020 nun abgeschlossen ist. Somit wurde eine Maßnahme umgesetzt, die der NAP so vorgesehen hat. Der Interministerielle Ausschuss (IMA) unter Vorsitz des Auswärtigen Amtes wurde über den gesamten Monitoring-Prozess hinweg eng eingebunden. Die Aktivitäten des IMA hat die Arbeitsgruppe Wirtschaft und Menschenrechte des Nationalen CSR-Forums (AG Wirtschaft und Menschenrechte), die VertreterInnen der Sozialpartner, aus der Wirtschaft und der Zivilgesellschaft umfasst, begleitet. Das Monitoring hatte das im NAP festgelegte Ziel zu ermitteln, ob bis 2020 mindestens 50 Prozent der in Deutschland ansässigen Unternehmen mit über 500 Beschäftigten die im NAP beschriebenen Kernelemente menschenrechtlicher Sorgfaltspflicht in ihre Unternehmensprozesse integriert haben. Neben dem Stand der Umsetzung der menschenrechtlichen Sorgfaltspflicht sollten auch qualitative Aussagen zur inhaltlichen Tiefe der Umsetzungen, zu den Herausforderungen sowie zum Aufwand für Unternehmen getroffen werden.

Sofern keine ausreichende Umsetzung der Kernelemente erfolgt, sieht der NAP vor, dass die Bundesregierung weiter gehende Schritte bis hin zu gesetzlichen Maßnahmen prüft (NAP, S. 10). Der

aktuelle Koalitionsvertrag<sup>1</sup> besagt: „Falls die wirksame und umfassende Überprüfung des NAP 2020 zu dem Ergebnis kommt, dass die freiwillige Selbstverpflichtung der Unternehmen nicht ausreicht, werden wir national gesetzlich tätig und uns für eine EU-weite Regelung einsetzen.“

Die Leistungsbeschreibung<sup>2</sup> zum Monitoring sah insgesamt drei Erhebungsphasen vor, je eine in den Jahren 2018 (explorative Erhebungsphase), 2019 (erste repräsentative Erhebung) und 2020 (zweite repräsentative Erhebung).<sup>3</sup> Die im Inception Report (veröffentlicht am 4. September 2018) beschriebene Struktur des Bewertungssystems wurde im Rahmen der explorativen Phase 2018 kritisch überprüft und weiterentwickelt.<sup>4</sup> Hierfür wurden Interviews mit 30 Unternehmen und neun Stakeholder-VertreterInnen, die von der AG Wirtschaft und Menschenrechte ausgewählt wurden, geführt, um qualitative Erkenntnisse zur Umsetzung menschenrechtlicher Sorgfaltspflichten zu generieren. Im Zwischenbericht zur explorativen Phase 2018 vom 5. Juli 2019<sup>5</sup> wurden die Erkenntnisse der explorativen Phase, die Planung des späteren Vorgehens und die Weiterentwicklung des methodischen Ansatzes für die nachfolgenden, repräsentativen Erhebungsphasen dargelegt.<sup>6</sup>

In den repräsentativen Monitoring-Phasen von 2019 und 2020 wurden quantitative Ergebnisse im Hinblick auf den Stand der Umsetzung der Kernelemente generiert. Hierfür wurden per Zufallsverfahren statistisch repräsentative Stichproben gezogen. Die Unternehmen der Stichproben wurden jeweils um die Teilnahme an einer Online-Umfrage gebeten, in der die Unternehmen zum Umsetzungsstand der Anforderungen jedes NAP-Kernelements befragt wurden. Mithilfe eines durch den NAP vorgesehenen „Comply or explain“-Mechanismus konnten die Unternehmen darüber hinaus zu jeder Frage erläuternde Angaben machen, falls ihr Umsetzungsstand nicht den zuvor vom IMA festgelegten Anforderungen entsprach. Im Anschluss wurden die Selbstauskünfte der Unternehmen nach einem festgelegten Verfahren vom Konsortium überprüft. Die Zwischenberichte 2019<sup>7</sup> und 2020<sup>8</sup> erläuterten jeweils die Befunde und beschrieben das methodische Vorgehen der repräsentativen Erhebungsphasen von 2019 und 2020. Für die Beantwortung der Kernfrage des Monitorings, ob im Jahr 2020 mindestens 50 Prozent der Unternehmen mit Sitz in Deutschland und mehr als 500 Beschäftigten die im NAP beschriebenen Prozesse menschenrechtlicher Sorgfalt in ihre Unternehmensprozesse integriert haben, sind gemäß der Zielvorgabe im NAP die Ergebnisse der repräsentativen Erhebung 2020 maßgeblich.

---

<sup>1</sup> Siehe Koalitionsvertrag zwischen CDU, CSU und SPD (Berlin, 12.03.2018): „Ein neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land“, online verfügbar unter:

<https://www.bundesregierung.de/resource/blob/975226/847984/5b8bc23590d4cb2892b31c987ad672b7/2018-03-14-koalitionsvertrag-data.pdf?download=1>, letzter Zugriff am 31.07.2020.

<sup>2</sup> Ausschreibung „Monitoring NAP“ vom 29.12.2017, Referenznummer der Bekanntmachung: VV-118-2017-0201 durch das Auswärtige Amt

<sup>3</sup> Für weitere Informationen siehe Abbildung 1.

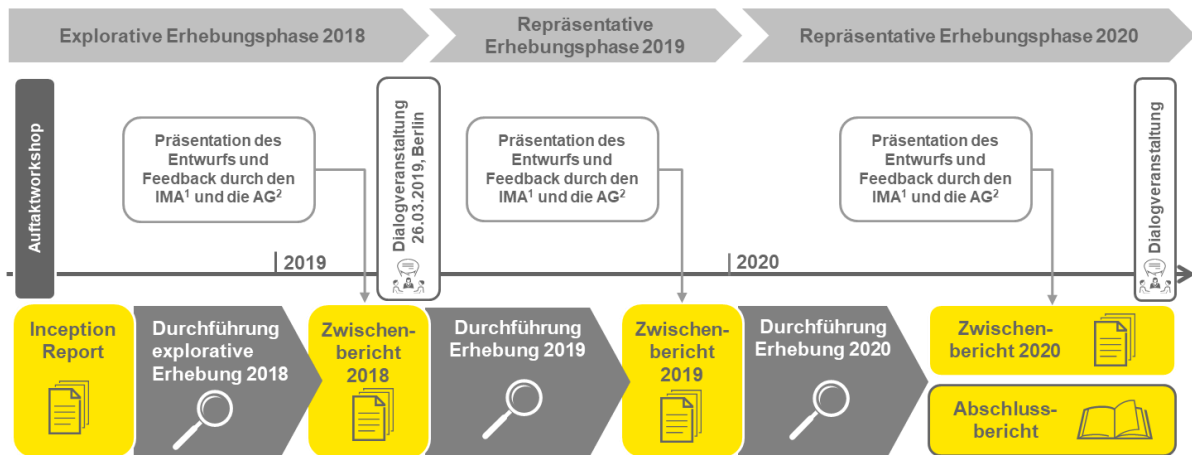
<sup>4</sup> Der Gesamtprojektverlauf des Monitorings wurde im Inception Report vom 04.09.2018 ausführlich dargestellt. Der Inception Report ist hier zu finden: <https://www.auswaertiges-amt.de/blob/2232456/8d910dd921d773bdafccfb106dff46f2/inception-report-data.pdf>, letzter Zugriff am 26.06.2020.

<sup>5</sup> Der Zwischenbericht zur explorativen Erhebungsphase 2018 ist hier zu finden: <https://www.auswaertiges-amt.de/blob/2232418/1531aad304f1dec719954f7292ddbc05/190710-nap-zwischenbericht-data.pdf>, letzter Zugriff am 26.06.2020.

<sup>6</sup> Siehe Zwischenbericht 2018, Fußnote 6 für weitere Informationen.

<sup>7</sup> Der Zwischenbericht zur repräsentativen Erhebungsphase 2019 ist hier zu finden: <https://www.auswaertiges-amt.de/blob/2314274/3a52de7f2c6103831ba0c24697b7739c/20200304-nap-2-zwischenbericht-data.pdf>; letzter Zugriff am 23.09.2020.

<sup>8</sup> Der Zwischenbericht zur repräsentativen Erhebungsphase 2020 ist hier zu finden: <https://www.auswaertiges-amt.de/blob/2384126/f793a1412234e5c281abc876f05e6974/200915-nap-3-bericht-data.pdf>; letzter Zugriff am 23.09.2020.



- 1) Interministerieller Ausschuss (IMA) unter Federführung des Auswärtigen Amtes
- 2) Arbeitsgruppe (AG) Wirtschaft und Menschenrechte des Nationalen CSR-Forums

Abbildung 1: Projektablauf im Überblick

Das Monitoring wurde im Auftrag der Bundesregierung durch ein Konsortium bestehend aus der Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft (EY), der adelphi consult GmbH (adelphi), der Systain Consulting GmbH (Systain) und der focusright GmbH (focusright) durchgeführt. EY trägt dabei die Verantwortung für das Gesamtprojektmanagement. Detaillierte Informationen zum Konsortium enthält der Inception Report.

### 1.3 Inhalte und Funktion des Abschlussberichts

Die in diesem Abschlussbericht dargestellten Gesamtergebnisse basieren auf den vor Auftragsvergabe definierten Vorgaben durch das Auswärtige Amt (Leistungsbeschreibung vom 29. Dezember 2017). Diese Vorgaben sehen vor, dass die Erhebungen 2019 und 2020 auf der Basis einer repräsentativen Stichprobe aller in Deutschland ansässigen Unternehmen mit mehr als 500 Beschäftigten durchgeführt werden. Dies wurde sowohl mit der Erhebung 2019 als auch mit der Erhebung 2020 erfüllt.

Die Schwerpunkte des vorliegenden Abschlussberichts liegen zum einen in der Analyse und Bewertung der Gesamtergebnisse und zum anderen in der Rekapitulation und Reflexion des Monitoring-Prozesses, der Methodik und der Ergebnisse.

Zunächst werden die Ergebnisse aller Erhebungsphasen (2018–2020) zusammengefasst, ausgewertet und einander gegenübergestellt (Kapitel 2.1). Beginnend mit einer Zusammenfassung werden die Ergebnisse vor allem hinsichtlich der unterschiedlichen Erfüllungsquoten pro Kernelement und den damit verbundenen Herausforderungen für Unternehmen bewertet. Ein besonderer Fokus liegt auf der Risikoanalyse, da deren Erfüllungsquote mitunter am geringsten ist (Kapitel 2.2). Außerdem wurde eine vertiefte Analyse der Erhebungsdaten 2019 und 2020 bezogen auf einzelne Schichten, Branchen und Wertschöpfungsstufen vorgenommen (Kapitel 2.3).

