



# Eine Konstruktvalidierung des Work Ability Index anhand einer repräsentativen Stichprobe von Erwerbstätigen in Deutschland

baua: Bericht

**Forschung  
Projekt F 2250**

M. Freyer

**Eine Konstruktvalidierung des  
Work Ability Index anhand einer  
repräsentativen Stichprobe von  
Erwerbstätigen in Deutschland**

1. Auflage 2019  
Dortmund/Berlin/Dresden

Diese Veröffentlichung ist 2016 als Masterarbeit zur Erlangung des akademischen Grades Master of Science (M.Sc.) im Master-Studiengang Psychologie der Fakultät KSW an der FernUniversität in Hagen im Arbeitsbereich Arbeits- und Organisationspsychologie entstanden. Betreut wurde die Arbeit durch Prof. Bernd Marcus (Universität Rostock, ehemals FernUniversität Hagen) und Dr. Uwe Rose (BAuA). Die Bearbeitung der Fragestellungen erfolgte im Rahmen des Forschungsprojektes F 2250 „Entwicklung einer empirischen Basis zur Arbeitsfähigkeit und funktionellen Gesundheit – Repräsentativerhebung an Erwerbstätigen“ der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. Das Forschungsprojekt beruht auf Stichproben aus Statistiken der Bundesagentur für Arbeit, die im Institut für Arbeitsmarkt- und Berufsforschung (IAB) in den Integrierten Erwerbsbiografien zusammengeführt wurden. Ein Zugang zu den Forschungsdaten als Scientific Use File ist beim Forschungsdatenzentrum des IAB möglich. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin.

Autorin: Marion Freyer  
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Projektleitung: Dipl. Soz. Eberhard Pech (†)  
Dr. Uwe Rose  
Dr. Gabriele Freude  
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Titelfoto: alvarez/iStock.com

Umschlaggestaltung: Susanne Graul  
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Herausgeber: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA)  
Friedrich-Henkel-Weg 1 – 25, 44149 Dortmund  
Postanschrift: Postfach 17 02 02, 44061 Dortmund  
Telefon 0231 9071-2071  
Telefax 0231 9071-2070  
E-Mail [info-zentrum@buaa.bund.de](mailto:info-zentrum@buaa.bund.de)  
Internet [www.buaa.de](http://www.buaa.de)

Berlin: Nöldnerstraße 40 – 42, 10317 Berlin  
Telefon 030 51548-0  
Telefax 030 51548-4170

Dresden: Fabricestraße 8, 01099 Dresden  
Telefon 0351 5639-50  
Telefax 0351 5639-5210

Die Inhalte der Publikation wurden mit größter Sorgfalt erstellt und entsprechen dem aktuellen Stand der Wissenschaft. Für die Richtigkeit, Vollständigkeit und Aktualität der Inhalte übernimmt die BAuA jedoch keine Gewähr.

Nachdruck und sonstige Wiedergabe sowie Veröffentlichung, auch auszugsweise, nur mit vorheriger Zustimmung der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.



doi:10.21934/buaa:bericht20190529 (online)

[www.buaa.de/dok/8818760](http://www.buaa.de/dok/8818760)

# Inhaltsverzeichnis

	Seite
<b>Kurzreferat</b>	<b>5</b>
<b>Abstract</b>	<b>6</b>
<b>1 Einleitung</b>	<b>7</b>
<b>2 Theorie</b>	<b>9</b>
2.1 Der Begriff der Arbeit und ihre Bedingungen	9
2.2 Die Entwicklung des Konzepts der Arbeitsfähigkeit	11
2.2.1 Funktionsfähigkeit	11
2.2.2 Arbeitsfähigkeit	13
2.2.3 Das Haus der Arbeitsfähigkeit	15
2.3 Operationalisierung von Arbeits- und Funktionsfähigkeit	17
2.4 Work Ability Index (WAI)	19
2.5 Empirischer Forschungsstand und Herleitung der Fragestellungen	21
<b>3 Methode</b>	<b>27</b>
3.1 Untersuchungsdesign und Durchführung	27
3.2 Messinstrumente	29
3.2.1 Work Ability Index (WAI)	29
3.2.2 Short-Form-12 (SF-12) Health Survey	32
3.2.3 Norwegian Function Assessment Scale (NFAS)	33
3.3 Stichprobenbeschreibung	34
3.4 Datenauswertung und statistische Methoden	35
3.4.1 Prüfung der Voraussetzungen	35
3.4.2 Konfirmatorische Faktorenanalysen	36
3.4.3 Multiple Gruppenvergleiche	38
3.4.4 Gewichtung der Items	39
3.4.5 Konvergenzvalidierung	40
<b>4 Ergebnisse</b>	<b>42</b>
4.1 Faktorielle Validität des WAI	42
4.2 Messinvarianz des WAI	47
4.2.1 Geschlecht	47
4.2.2 Altersgruppen	48
4.3 Gewichtete Summenindizes und Normwerttabellen	50
4.4 Konvergenzvalidität des WAI	57
<b>5 Diskussion</b>	<b>59</b>
5.1 Ergebnisdiskussion	59
5.1.1 Dimensionalität des WAI	59
5.1.2 Zulässigkeit von Mittelwertvergleichen	62
5.1.3 Itemgewichtung	63
5.1.4 Zusammenhang mit konstrukt-nahen Instrumenten	64
5.2 Kritische Methodenwürdigung	66
5.3 Fazit und Ausblick	67