Die Reflexion des Monitorings betrachtet sowohl die Erhebungsmethode als auch den Monitoring-Prozess insgesamt (Kapitel 3.1). Darin wird die Entwicklung der Erhebungsmethode rekapituliert und erläutert, inwiefern sich die methodischen Ansätze bewährt haben (Kapitel 3.2). Dahin gehend wird auch der Anforderungsrahmen des Monitorings reflektiert und zu den verfügbaren zeitlichen Ressourcen und anderen Herausforderungen, z. B. bezüglich der „Strenge der Bewertung“ von Antworten, in Relation gesetzt (Kapitel 3.3), um schließlich die Limitationen der Methodik aufzuzeigen (Kapitel 3.4). Zuletzt werden die Ergebnisse des Monitorings und ihre Bedeutung für die weitere Diskussion um die menschenrechtliche Sorgfaltspflicht von Unternehmen reflektiert (Kapitel 3.5).



## 2 Analyse und Bewertung der Gesamtergebnisse

In diesem Kapitel werden die wichtigsten Ergebnisse der drei Erhebungsphasen 2018 bis 2020 rekapituliert. Außerdem werden Erkenntnisse aus vertieften Analysen des Datenmaterials der Erhebungen 2019 und 2020 aufgezeigt.

### 2.1 Zusammenfassung der Ergebnisse der Erhebungsphasen 2018–2020

Die drei im Rahmen des NAP-Monitorings durchgeführten Erhebungen sind unabhängig voneinander. Das bedeutet, dass die Ergebnisse der einzelnen Erhebungsphasen keine Rückschlüsse auf eine zeitliche Entwicklung des Umsetzungsstandes der NAP-Anforderungen der Unternehmen zulassen. Sämtliche Erhebungsergebnisse sind Momentaufnahmen des jeweiligen Erhebungszeitraums. Während die Unternehmen der explorativen Erhebung 2018 gezielt ausgewählt wurden, wurden diejenigen der repräsentativen Erhebungsphasen 2019 und 2020 mittels Zufallsstichprobe bestimmt.

Für die Beantwortung der Kernfrage des Monitorings, ob im Jahr 2020 mindestens 50 Prozent der Unternehmen mit Sitz in Deutschland und mehr als 500 Beschäftigten die im NAP beschriebenen Prozesse menschenrechtlicher Sorgfalt in ihre Unternehmensprozesse integriert haben, sind gemäß Zielvorgabe im NAP die Ergebnisse der repräsentativen Erhebung 2020 maßgeblich. Auf der Basis der Auswertung und der angewandten statistischen Gewichtungsverfahren<sup>9</sup> gelten 13 bis 17 Prozent der Unternehmen als „Erfüller“, 83 bis 87 Prozent der Unternehmen sind „Nicht-Erfüller“ und weniger als 1 Prozent sind „Unternehmen mit Umsetzungsplan“. Innerhalb der Gruppe der „Nicht-Erfüller“ befinden sich bezogen auf die Grundgesamtheit 10 bis 12 Prozent der Unternehmen „auf einem guten Weg“ zur Erfüllung der NAP-Anforderungen.<sup>10</sup>

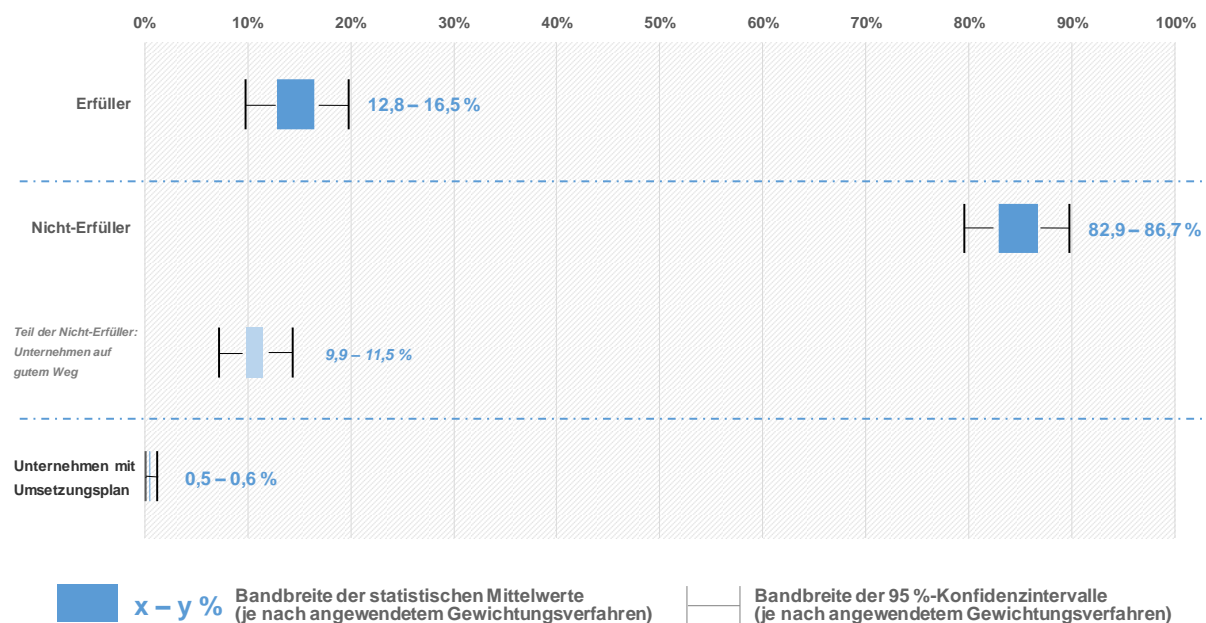


Abbildung 2: Ergebnisse der repräsentativen Erhebung 2020

<sup>9</sup> Zur Gewährleistung eines repräsentativen Ergebnisses und zum Ausgleich der Selektion nach beobachtbaren Faktoren wurden insgesamt vier verschiedene Gewichte berechnet. Aus statistischer Perspektive ist die Anwendung jedes der berechneten Gewichte begründbar. Die Ergebnisse bezüglich der Erfüllungskategorien wurden daher als Bandbreite dieser vier Gewichte dargestellt.

<sup>10</sup> Siehe Zwischenbericht 2020, S. 18 f.

Die Ergebnisse der repräsentativen Erhebung 2019 bewegen sich in ähnlichen Größenordnungen wie 2020. Demnach sind 17 bis 19 Prozent der Unternehmen „Erfüller“ und 78 bis 81 Prozent „Nicht-Erfüller“. „Unternehmen mit Umsetzungsplan“ haben einen Anteil von 2 bis 3 Prozent. Die „Unternehmen auf gutem Weg“ als Teilmenge der „Nicht-Erfüller“ liegen bei einem Anteil von 9 bis 12 Prozent.<sup>11</sup>

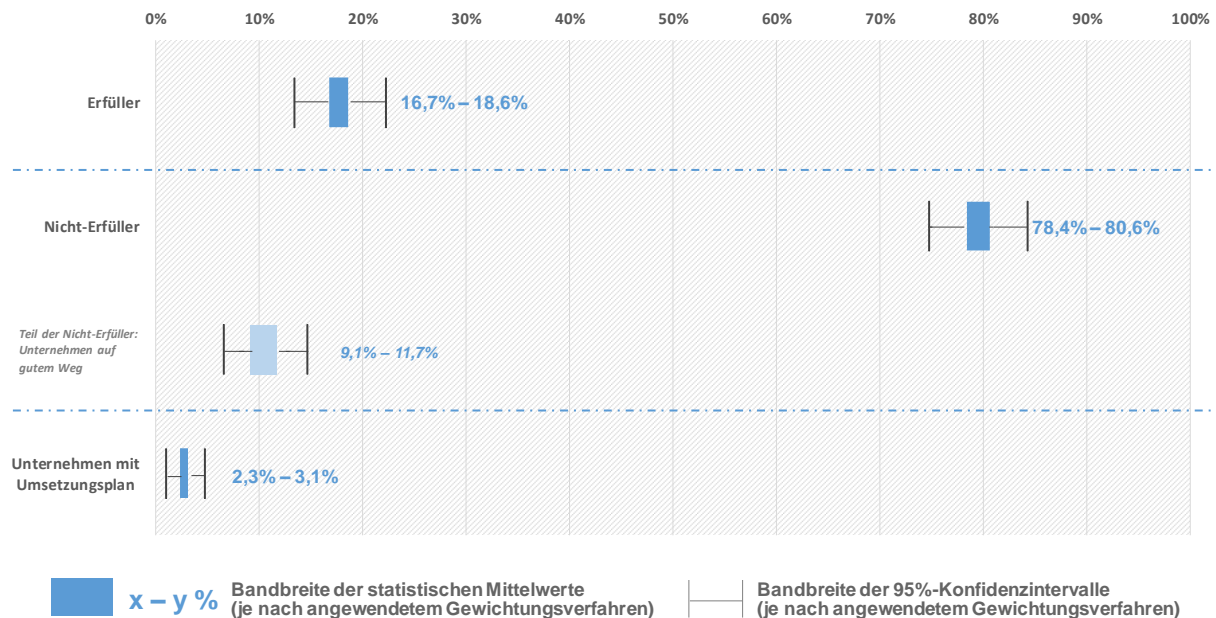


Abbildung 3: Ergebnisse der repräsentativen Erhebung 2019

Aufgrund der im Wesentlichen gleich gebliebenen Methodik der Erhebungen 2019 und 2020 lassen sich die Ergebnisse sinnvoll miteinander vergleichen. Die Qualität und Belastbarkeit der Erhebungsergebnisse zeigt sich an der Überlappung der Bandbreiten der 95-Prozent-Konfidenzintervalle<sup>12</sup>. Diese ist nicht nur im Hinblick auf den Erfüllerstatus insgesamt gegeben, sondern auch mit Blick auf den Erfüllungsgrad der einzelnen Kernelemente.

<sup>11</sup> Siehe Zwischenbericht 2019, S. 18f.

<sup>12</sup> Ein Konfidenzintervall ist die Bandbreite, innerhalb der der wahre Wert des arithmetischen Mittelwerts der statistischen Erhebung bezogen auf die Grundgesamtheit mit einer definierten Wahrscheinlichkeit (Konfidenzniveau; hier 95 Prozent) liegt.

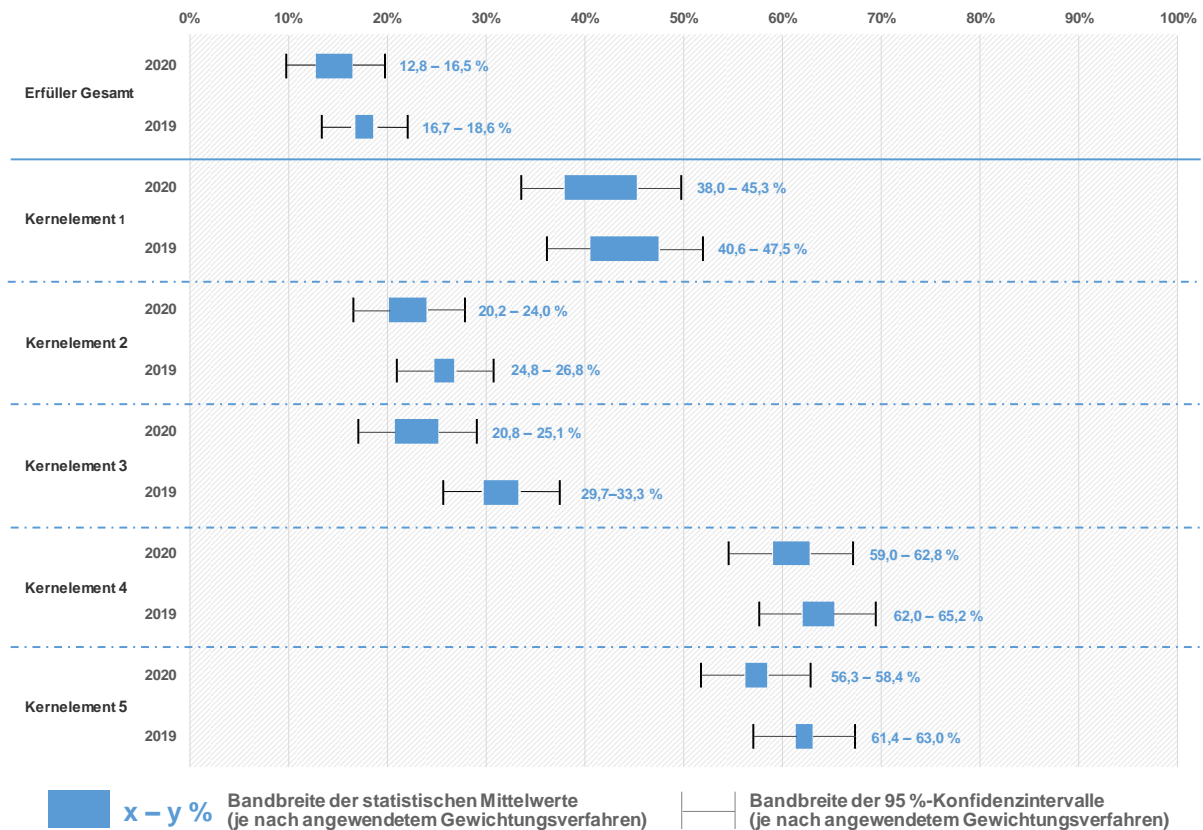


Abbildung 4: Ergebnisse der repräsentativen Erhebungen 2019 und 2020 im Vergleich

Auffällig ist, dass die Erfüllungsgrade 2020 im Vergleich zu 2019 leicht nach links verschoben sind. Das liegt daran, dass die angewendeten Gewichte einen stärkeren Einfluss auf die Ergebnisse 2020 haben, als dies 2019 der Fall war, und nicht an einer tatsächlichen Verschlechterung der Erfüllerquote. Das wiederum ist auf eine verbesserte Datenbasis 2020 gegenüber 2019 bei der Bruttostichprobe zurückzuführen.<sup>13</sup> Die verbesserte Datenbasis führte dazu, dass Verzerrungen durch beobachtbare Faktoren 2020 besser in der Gewichtung berücksichtigt werden konnten. Es lässt sich zudem – mit einer Ausnahme – nicht erkennen, dass die geringfügige Anpassung des Fragebogens zwischen den repräsentativen Erhebungen Einfluss auf die Ergebnisse hatte. Lediglich der vergleichsweise deutlich geringere Erfüllungsgrad von Kernelement 3 kann auf eine Änderung der Antwortoptionen in der Eingangsfrage zum Kernelement zurückgeführt werden. Weitere ausführliche Erläuterungen zur Vergleichsanalyse finden sich in Kapitel 3.3 des Zwischenberichts 2020.

Die Unternehmen hatten in den beiden repräsentativen Erhebungen außerdem die Möglichkeit, im Online-Fragebogen ihre Herausforderungen und Wünsche bezüglich der Umsetzung der NAP-Anforderungen durch eine Freitextoption zu den Kernelementen zu äußern. Auch hier ähneln sich die Angaben in den Erhebungen 2019 und 2020 stark. Die mit Abstand am häufigsten genannte Herausforderung ist die Komplexität der Wertschöpfungskette, was die Identifikation und die Minderung von Menschenrechtsrisiken erschwert. Ebenfalls oft genannt wurden Herausforderungen aufgrund von unterschiedlichen internationalen Rahmenbedingungen – sowohl bezogen auf das regulatorische Umfeld als auch auf kulturelle Aspekte. Bei den Wünschen beziehen sich die Angaben auf mehr Unterstützung der Bundesregierung und von Verbänden, zum Beispiel in Form von

<sup>13</sup> Für die Unternehmen der erweiterten Stichprobe 2019 wurde keine Internetrecherche hinsichtlich der Existenz einer Grundsatzklärung durchgeführt, sodass die Information in dieser Eigenschaft fehlte. Bei der Stichprobe 2020 wurde dagegen für alle Unternehmen eine solche Recherche durchgeführt.

Guidelines zur menschenrechtlichen Sorgfaltspflicht oder der Etablierung einheitlicher Standards. Weiterhin haben sich einige Unternehmen für Regulierungen auf internationaler Ebene ausgesprochen, um einheitliche Voraussetzungen und Rahmenbedingungen für eine systematische Umsetzung des Themas über die Landesgrenzen hinaus zu schaffen.<sup>14</sup>

## 2.2 Fokus Risikoanalyse und Maßnahmen

Im Rahmen der Auswertung wurden in beiden repräsentativen Erhebungen Detailanalysen der einzelnen Kernelemente durchgeführt, um zu ergründen, welche Faktoren der Umsetzung der menschenrechtlichen Sorgfalt durch die Unternehmen das Ergebnisbild entscheidend beeinflusst haben. Wie auch die Kernergebnisse stimmen diese Faktoren in weiten Teilen überein. So bestehen die größten Herausforderungen für Unternehmen im Bereich der menschenrechtlichen Risikoanalyse bzw. in der Einrichtung eines Verfahrens zur Ermittlung nachteiliger Auswirkungen auf die Menschenrechte und der Umsetzung angemessener und risikobasierter Maßnahmen. Zum einen existiert eine solche Risikobetrachtung bei rund der Hälfte der Unternehmen nicht, zum anderen bezieht die Analyse in einigen Fällen nicht alle relevanten Bestandteile der Wertschöpfungskette ein.<sup>15</sup> Dies entspricht auch den Ergebnissen der explorativen Erhebung 2018. In den Interviews betonten die Unternehmen methodische Herausforderungen wie eine eindeutige Abgrenzung von Menschenrechtsthemen, die Wesentlichkeitseinschätzung von Risiken und fehlendes Wissen und Ressourcen zur Durchführung einer Risikoanalyse.<sup>16</sup>

Der angemessene Umfang der Risikoanalyse – also die angemessene Einbeziehung aller relevanten Stufen der Wertschöpfungskette – ist eine zentrale Anforderung des NAP. Die nähere Analyse der Daten der repräsentativen Erhebungen zeigt, dass ein Zusammenhang zwischen einem angemessenen Umfang der Risikoanalyse und einem umfassenderen Risikoprofil besteht. So ist die Korrelation zwischen dem Erfüllungsgrad von Frage 2.2.1<sup>17</sup> und den Variablen des Risikoprofils (Fragen 2.3.1 bis 2.3.4)<sup>18</sup> für fast jede dieser Variablen positiv signifikant. Das bedeutet, dass Unternehmen, die einen ausreichenden Umfang der Risikoanalyse aufweisen, eher Risiken in den einzelnen Wertschöpfungsstufen und dabei eher hohe Risiken identifizieren und mehr Menschenrechtsthemen und potenziell betroffene Gruppen angeben. Dagegen ist die Korrelation zwischen dem Erfüllungsgrad von Frage 2.2.1 und der Antwort auf Frage 2.3.1, dass keine Risiken identifiziert wurden, stark negativ signifikant. Das bedeutet, dass Unternehmen, deren Risikoanalyse einen ausreichenden Umfang hatte, nur in ganz wenigen Fällen keine Menschenrechtsrisiken identifiziert haben.

Es wurde außerdem untersucht, ob Unternehmen, die spezifische Wertschöpfungsstufen in ihre Risikoanalyse einbeziehen, in der Gesamtbetrachtung eher „Erfüller“ sind. Diese Korrelation ist für die direkte Lieferkette, die indirekte Lieferkette und Produkte und Dienstleistungen positiv signifikant. Das heißt, dass Unternehmen, die die genannten drei Stufen aus ihrer Risikoanalyse

---

<sup>14</sup> Siehe Zwischenbericht 2020, S. 31 sowie Zwischenbericht 2019, S. 28.

<sup>15</sup> Siehe Zwischenbericht 2020, S. 22.

<sup>16</sup> Siehe Zwischenbericht 2018, S. 36.

<sup>17</sup> Frage 2.2.1: Welche Stufen der Wertschöpfungskette analysieren Sie im Hinblick auf Menschenrechtsrisiken?

<sup>18</sup> Frage 2.3.1: Für welche Stufen der Wertschöpfungskette haben Sie potenziell negative Auswirkungen auf Menschenrechte identifiziert?

Frage 2.3.2: Für welche Stufen der Wertschöpfungskette haben Sie potenziell besonders hohe negative Auswirkungen auf Menschenrechte identifiziert?

Frage 2.3.3: Welche Menschenrechtsthemen haben Sie bei Ihrer Risikoanalyse identifiziert?

Frage 2.3.4: Welche potenziell betroffenen Gruppen haben Sie bei Ihrer Risikoanalyse identifiziert?

ausgeklammert haben, in der Gesamtbewertung eher „Nicht-Erfüller“ waren. Zusammenfassend lässt sich also sagen, dass in diesem Zusammenhang nicht eine bestimmte Wertschöpfungsstufe allein zur Nicht-Erfüllung vieler Unternehmen beigetragen hat; die Nicht-Erfüllung ist vielmehr u. a. darauf zurückzuführen, dass bei der Risikoanalyse besonders häufig all jene Stufen der Wertschöpfungskette ausgeklammert werden, die generell über das eigene Unternehmen hinausgehen.

Eine weitere große Herausforderung scheint es in der Umsetzung angemessener Maßnahmen zur Abwendung von Menschenrechtsrisiken zu geben. Diese sollen laut NAP auf der Basis der Ergebnisse der Risikoanalyse festgelegt werden. Die beiden Kernelemente hängen also eng miteinander zusammen, was vermutlich auch erklärt, warum ihr Erfüllungsgrad über beide repräsentative Erhebungen hinweg ähnlich (gering) ist. Dadurch, dass bei einem Großteil der Unternehmen bereits die menschenrechtliche Risikoanalyse entweder nicht vorhanden oder unzureichend im Sinne der NAP-Anforderungen ist, lassen sich aus einer solchen (nicht vorhandenen oder unzureichenden) Risikoanalyse keine adäquaten bzw. angemessenen Maßnahmen ableiten. Das wird auch daran deutlich, dass in beiden repräsentativen Erhebungen nur rund ein Drittel der Unternehmen angaben, dass sie Maßnahmen auf der Grundlage der menschenrechtlichen Risikoanalyse identifiziert haben.<sup>19</sup> Die Unternehmen der explorativen Erhebung 2018 gaben an, dass die wesentlichen Herausforderungen im Bereich des mangelnden Einflusses auf die risikobehafteten Teile der Wertschöpfungskette und der unterschiedlichen rechtlichen und kulturellen weltweiten Rahmenbedingungen liegen.<sup>20</sup> Bei den Teilnehmern der explorativen Erhebung handelte es sich überwiegend um Unternehmen, die sich bereits eingehender mit menschenrechtlichen Sorgfaltspflichten auseinandergesetzt hatten.<sup>21</sup> Der Befund aus der explorativen Erhebung deckt sich mit den ergänzenden Angaben der Unternehmen am Ende des Fragebogens in den beiden repräsentativen Erhebungsphasen.<sup>22</sup> Zusammenfassend lässt sich festhalten, dass die Herausforderungen im Bereich der Maßnahmen menschenrechtlicher Sorgfalt auf zwei Ebenen liegen: Zum einen bestehen Herausforderungen hinsichtlich der grundlegenden Identifikation geeigneter Maßnahmen zur Minderung von Menschenrechtsrisiken, zum anderen fällt es den Unternehmen mitunter schwer, identifizierte notwendige Maßnahmen erfolgreich operativ umzusetzen und deren Wirksamkeit zu überprüfen.

Im Rahmen der Erhebung wurden zusätzlich die Freitextangaben der Unternehmen zu den menschenrechtsbezogenen Maßnahmen in den jeweiligen Stufen der Wertschöpfungskette ausgewertet, um einen Überblick über das mögliche Vorgehen zur Minderung von Menschenrechtsrisiken zu erhalten:

Stufe der Wertschöpfungskette	Häufig genannte menschenrechtsbezogene Maßnahmen
Eigener Betrieb	<ul style="list-style-type: none"> <li>• Schulungen</li> <li>• Verhaltenskodex bzw. Grundsatzerklärung</li> <li>• Beschwerdemechanismus bzw. Melde-Hotline</li> <li>• Integration in Managementsysteme (z. B. Arbeitssicherheit oder Personal)</li> <li>• Integration in Compliance-Prozesse</li> </ul>

<sup>19</sup> Siehe Zwischenbericht 2020, Anhang 4 sowie Zwischenbericht 2019, Anhang 4.

<sup>20</sup> Siehe Zwischenbericht 2018, S. 38 f.

<sup>21</sup> Siehe Zwischenbericht 2018, S. 31 f.

<sup>22</sup> Siehe Zwischenbericht 2020, S. 31 sowie Zwischenbericht 2019, S. 28.

Verbundene Unternehmen	<ul style="list-style-type: none"> <li>• Schulungen</li> <li>• Verhaltenskodex bzw. Grundsatzerklärung</li> <li>• Beschwerdemechanismus bzw. Melde-Hotline</li> <li>• Integration in Compliance-Prozesse</li> <li>• Interne Audits/Revision</li> </ul>
Direkte Lieferkette	<ul style="list-style-type: none"> <li>• Lieferantenverpflichtung (z. B. Verhaltenskodex für Lieferanten)</li> <li>• Durchführung von Audits</li> <li>• Lieferantenbewertungen</li> <li>• Schulungen</li> </ul>
Indirekte Lieferkette	<ul style="list-style-type: none"> <li>• Durchführung von Audits</li> <li>• Beteiligung an Brancheninitiativen</li> <li>• Pläne für Abhilfemaßnahmen</li> <li>• Förderung der Lieferkettentransparenz</li> <li>• Integration in Managementsysteme (z. B. Einkauf)</li> </ul>
Produkte/Dienstleistungen/Projekte	<ul style="list-style-type: none"> <li>• Qualitätsmanagement</li> <li>• Beteiligung an Brancheninitiativen</li> <li>• Zertifizierungen</li> </ul>
Investitionstätigkeit	<ul style="list-style-type: none"> <li>• Monitoring der Einhaltung der vom Unternehmen definierten Standards</li> <li>• Berücksichtigung externer Standards (z. B. PRI – Principles for Responsible Investment)</li> </ul>

Tabelle 1: Menschenrechtsbezogene Maßnahmen in der Wertschöpfungskette

### 2.3 Vertiefte Analyse der Erhebungsdaten 2019 und 2020

Für die Erstellung dieses Abschlussberichts wurden vertiefte Analysen des Datenmaterials der repräsentativen Erhebung 2019 und 2020 durchgeführt. Zu diesem Zweck wurden die Nettostichproben 2019 und 2020 miteinander kombiniert. Eine detaillierte Erläuterung der Methodik der Analysen findet sich im Anhang. **Durch die Betrachtung definierter Teilmengen der Nettostichproben 2019 und 2020 sind die folgenden Analyseergebnisse nicht repräsentativ für die Grundgesamtheit. In diesem Kapitel können daher lediglich vergleichende qualitative Tendenzen aufgezeigt werden.** Es lässt sich zudem nicht ausschließen, dass einzelne Ausprägungen auf Selbstselektion der Unternehmen bei der Teilnahme am Monitoring zurückzuführen sind. Dennoch geben die Tendenzen Aufschluss über den Erfüllungsgrad u. a. in einzelnen Schichten, Branchen und Wertschöpfungsstufen verglichen mit dem Erfüllungsgrad der Unternehmen insgesamt.

Zur Bestimmung der Bruttostichprobe wurden die Unternehmen in sechs Schichten eingegliedert. Sie umfassen die drei Größenklassen 501 bis 1.000 Beschäftigte, 1001 bis 2.000 Beschäftigte und mehr als 2.000 Beschäftigte sowie die beiden Wirtschaftszweige produzierendes Gewerbe sowie Handel und Dienstleistungen. Die Informationen stammen aus der Bisnode-Datenbank. Die Betrachtung dieser Schichten zeigt, dass die Unternehmen aus dem Bereich Handel und Dienstleistungen über alle Größenklassen hinweg einen leicht überdurchschnittlichen<sup>23</sup> Erfüllungsgrad aufweisen. Im produzierenden Gewerbe ist der Erfüllungsgrad bei Unternehmen mit mehr als 2.000 Beschäftigten

<sup>23</sup> Der Durchschnitt entspricht dem 95-Prozent-Konfidenzintervall der „Erfüller“ auf der Basis der gewichteten Daten der kombinierten Nettostichproben (siehe Anhang).

ähnlich. Dagegen nimmt die Erfüllerquote im produzierenden Gewerbe mit sinkender Größe der Unternehmen ab. Bei produzierenden Unternehmen mit 501 bis 1.000 Beschäftigten liegt sie stark unter dem Durchschnitt.

Außerdem wurde der Erfüllungsgrad nach Branchen ausgewertet. Die Unternehmen konnten sich zu Beginn des Online-Fragebogens selbst bis zu 3 von 30 Branchen zuordnen. Um genügend Datenpunkte für eine Auswertung zu erhalten, wurden diese 30 Branchen in 10 Branchen geclustert (siehe Anhang). Das Branchencluster „Nahrungsmittel, Holzprodukte, Textilien und Leder“ weist einen stark überdurchschnittlichen Erfüllungsgrad auf, die Branchencluster „Finanzdienstleistungen“, „Transport und Logistik“ und „Chemie, Pharmazie, Kunststoffe und Bergbau“ haben ein durchschnittliches Erfüllungsniveau und das Branchencluster „Automobil und Fahrzeuge“ schneidet leicht unterdurchschnittlich ab. Alle anderen Branchencluster – „Maschinenbau und Metallindustrie“, „Elektronik, Telekommunikation und Digitales“, „Dienstleistungen“ (nicht deckungsgleich mit „Handel und Dienstleistungen“ bei den Schichten, siehe oben), „Bau und Immobilien“ und „Energie, Ver- und Entsorgung“ – haben einen stark unterdurchschnittlichen Erfüllungsgrad.

Daneben konnten sich die Unternehmen zu Beginn des Online-Fragebogens einer oder mehrerer von sieben Wertschöpfungsstufen zuordnen, in denen sie tätig sind. In der Analyse bestätigte sich das oben gezeigte Bild. Unternehmen im Vertrieb/Handel haben einen überdurchschnittlichen Erfüllungsgrad, während Unternehmen im Bereich Kreditvergabe/Finanzierung/Versicherung durchschnittlich abschneiden. Die Wertschöpfungsstufen Herstellung von Komponenten/Zwischenprodukten, Herstellung von Endprodukten und Dienstleistungen weisen ein unterdurchschnittliches Erfüllungsniveau auf. Für die Wertschöpfungsstufen Rohstoffgewinnung und Abfallbehandlung/Recycling liegen nicht genügend Datenpunkte für eine Aussage vor.

Die vertiefte Analyse zeigt einen Zusammenhang zwischen der Teilnahme an Brancheninitiativen und der NAP-Erfüllung. Die Gruppe der Unternehmen, die angaben, an einer Brancheninitiative teilzunehmen, um tatsächlich und potenziell nachteiligen Auswirkungen auf Menschenrechte zu begegnen, erreicht einen überdurchschnittlichen Anteil in den Kategorien „Erfüller“ und „Unternehmen auf gutem Weg“. Auch ihre Erfüllung der Kernelemente liegt über den Werten der Nettostichprobe. Dieser Vorsprung war in beiden repräsentativen Phasen beobachtbar. Allerdings zeigt die Auswertung, dass selbst in dieser Gruppe der Anteil der Erfüller unter 50 Prozent liegt. Hervorzuheben ist außerdem, dass die Ergebnisse keine kausalen Zusammenhänge darstellen, sondern lediglich Korrelationen aufzeigen. Der Beitritt zu einer Brancheninitiative kann sowohl Ausgangspunkt für die Beschäftigung eines Unternehmens mit seiner menschenrechtlichen Sorgfaltspflicht als auch deren Ergebnis sein.

Bereits in der explorativen Phase 2018 brachten die Unternehmen und Stakeholder zum Ausdruck, dass die Teilnahme an Brancheninitiativen einen wichtigen Beitrag zur Umsetzung der menschenrechtlichen Sorgfaltspflicht leiste, aber nicht notwendigerweise ausreiche, um alle Anforderungen des NAP zu erfüllen. Standards und Brancheninitiativen wurden als hilfreich für die Informationsfindung und zum Austausch unter den Mitgliedern bezeichnet. Darüber hinaus leisteten sie Unterstützung beim Einstieg in das Management von Lieferanten – auch über die direkte Lieferkette hinaus – und bei der Auswahl, Kontrolle und Schulung der GeschäftspartnerInnen durch Standardsetzung und Verifizierungsangebote. Jedoch wurde kritisiert, dass es zu viele unterschiedliche Standards und Initiativen gebe, die der Komplexität eines Unternehmens nicht genügend Rechnung trügen oder nur einen Ausschnitt an Menschenrechtsthemen berücksichtigten.<sup>24</sup>

---

<sup>24</sup> Siehe Zwischenbericht 2018, S. 32 f.

### 3 Reflexion des Monitorings

Die Anforderungen an das Monitoring waren in der Leistungsbeschreibung<sup>25</sup> zum Vergabeverfahren des Monitorings festgehalten. Das Konsortium hat im Rahmen regelmäßiger Berichte jeweils das grundlegende Konzept, Details des methodischen Vorgehens und schließlich die Darstellung der Ergebnisse erstellt. Der IMA hat die Berichte jeweils im Konsens abgenommen. Die Methodik des NAP-Monitorings wurde als lernendes System angelegt. Sie wurde durch die jeweiligen Erkenntnisse aus den drei Monitoring-Phasen iterativ weiterentwickelt.

#### 3.1 Reflexion des Monitoring-Prozesses

Kapitel III des NAP und die Leistungsbeschreibung bildeten die Grundlage des Monitoring-Prozesses. Im NAP sind die in Kapitel 1.2 genannten Ziele des Monitorings beschrieben und in der Leistungsbeschreibung wird das methodische Vorgehen u. a. mit den drei Erhebungsphasen, dem Stufenverfahren, dem „Comply or explain“-Ansatz und der Repräsentativität skizziert. Diese Vorgaben waren über den gesamten Monitorings-Prozess maßgeblich für das Vorgehen des Konsortiums. Mit den Ergebnissen der repräsentativen Erhebung 2020 und ihrer Darstellung im Zwischenbericht 2020 wurde ermittelt, inwieweit die Unternehmen in Deutschland mit mehr als 500 Beschäftigten die Anforderungen des NAP umgesetzt haben. Dabei wurden die relevanten Standards wissenschaftlichen Arbeitens berücksichtigt, sodass die ermittelten Ergebnisse objektiv und belastbar sind. Daneben wurde auch das Ziel erreicht, über den Umsetzungsstand des NAP hinaus weitere quantitative und qualitative Informationen zur inhaltlichen Tiefe und Herausforderungen bei der Umsetzung der Elemente menschenrechtlicher Sorgfalt zu ermitteln.

Für die Erstellung der Berichte war das Konsortium unter Federführung von EY zuständig. Die Berichtsentwürfe wurden zunächst mit dem Auswärtigen Amt als Auftraggeberin besprochen. Das Ergebnis der Beratung wurde anschließend dem IMA und der AG Wirtschaft und Menschenrechte zur Kommentierung vorgelegt. Das Konsortium hat dem IMA und der AG zudem alle Berichtsentwürfe in einer Gremiensitzung in einem Vortrag mit begleitendem Foliensatz mündlich präsentiert und Fragen beantwortet. Die schriftlichen Kommentare wurden vom Konsortium je Bericht in einem Kommentierungsraster festgehalten und bei Bedarf mit Anmerkungen versehen. Die Kommentierungsraster und die auf der Basis der Kommentare angepassten Berichte wurden in Abstimmung mit dem Auswärtigen Amt wieder dem IMA und der AG Wirtschaft und Menschenrechte zugänglich gemacht.<sup>26</sup> Je nach Komplexität und Inhalt der Berichte waren mehrere Kommentierungsrunden erforderlich. Bei Bedarf wurden einzelne Detailfragen im IMA nachträglich erörtert. Der Anforderungsrahmen für die repräsentativen Erhebungen wurde lediglich im IMA erörtert und der AG Wirtschaft und Menschenrechte nicht vorgelegt, um keine systematische Verzerrung der Antworten durch Bekanntgabe des Anforderungsrahmens zu begünstigen. Der Anforderungsrahmen wurde nach Abschluss des Monitorings als Anlage des Zwischenberichts 2020 veröffentlicht. Sämtliche Berichte, der Anforderungsrahmen und die Leistungsbeschreibung wurden vom Auswärtigen Amt nach konsensueller Zustimmung aller IMA-Ressorts freigegeben.

---

<sup>25</sup> Ausschreibung „Monitoring NAP“ vom 29. Dezember 2017, Referenznummer der Bekanntmachung: VV-118-2017-0201 durch das Auswärtige Amt

<sup>26</sup> Auf dieses Vorgehen wurde beim Zwischenbericht 2019 verzichtet, da es im IMA vorab keinen Konsens bzgl. des Vorschlags des Auswärtigen Amtes gab, der AG schon vor dem finalen IMA-Freigabebeschluss des Berichts einen Entwurf zur Kommentierung vorzulegen.



Der Inception Report beschrieb das methodische Vorgehen bei der explorativen Erhebung 2018 und bot damit eine wichtige Grundlage für die Vorbereitung der repräsentativen Erhebungsphasen. Der Inception Report enthielt außerdem eine theoretische Beschreibung des angedachten Bewertungssystems für die repräsentativen Erhebungen. Wie in den Kapiteln 3.2 und 3.3 beschrieben, wurde das Bewertungssystem im Zuge des Zwischenberichts 2018 grundlegend überarbeitet. Es lässt sich daher feststellen, dass es ratsam gewesen wäre, im Inception Report einen Fokus auf das Vorgehen in der explorativen Erhebungsphase und die Entwicklung des Fragebogens zu legen.

Der Zwischenbericht 2018 zu den Ergebnissen der explorativen Erhebung enthielt mit dem Fragebogen und dem detaillierten Bewertungssystem die methodischen Kernstücke des Monitorings. Die Herausforderung bestand darin, ein möglichst detailliertes methodisches Vorgehen aufzuzeigen, das dem IMA als Entscheidungsgrundlage dienen konnte. Da das Monitoring weltweit das erste seiner Art war, gab es noch keine umfassenden Erfahrungswerte, die das Konsortium berücksichtigen konnte. Gleichzeitig mussten insbesondere im Bewertungssystem die verschiedenen Perspektiven der im IMA vertretenen Ressorts berücksichtigt werden. Die notwendige Konsensbildung im IMA zum Zwischenbericht 2018 und damit zur Methodik konnte nicht im ursprünglich festgelegten Zeitrahmen erreicht werden. Dadurch verspätete sich der Start der Erhebung 2019 um ca. drei Monate, wodurch sich der gesamte Zeitplan bis zur Vorlage der Ergebnisse der Erhebung 2020 im Sommer 2020 – wie ursprünglich geplant – entsprechend verdichtete. Dies hatte zur Folge, dass die Einbindung externer Stakeholder wie z. B. die Kommentierung des Zwischenberichts 2019 durch die AG Wirtschaft und Menschenrechte teilweise nicht wie im vorgesehenen Maße stattfinden konnte.

Die Ergebnisse der beiden repräsentativen Erhebungen 2019 und 2020 wurden in den jeweiligen Zwischenberichten festgehalten. Wie in Kapitel 2.1 beschrieben, sind die Ergebnisse in ihren Größenordnungen ähnlich. Dadurch decken sich auch die Zwischenberichte inhaltlich in weiten Teilen. Dies war jedoch notwendig, da jedem Zwischenbericht eine eigene, unabhängige Erhebung zugrunde lag, sodass der jeweilige Zwischenbericht für externe LeserInnen in sich schlüssig sein und ohne viele Verweise auskommen sollte. Aufgrund der Auswirkungen der COVID-19-Pandemie wurde der Erhebungszeitraum 2020 um einen Monat verlängert. Dementsprechend wurden auch die Vorlage und die Abnahme des Zwischenberichts 2020 verschoben. Als Entscheidungsgrundlage für den weiteren politischen Prozess wurde dem Auswärtigen Amt bereits im Vorfeld am 14. Juli 2020 eine Ergebnisindikation der repräsentativen Erhebung 2020 mitgeteilt.

Die Zwischenberichte enthielten neben den Gesamtergebnissen auch Detailergebnisse zu den einzelnen Kernelementen sowie zum Einfluss der jeweiligen Methodikbestandteile „Comply or explain“-Mechanismus, Freitexte und Widersprüche auf das Ergebnis. Das bietet eine höhere Transparenz für die beteiligten Ressorts und Stakeholder.

Insgesamt haben sämtliche während des Monitorings erstellten Berichte ihre in der Leistungsbeschreibung jeweils vorgesehene Funktion erfüllt. Der Zeitplan wurde am Ende trotz der politischen und pandemiebedingten Verlängerung von Diskussions- und Erhebungsphasen eingehalten, sodass die Bundesregierung wie geplant im Sommer 2020 mit der Diskussion von Folgemaßnahmen beginnen konnte.

### **Kontaktierungsverfahren in den repräsentativen Erhebungsphasen**

In der ursprünglichen Planung der repräsentativen Erhebungen des Monitorings war eine Kontaktaufnahme mit den zuständigen Personen der Unternehmen der Stichprobe geplant. Hierfür sollten sowohl offizielle Kontakte wie beispielsweise die öffentlich einsehbaren AnsprechpartnerInnen der Einkaufs- oder Nachhaltigkeitsabteilungen als auch dem Konsortium

bekannte AnsprechpartnerInnen in den Unternehmen genutzt werden. Aus datenschutzrechtlichen Gründen war dies nicht möglich. Stattdessen war eine Kontaktierung über allgemeine, öffentliche E-Mail-Adressen und die schriftliche Kontaktierung der Geschäftsführung notwendig. Der Ablauf dieses Vorgehens wurde in der dritten Monitoring-Phase auf der Basis von Erfahrungen der vorherigen Phase nachgebessert:

Die Unternehmen der Stichprobe 2019 wurden zuerst per E-Mail über allgemein verfügbare Adressen kontaktiert. Erst zu einem späteren Zeitpunkt wurde eine postalische Kontaktaufnahme an die Geschäftsführung nachgeschickt, die zu einer höheren Teilnahme geführt hat. In der Erhebung 2020 wurde daher zu Beginn neben der E-Mail-Ansprache direkt eine postalische Kontaktaufnahme an die Geschäftsleitung versandt. Dies hat zu einer ausreichenden Teilnahme trotz der allgemeinen Sondersituation angesichts der Corona-Pandemie geführt.

### 3.2 Reflexion der Erhebungsmethodik

Nach Auftragserteilung wurde basierend auf dem Angebotskonzept eine erste Methodik zur Erhebung ausgearbeitet. Als Basis für das methodische Vorgehen diente die Leistungsbeschreibung aus den Ausschreibungsunterlagen zum NAP-Monitoring. Die Methodik wurde durch die Abnahme des Inception Report sowie der nachfolgenden Zwischenberichte 2018 und 2020 durch den IMA im Konsens angenommen.

Anschließend wurden in der explorativen Phase Fragebogen und Anforderungsrahmen ausgearbeitet und mit dem IMA und den Stakeholdern der AG Wirtschaft und Menschenrechte diskutiert. Die Erkenntnisse aus dieser Phase flossen in die Ausgestaltung der Befragung 2019 ein. Letztere war methodisch so robust, dass für die Erhebung 2020 nur geringfügige Anpassungen notwendig waren.

Die wesentlichen Grundpfeiler hatten in der gesamten Projektzeit Bestand, wie auch die Festlegung auf eine Umsetzung, die eine ausreichende Aussagekraft mit einem vertretbaren Aufwand für alle Beteiligten verbindet. Nachfolgend werden kritische Aspekte der Methode dargestellt.

#### 3.2.1 Definition der Grundgesamtheit

Die Grundgesamtheit war in der Leistungsbeschreibung zur Vergabe des Monitorings definiert als Größe der Ergebnistabellen aus dem statistischen Unternehmensregister „Unternehmen nach Wirtschaftsabschnitten mit 500 und mehr sozialversicherungspflichtig Beschäftigten“ mit Stand zum Zeitpunkt der Abnahme des jeweiligen Zwischenberichts. Da die Rohdaten des Statistischen Bundesamtes aufgrund des Statistikgeheimnisses weder dem Konsortium noch dem Auftraggeber zur Verfügung gestellt wurden, musste für die Ermittlung der Grundgesamtheit eine kommerzielle Unternehmensdatenbank genutzt werden. Die dafür vom Konsortium im Inception Report zunächst ins Auge gefasste Amadeus-Datenbank stellte sich als ungeeignet heraus, da darin die vom NAP erfassten Unternehmen, insbesondere aus dem Finanzsektor, nicht vollständig abgebildet waren. Für beide repräsentative Erhebungen wurde daher die Unternehmensdatenbank Bisnode genutzt, die alle notwendigen Anforderungen erfüllte.

Bei der Bestimmung der Anzahl der Beschäftigten der Unternehmen wurde über alle Erhebungsphasen hinweg auf ArbeitnehmerInnen im Sinne des § 267 Abs. 5 HGB abgestellt. Die Anzahl der Beschäftigten gilt jeweils für die einzelne rechtliche Entität bzw. bei Konzernen mit Sitz in Deutschland gilt somit die weltweite (konsolidierte) Anzahl der MitarbeiterInnen.

In Deutschland gibt es keine Legaldefinition und auch keine andere einheitliche Definition für die Institution „Unternehmen“. Für den Zweck des Monitorings hat das Erhebungsteam eine dem Geist des NAP folgende Definition des Unternehmensbegriffs herausgearbeitet: „Unternehmen der Grundgesamtheit des NAP-Monitorings werden als solche Unternehmen definiert, die dem erwerbswirtschaftlichen Prinzip folgen.“ Um die Unternehmen zur Bestimmung der Grundgesamtheit filtern zu können, wurde im ersten Schritt die Rechtsform als Indiz für das erwerbswirtschaftliche Prinzip genutzt und im zweiten Schritt eine manuelle Prüfung bestimmter Unternehmen vorgenommen. Auf diese Weise konnten rein gemeinnützige Unternehmen herausgefiltert werden.<sup>27</sup>

### 3.2.2 Bewertungskategorien

Die ursprünglich geplante Einteilung in zwei Ergebniskategorien („Erfüller“ und „Nicht-Erfüller“) hat sich in der explorativen Phase als zu wenig differenziert herausgestellt. Es wurde deutlich, dass sich verschiedene Stakeholder eine Ergebnisdefinition wünschen, die es erlaubt, Unternehmen, die bei der Umsetzung schon weiter fortgeschritten sind, von jenen zu unterscheiden, die noch nichts oder sehr wenig umgesetzt haben. Daher wurde mit der Abnahme des Zwischenberichts 2018 durch den IMA die Gruppe der Nicht-Erfüller konkretisiert und durch die Untergruppe der „Unternehmen auf einem guten Weg“ ergänzt. Mit Blick auf die eher vage NAP-Vorgabe des Umsetzungshorizonts „bis 2020“ wurde zudem eine neue Ergebniskategorie „Unternehmen mit Umsetzungsplan“ geschaffen. Mit dieser Anpassung konnte in den Zwischenberichten 2019 und 2020 aufgezeigt werden, dass sich das Gesamtergebnis – nämlich weniger als 50 Prozent der betreffenden Unternehmen erfüllen den NAP – nicht ändert, wenn diejenigen Unternehmen gesondert ausgewiesen werden, die die NAP-Vorgaben zum Zeitpunkt der Erhebung noch nicht vollständig umgesetzt hatten, aber über einen hinreichenden Umsetzungsplan bis Ende 2020 verfügen oder auf „einem gutem Weg“ der Umsetzung sind. Im Lichte des sehr kleinen und für das Ergebnis nicht relevanten Umfangs der Gruppe „Unternehmen mit Umsetzungsplan“ hat der Auftraggeber beschlossen, dass eine gesonderte Nachvalidierung der im dritten Zwischenbericht benannten „Unternehmen mit Umsetzungsplan“ zu Anfang 2021 unverhältnismäßig wäre.

### 3.2.3 Fragebogen und Antwortmöglichkeiten

Im Inception Report war vorgesehen, dass jede Antwortmöglichkeit ein bestimmtes Umsetzungsniveau von „0 – nicht vorhanden“ bis „5 – vollumfänglich“ pro Frage/Kriterium darstellt. Diese Logik hat sich basierend auf den Ergebnissen der explorativen Phase als nicht geeignet erwiesen und wurde bei der Erstellung des Fragebogens verworfen. So wurden etwa bei einigen Fragen zum einen weniger als fünf Antwortmöglichkeiten definiert, etwa auch rein binäre Möglichkeiten wie „Ja, vorhanden“ und „Nein, nicht vorhanden“. Dies war für bestimmte Sachverhalte eine passendere Bewertungsgrundlage als das ursprünglich 5-stufige System. Zum anderen wurden auch teilweise mehrere gleichwertige Antwortmöglichkeiten definiert, die zur Erfüllung der Frage geführt haben. Dies hat sich als notwendig herausgestellt, um verschiedene Möglichkeiten der Ausgestaltung der Anforderungen abbilden zu können. Diese Veränderungen haben sich als sinnvoll erwiesen, um Unternehmen eine klar verständliche, passende und adäquate Antwortmöglichkeit zu geben. Auch die Nutzung von Freitexten hat dazu beigetragen, die unternehmensspezifischen Aspekte abbilden und dennoch eine zielgerichtete Bewertung einer adäquaten Umsetzung vornehmen zu können.

---

<sup>27</sup> Siehe Zwischenbericht 2018, Kapitel 6.2.1.

Die Erhebungen 2019 und 2020 wurden im Wesentlichen mit dem in der explorativen Phase entwickelten Fragebogen und Anforderungsrahmen durchgeführt. Das Grundgerüst hat sich als funktional, sehr anwendungsfähig und aussagekräftig erwiesen. Zwischen der Erhebung 2019 und der Erhebung 2020 wurden ausschließlich geringe Anpassungen im Fragebogen vorgenommen, die der Verbesserung der Verständlichkeit dienen sollten. Methodisch sind beide Erhebungen aber weitgehend identisch.

### 3.2.4 Anforderungsrahmen

Vor Beginn der Monitoring-Phasen war methodisch vorgesehen, verschiedene vordefinierte Erfüllungsmuster basierend auf der Entwicklung eines systematischen Anforderungsrahmens zu definieren. Diese sollten sich nach objektiven Unternehmensmerkmalen (z. B. Größe, Sektor etc.) richten und in Anforderungsprofilen mit spezifischem Umsetzungsniveau pro Kernelement zusammengefasst werden („Bewertungscluster“). Aufgrund der sehr variantenreichen Merkmalskombinationen von Unternehmen und ihrer spezifischen Eigenschaften im Hinblick auf potenzielle menschenrechtliche Auswirkungen wurde dieser Ansatz auf der Basis der Ergebnisse der explorativen Phase weiterentwickelt.

Eine erste Weiterentwicklung war die Definition der unternehmensindividuellen Risikodisposition als zentrales Element des menschenrechtlichen Sorgfaltsprozesses des Unternehmens. So wurde das Bewertungssystem im Hinblick auf die Einbeziehung der Ergebnisse der Risikoanalyse (Risikodisposition) aus Kernelement 2 des NAP überarbeitet. Die Ergebnisse der Risikoanalyse des antwortenden Unternehmens entscheiden folglich über die weitere Ausgestaltung der menschenrechtlichen Sorgfaltspflicht hinsichtlich Maßnahmen und deren Wirksamkeitskontrolle.

Ergänzend wurden bei vielen Fragen unterschiedliche Umsetzungswege der Unternehmen berücksichtigt. So spiegelten verschiedene zur Erfüllung führende Antwortmöglichkeiten die Bandbreite der unternehmerischen Umsetzung.

Eine Veränderung zu der im Inception Report vorgestellten Methodik war auch, dass bei manchen Fragen die Anforderungen für alle Unternehmen identisch definiert wurden, da keine Abhängigkeit von variierenden Unternehmenseigenschaften wie Größe oder Branche als relevant für die Erfüllung erachtet werden kann. So ist in Kernelement 1 die Existenz einer Grundsatzerklärung für alle Unternehmen die gleiche Anforderung. Bei anderen Fragen unterscheiden sich die Anforderungen unternehmensspezifisch nach der konkreten Risikodisposition des Unternehmens – also dem Vorhandensein und die Höhe bestimmter Menschenrechtsrisiken bzw. bestimmter betroffener Gruppen.

Dieser Anforderungsrahmen, der die individuelle Risikodisposition der Unternehmen als eine Bewertungsgrundlage verwendet und somit berücksichtigt, dass die konkrete Ausgestaltung einer menschenrechtlichen Sorgfaltsprüfung immer unternehmensspezifisch erfolgen muss, hat sich während der ersten repräsentativen Erhebung bewährt und wurde auch in der zweiten Erhebung angewandt.

### 3.2.5 Stufenverfahren

Bereits der Inception Report sah ein vierstufiges Verfahren für die Prüfung der Selbstauskünfte der Unternehmen vor. Dieses hat sich grundsätzlich bewährt und wurde in beiden repräsentativen Phasen angewandt.

Auf Stufe 1 erfolgte die Online-Befragung, die durch eine Medienanalyse und eine Recherche branchentypischer Menschenrechtsthemen zur Plausibilisierung der Angaben ergänzt wurde. Auf Stufe 2 wurden die Angaben auf Widersprüche hin untersucht. Solche Widersprüche konnten sich beispielsweise durch Abweichungen zwischen verschiedenen Antworten innerhalb der Selbstauskunft oder zwischen der Selbstauskunft und öffentlich verfügbaren Informationen ergeben. Bei vorliegenden Widersprüchen und einer vom Unternehmen angegebenen Kontaktmöglichkeit wurden die Unternehmen auf Stufe 3 kontaktiert, um die Widersprüche aufzuklären. Es war vorgesehen, dass daraus verbleibende Widersprüche auf Stufe 4 durch die Befragung weiterer Stakeholder endgültig aufgelöst werden. Da sich sämtliche bewertungsrelevanten Widersprüche auf Stufe 3 im Gespräch mit den Unternehmen abschließend bearbeiten ließen, kam Stufe 4 nicht zur Anwendung.

### 3.2.6 Prüfung der „Comply or explain“- und Freitext-Antworten

Der „Comply or explain“-Mechanismus war bereits in der Leistungsbeschreibung definiert. Er sollte den Unternehmen erlauben zu begründen, warum sie bestimmte, im NAP vorgesehene Verfahren nicht oder in anderer Form umgesetzt haben. Darüber hinaus enthielt der Fragebogen mehrere Freitextoptionen, sodass Unternehmen über die vorgegebenen Antworten hinausgehende Angaben machen konnten. Unternehmen wurden auf der Ebene jeder einzelnen bewertungsrelevanten Frage aufgefordert, eine „Explain“-Antwort zu ergänzen, wenn die gewählte vorformulierte Antwortoption nicht dem Anforderungsrahmen entsprach.

In der explorativen Phase wurden fünf Prüfprinzipien für Antworten bei „Comply or explain“ und in Freitexten definiert: Vollständigkeit und Bezüglichkeit (formelle Prüfprinzipien) sowie Konkretheit, funktionale Äquivalenz und Kontexteinbettung (inhaltliche Prüfprinzipien). Diese wurden als Leitprinzipien verwendet, um die Bewertung der „Comply or explain“- und Freitext-Angaben zu erleichtern und zu vereinheitlichen. Alle „Comply or explain“- und Freitextangaben wurden hinsichtlich der formellen und inhaltlichen Dimensionen einzeln geprüft und bewertet. Für „Explain“-Antworten von Unternehmen, deren Bewertung sich nicht direkt aus dem Anforderungsrahmen ableiten lassen konnten, wurde zunächst eine Einschätzung innerhalb des jeweiligen Beurteilungsteams definiert und dann im Konsortium besprochen. Die Ergebnisse dieser Erörterungen flossen in eine interne Prüfungsanleitung ein. Bei diesen Bewertungen wurde auch das Prinzip der Angemessenheit zur Berücksichtigung kontextabhängiger Faktoren umgesetzt. Hier wurde geprüft, inwieweit Informationen zu Branche, Größe und Position in der Liefer- und Wertschöpfungskette zur Bewertung einer angemessenen Umsetzung der Anforderungen dienen können. Als Beispiel: Unternehmen ohne wesentliche Investitionstätigkeiten müssen die Perspektive der Investitionen nicht in der Risikoanalyse berücksichtigen, um die Anforderungen zu erfüllen.

Dieses Verfahren hat sich über beide Befragungen hinweg bewährt. Zum einen hat in beiden repräsentativen Erhebungen ein Großteil der Unternehmen von der Möglichkeit Gebrauch gemacht, „Comply or explain“- und Freitext-Angaben einzutragen; zum anderen konnte über das formelle und inhaltliche Prüfverfahren mit der Nutzung und Weiterentwicklung der Prüfanleitung eine personenunabhängige und einheitliche Bewertung sichergestellt werden.

### 3.3 Reflexion der NAP-Anforderungen und des Anforderungsrahmens

Ein zentraler Aspekt der Erhebung und der dazugehörigen Bewertung ist der Anforderungsrahmen. Der Anforderungsrahmen bestimmt, welche Antworten im Fragebogen am Ende als eine

angemessene Umsetzung der menschenrechtlichen Sorgfalt zu bewerten sind. Hierbei galt es, einen guten Kompromiss zwischen effizienter Bewertung und der intensiveren Betrachtung von spezifischen und unternehmensindividuellen Umständen zu ermöglichen.

Ausgehend vom Anspruch des Monitorings, den engen zeitlichen Vorgaben und dem Anspruch, Unternehmen nicht durch eine Überkomplexität des Fragebogens von der freiwilligen Teilnahme abzuschrecken und zugleich besonderen Unternehmenskonstellationen und -kontexten gerecht zu werden, war die methodische Kombination aus einem effizient ausfüllbaren Online-Fragebogen mit vordefinierten Antworten auf der einen Seite und der Einzelfallprüfung durch die Nutzung von Freitext und „Comply or explain“-Angaben (siehe Kapitel 3.2.6) auf der anderen Seite angemessen. Die Auseinandersetzung mit den Angaben der Unternehmen zeigte deutlich, dass ein reiner Bewertungsautomatismus ohne qualitative Einzelfallprüfung nicht funktioniert. Daneben wurde die Bewertung durch die Integration der spezifischen Risikodisposition des Unternehmens in den Anforderungsrahmen weiter individualisiert.

Bereits bei der Erstellung des Fragebogens und des Anforderungsrahmens ist auffällig geworden, dass der derzeitige NAP in seinem Kapitel III weniger konkrete Orientierung und Leitlinie für die Unternehmen bietet als die umfangreicheren VN-Leitprinzipien. Auch inhaltlich sind die NAP-Anforderungen und die VN-Leitprinzipien nicht immer deckungsgleich. So sind die Anforderungen des NAP bezüglich der Grundsatzklärung vergleichsweise ausführlich und detailliert, während an anderen Stellen, beispielsweise zur Berichterstattung oder zum Beschwerdemechanismus, die Anforderungen deutlich weniger spezifisch formuliert sind. Dies war eine zentrale Herausforderung bei der Entwicklung des Anforderungsrahmens. Maßgeblich für die Ausgestaltung des vom IMA verabschiedeten Anforderungsrahmens waren ausschließlich die Anforderungen des NAP. Zu betonen ist daher, dass eine Erfüllung des NAP-Anforderungsrahmens nicht zwingend gleichzusetzen ist mit einer Aussage zum Umsetzungsgrad der VN-Leitprinzipien.

Auch die Rückmeldungen der Unternehmen deuten darauf hin (vgl. Kapitel 2.1), dass die Übersetzung der allgemein gehaltenen Anforderungen in konkrete Handlungsvorgaben und Kriterien in der vielgestaltigen Unternehmenspraxis mitunter schwierig ist. Eine Unsicherheit besteht hinsichtlich der Erwartungshaltung an den genauen Umfang der menschenrechtlichen Sorgfaltspflicht. So ist den Unternehmen mehrheitlich nicht ausreichend bewusst, dass der NAP vorgibt, alle für die Geschäftsaktivitäten relevanten Stufen der Wertschöpfungskette des Unternehmens in die menschenrechtliche Risikoanalyse einzubeziehen, um anhand des Ergebnisses risikobasierte Maßnahmen zur Minderung von negativen menschenrechtlichen Auswirkungen identifizieren zu können. Dabei zeigen sich insbesondere Schwierigkeiten bei der Interpretation der Angemessenheitskriterien im NAP, also der Größe, Branche und Position in der Wertschöpfungskette des Unternehmens. Der NAP erwähnt diese lediglich allgemein, ohne sie näher auszuführen. Dieser Aspekt betrifft insbesondere die Grenzen der Verantwortung und Einflussnahme in globalen und komplexen Wertschöpfungsketten. Hier kommt erschwerend hinzu, dass der NAP die für die VN-Leitprinzipien zentrale, abgestufte Verursachungslogik von „cause“, „contribute“ und „directly linked“ in ihrer Bedeutung für die Ermittlung der Verantwortungstiefe kaum diskutiert bzw. sogar abweichend von den VN-Leitprinzipien benennt („indirekte Verbundenheit“). Der Anforderungsrahmen hat den Unternehmen daher an dieser Stelle über den „Comply or explain“-Mechanismus und Freitextoptionen viel Raum zur individuellen Darstellung ihres Vorgehens bei Risikoanalyse und Maßnahmen und der Begründung dieses Vorgehens gegeben. Auf diese Weise konnte bei der Bewertung berücksichtigt werden, inwiefern die Unternehmen die genannten Angemessenheitskriterien bei der Ausgestaltung der menschenrechtlichen Sorgfalt und der Priorisierung ihrer Maßnahmen reflektiert hatten.

Der Gestaltung des Anforderungsrahmens kam in der Methodik des Monitorings eine zentrale Bedeutung zu, da dieser den Erfüllerstatus der einzelnen Unternehmen und somit die Erfüllerquote der teilnehmenden Unternehmen insgesamt bestimmte. Alles in allem war der Definitionsprozess des Anforderungsrahmens durch divergierende Positionen der beteiligten Stakeholder geprägt. Da der NAP selbst keine differenzierte Abstufung der Anforderungen definiert, war es methodisch geboten, sich bei der Ausgestaltung des Anforderungsrahmens eng an die Formulierungen in Kapitel 3 des NAP zu halten. Auf diese Art und Weise konnte eine angemessene und vom IMA bestätigte Bewertungsgrundlage entwickelt werden, die von den involvierten Stakeholdern weiterhin kritisch begleitet wurde, aber aufgrund der Nähe zum Wortlaut des NAP gerechtfertigt war. Bei solchen Fragen, bei denen der NAP keine ausreichende Handhabe bot, um für die Antworten einen soliden Anforderungsrahmen zu hinterlegen, wurde folgerichtig auf die Definition einer Anforderung verzichtet. Die Antworten auf diese Fragen waren also nicht erfüllungsrelevant und dienten der Gewinnung qualitativer Zusatzinformationen.

### 3.4 Limitationen der Methodik

Verschiedene Aspekte erwiesen sich als herausfordernd. Insbesondere sind die ausreichende Teilnahme und das Erreichen einer für die Grundgesamtheit repräsentativen Netto-Stichprobe zu nennen. Die Beantwortung der Fragen war für Unternehmen freiwillig und das Kontaktmanagement musste konsequent die datenschutzrechtlichen Vorgaben berücksichtigen. Dazu kamen das aus der Freiwilligkeit der Erhebungen abgeleitete Anonymitätsgebot gegenüber der Bundesregierung und der Öffentlichkeit sowie das Risiko einer noch stärkeren Selbstselektion bei der Teilnahme, wenn die teilnehmenden Unternehmen und ihre Erfüllungsbewertung nicht anonym behandelt worden wären. So mussten viele Unternehmen angeschrieben und aufwendige Kontaktschritte unternommen werden, um eine ausreichend hohe Nettostichprobe in den jeweiligen repräsentativen Erhebungsphasen zu erreichen. Dadurch kam es fast zwangsläufig zur Selbstselektion der teilnehmenden Unternehmen. Die daraus resultierende Ergebnisverzerrung konnte durch statistische Methoden im Hinblick auf beobachtbare Faktoren ausgeglichen werden, nicht jedoch im Hinblick auf unbeobachtbare Faktoren.

Im Vordergrund des Monitorings stand die Frage: Handelt es sich bei dem Unternehmen um einen Erfüller oder nicht? Die angestrebte Effizienz der Befragung in Form eines Online-Fragebogens ist ein Vorteil zur schnellen Generierung von Erhebungsergebnissen, führt jedoch zu einer geringeren Informationstiefe in den Erhebungsergebnissen als beispielsweise direkte Interviews. So ermöglicht die Erhebungsform der Selbstauskunft eine statistisch zulässige und für den Zweck des NAP-Monitorings effektive Erhebung, aber nur zu einem geringeren Teil eine Generierung detaillierter und vollständig überprüfter Informationen.

Eine qualitative Plausibilisierung der Angaben wurde ausschließlich anhand öffentlicher Quellen und der Identifikation von Logikfehlern innerhalb der Antworten durchgeführt. Eine abschließende Überprüfung und Validierung der Sorgfaltsprozesse der Unternehmen war in der Methodik des Monitorings nicht vorgesehen, da dies erweiterte Beleg- oder gar Vor-Ort-Prüfungen bei den Unternehmen notwendig gemacht und den Zeit- und Ressourcenaufwand vervielfacht hätte.

Eine weitere Herausforderung war die eingeschränkte Kontaktmöglichkeit mit Unternehmen zur Durchführung dieser Plausibilisierung. Im Sinne des Datenschutzes bat der Online-Fragebogen die Teilnehmer um Zustimmung zu einer Kontaktierung durch das Erhebungsteam im Falle von Nachfragen. Eine gewisse Anzahl Unternehmen verweigerte diese Zustimmung. Bei diesen Unternehmen war ein Auflösen von Widersprüchen nicht möglich.

Bei nicht persönlich oder gar anonym durchgeführten Befragungen ist es zudem möglich, dass die Angaben des einzelnen Unternehmens davon abhängig sind, wer den Fragebogen bearbeitet hat. Dabei können sich fachliches Vorwissen, die Verfügbarkeit von Informationen über Unternehmensinterna und subjektive Positionen auf das Antwortverhalten auswirken. Um solche Effekte abzumildern, wurde den teilnehmenden Unternehmen explizit empfohlen, verschiedene Akteure im Unternehmen in die Beantwortung des Fragebogens einzubeziehen.

Abschließend lässt sich festhalten, dass trotz der dargestellten Herausforderungen das Monitoring zielgenau und verlässlich feststellen konnte, dass weniger als 50 Prozent der deutschen Unternehmen mit mehr als 500 Beschäftigten die NAP-Anforderungen umgesetzt haben. Das Monitoring war somit eine umfassende, konsequente und wirksame Überprüfung der Umsetzung von Kapitel III des NAP und konnte in aggregierter Form auch besondere Schwachpunkte der Erfüllung in Bezug auf einzelne Kernelemente aufzeigen. Dessen ungeachtet ist auch festzuhalten, dass das NAP-Monitoring weniger aussagekräftig in Bezug auf die Frage ist, warum diese Schwächen bestehen. Die Herausforderungen der einzelnen Unternehmen lassen sich nur skizzieren, da die Methodik des NAP-Monitorings nicht darauf ausgelegt ist, eine abschließende individuelle Erfüllungsprüfung bei einzelnen Unternehmen im Hinblick auf Anforderungen der VN-Leitprinzipien oder einer möglichen gesetzlichen Regulierung an die menschenrechtlichen Sorgfaltspflicht durchzuführen. Diese müsste zu diesem Zweck stärker in Richtung einer qualitativen Einzelfallprüfung mit ergänzenden Erhebungsinstrumenten wie Interviews und einer verstärkten Dokumentenprüfung weiterentwickelt werden.

#### **Erkenntnisse des Helpdesk Wirtschaft & Menschenrechte der Bundesregierung**

Im Rahmen der beiden repräsentativen Erhebungen des NAP-Monitorings 2019 und 2020 war der Helpdesk Wirtschaft & Menschenrechte der Bundesregierung offizieller Ansprechpartner für Unternehmen bei inhaltlichen Rückfragen zum NAP-Monitoring sowie zum Fragebogen. UnternehmensvertreterInnen konnten sich telefonisch und per E-Mail an die BeraterInnen des Helpdesk wenden.

Während der beiden Erhebungsphasen 2019 und 2020 haben jeweils ca. 140–145 Unternehmen das Beratungsangebot des Helpdesk Wirtschaft & Menschenrechte genutzt. Bei der Erhebung 2020 gab es mit rund 200 Anfragen trotz Herausforderungen aufgrund der COVID-19-Pandemie eine leicht erhöhte Anfragenanzahl im Vergleich zu 2019 (ca. 180 Anfragen). Die Erkenntnisse aus den Aktivitäten des Helpdesk Wirtschaft & Menschenrechte im Rahmen des Monitorings werden in einem eigenen Kurzbericht des Helpdesk veröffentlicht.

Dieser ist ab dem 19. Oktober 2020 auf der [Webseite](#) der Agentur für Wirtschaft & Entwicklung verfügbar.

Der Helpdesk Wirtschaft & Menschenrechte der Bundesregierung ist ein kostenfreies Unterstützungsangebot für Unternehmen, situiert in der Agentur für Wirtschaft & Entwicklung und finanziert vom Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. Der Helpdesk agiert unabhängig vom NAP-Monitoring-Konsortium und unter Wahrung des Datenschutzes.



## Anhang: Methodik der vertieften Analyse

Die vertiefte Analyse in Kapitel 2.3 dient dazu, das Datenmaterial der repräsentativen Erhebung 2019 und 2020 nach Schichten, Branchen und Wertschöpfungsstufen, denen die Unternehmen zugeordnet sind, auszuwerten. Die Schichten ergeben sich aus dem Prozess der Stichprobenziehung, bei dem jedes Unternehmen gemäß dem in der Bisnode-Datenbank hinterlegten Wirtschaftszweig (produzierendes Gewerbe, Dienstleistungen) und der Anzahl seiner Beschäftigten (501–1.000, 1.001–2.000, > 2.000) einer von sechs Schichten zugeordnet wird. Die Zuordnung zu den Branchen und Wertschöpfungsstufen ergab sich aus den Angaben der Unternehmen im Online-Fragebogen. Hier konnten die Unternehmen bis zu drei Branchen sowie mehrere von sieben Wertschöpfungsstufen auswählen, in denen sie tätig sind. Dementsprechend ist es möglich, dass ein Unternehmen in das Analyseergebnis mehrerer Branchen und/oder Wertschöpfungsstufen einfließt. Für die Branchen gab es 30 Auswahloptionen, inkl. einer Option „Weitere“ mit Freitext. Für die vertiefte Analyse wurden diese Freitexteinträge bei Unternehmen, die keine andere Branche ausgewählt hatten, manuell überprüft und das Unternehmen – falls möglich – einer der anderen Branche zugeordnet. Außerdem wurden die 30 Branchen aus dem Online-Fragebogen in zehn Branchencluster zusammengefasst, um genügend Datenpunkte für die vertiefte Analyse zu generieren.

Übersicht über die Branchencluster<sup>28</sup>:

Branchencluster	Branche(n) gemäß Online-Fragebogen
Automobil und Fahrzeuge	<ul style="list-style-type: none"> <li>• Automobil</li> </ul>
Finanzdienstleistung	<ul style="list-style-type: none"> <li>• Finanzdienstleistung</li> </ul>
Maschinenbau und Metallindustrie	<ul style="list-style-type: none"> <li>• Maschinenbau</li> <li>• Metallerzeugung und -bearbeitung</li> <li>• Sicherheit und Verteidigung</li> </ul>
Elektronik, Telekommunikation und Digitales	<ul style="list-style-type: none"> <li>• Elektronik</li> <li>• Telekommunikation, Software und Digitales</li> </ul>
Nahrungsmittel, Holzprodukte, Textilien und Leder	<ul style="list-style-type: none"> <li>• Nahrungsmittel, Getränke, Tabak und Landwirtschaft</li> <li>• Holz-, Papierwaren und Forstwirtschaft</li> <li>• Textilien und Leder</li> </ul>
Transport und Logistik	<ul style="list-style-type: none"> <li>• Transport und Logistik</li> </ul>
Dienstleistung	<ul style="list-style-type: none"> <li>• Gesundheit und Pflege</li> <li>• Personal-, Reinigungs- und Sicherheitsdienstleistungen</li> <li>• Werbung und Marketing</li> <li>• Beratung, Wirtschaftsprüfung und Kanzleien</li> <li>• Gastronomie und Beherbergung</li> <li>• Reise, Freizeit und Unterhaltung</li> <li>• Bildung</li> </ul>
Chemie, Pharmazie, Kunststoffe und Bergbau	<ul style="list-style-type: none"> <li>• Chemie</li> <li>• Pharmazie und Biotechnologie</li> <li>• Gummi- und Kunststoffwaren</li> <li>• Bergbau und Gewinnung von Steinen, Erden, Erdöl und Erdgas</li> </ul>
Bau und Immobilien	<ul style="list-style-type: none"> <li>• Baugewerbe</li> <li>• Immobilien</li> </ul>
Energie, Ver- und Entsorgung	<ul style="list-style-type: none"> <li>• Wasserversorgung</li> <li>• Entsorgung</li> <li>• Energieversorgung</li> </ul>

<sup>28</sup> Zu beachten ist, dass die Branchencluster nicht den Schichten der Stichprobenziehung entsprechen.

Für die vertiefte Analyse wurden die Daten der Erhebungen 2019 und 2020 miteinander kombiniert. Dies war notwendig, weil eine Unterteilung der Nettostichproben nach Schichten, Branchen und Wertschöpfungsstufen dazu führt, dass sich in den jeweiligen Gruppen viel weniger Unternehmen befinden als in den gesamten Nettostichproben. **Trotz der Zusammenführung der Daten befinden sich in den jeweiligen Gruppen nicht genügend Unternehmen, um für diese eine repräsentative Aussage treffen zu können.** Aus diesem Grund werden für diese Gruppen lediglich qualitative, vergleichende Tendenzen aufgezeigt.

Bei Zusammenführung der Daten aus den Jahren 2019 und 2020 wurden zunächst die Unternehmen identifiziert, die an beiden Erhebungen teilgenommen hatten. Für diese Unternehmen wurden lediglich die Daten von 2020 berücksichtigt. Anschließend wurden die kombinierten Daten im Hinblick auf den Erfüllungsgrad insgesamt und denjenigen der einzelnen Kernelemente ausgewertet. Dabei wurde ein statistisches Gewicht verwendet, um die in beiden Erhebungen vorliegenden Verzerrungen aufgrund beobachtbarer Faktoren auszugleichen. Dabei handelt es sich um das Gewichtungungsverfahren 4 (siehe Anhang 2 im Zwischenbericht 2020), da dieses am konservativsten ist. Das Gewicht berechnete sich neu aus den kombinierten Datensätzen 2019 und 2020.

Die in Kapitel 2.3 dieses Berichts aufgezeigten vergleichenden Tendenzen beziehen sich auf das 95-Prozent-Konfidenzintervall des jeweiligen Erfüllungsgrads (z. B. Erfüller insgesamt). Somit ist immer die Grundgesamtheit der Unternehmen Maßstab für den Vergleich. Gruppen (z. B. eine einzelne Branche), deren Erfüller-Mittelwert innerhalb dieses Konfidenzintervalls liegt, werden als „durchschnittlich“ definiert. Liegt der Mittelwert über dem Konfidenzintervall aller Unternehmen, wird die Gruppe als „überdurchschnittlich“ definiert, liegt er mehr als 5 Prozentpunkte darüber, als „stark überdurchschnittlich“. Umgekehrt sind Gruppen, deren Mittelwert unter dem Konfidenzintervall liegt, als „unterdurchschnittlich“ definiert bzw. als „stark unterdurchschnittlich“, wenn die Differenz mehr als 5 Prozentpunkte beträgt.