<b>Literaturverzeichnis</b>	<b>69</b>
<b>Abkürzungsverzeichnis</b>	<b>78</b>
<b>Anhang 1 Work Ability Index</b>	<b>79</b>
<b>Anhang 2 Soziodemografische Daten der Stichprobe</b>	<b>83</b>

# **Eine Konstruktvalidierung des Work Ability Index anhand einer repräsentativen Stichprobe von Erwerbstätigen in Deutschland**

## **Kurzreferat**

Der Work Ability Index (WAI) ist ein etabliertes Instrument zur Erfassung der Arbeitsfähigkeit, welches unter der umstrittenen Annahme der Eindimensionalität als ungewichteter Summenindex entwickelt wurde. Die Ziele dieser Arbeit waren die Konstruktvalidierung des WAI, unter Berücksichtigung der bisher nicht untersuchten Messinvarianz verschiedener Subgruppen, sowie die gewichtete Berechnung und Validierung individueller Indexwerte. Die Daten entstammen einer repräsentativen Stichprobe von Erwerbstätigen in Deutschland (N = 3 968). Anhand konfirmatorischer Faktorenanalysen am WAI wurden dessen Dimensionalität und die Messinvarianz für das Geschlecht und 3 Altersgruppen untersucht. Durch Korrelationsanalysen mit dem Short-Form-12 Health Survey und der Norwegian Function Assessment Scale erfolgte die Konvergenzvalidierung. Die Ergebnisse stützen die Annahme einer zweifaktoriellen Struktur des WAI mit skalarer Messinvarianz der untersuchten Subgruppen. Die Gewichtung der Summenindizes erfolgte aus den Ergebnissen zur Faktorenstruktur und konvergente Validität der untersuchten Instrumente wurde gefunden. Somit wird die Arbeitsfähigkeit im WAI durch die beiden Faktoren der subjektiven Arbeitsfähigkeit und Ressourcen sowie der Gesundheitsbedingungen abgebildet.

## **Schlagwörter:**

Arbeitsfähigkeit, Work Ability Index, Faktorenstruktur, Messinvarianz, gewichteter Summenindex

# **The construct validity of the Work Ability Index based on a representative sample of workers in Germany**

## **Abstract**

The Work Ability Index (WAI) is an established instrument for assessing work ability. It was developed under the controversial assumption of one-dimensionality as an unweighted sum score. The objectives of this work are the construct validation of the WAI, taking into consideration the until now unexamined measurement invariance for various subgroups, and the weighted scoring and validation of individual values. The data come from a representative sample of the working population in Germany (N = 3,968). Using confirmatory factor analysis, dimensionality and measurement invariance of the WAI were examined for sex and 3 age groups. Convergence validity was investigated by means of correlation analysis using the Short-Form-12 Health Survey and the Norwegian Function Assessment Scale. The results support an assumption of two-dimensionality along with scalar measurement invariance for the tested subgroups. The weighted sum scores were derived from the results on the factorial structure and convergent validity for the examined instruments was found. Thus, work ability in the WAI is represented by the factors of subjective work ability and resources, as well as the health conditions.

## **Key words:**

work ability, Work Ability Index, factor structure, measurement invariance, weighted sum score

# 1 Einleitung

In den letzten Jahrzehnten hat das Konzept der Arbeitsfähigkeit und dessen Operationalisierung zunehmend Beachtung gefunden (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin [BAuA], 2013; Ilmarinen, J. & Tuomi, 2004; Nordenfelt, 2008). Die Relevanz der Thematik für den betrieblichen und gesundheitspolitischen Kontext ergibt sich dabei aus dem sozialpolitischen Ziel zum Erhalt der Teilhabe am Erwerbsleben (§1 Sozialgesetzbuch III). Die Arbeitsfähigkeit, als das Vermögen eines Menschen seine Tätigkeit am Arbeitsplatz auszuführen, ist abhängig von seinen physischen und psychischen Ressourcen, welche wiederum im Wechselspiel mit den Anforderungen bei der Arbeit und der Umwelt stehen (Ilmarinen, J., Gould, Järvikoski & Järvisalo, 2008). Von besonderem Interesse sind hierbei gezielt veränderbare Einflussfaktoren auf die Arbeitsfähigkeit, durch die das individuelle Wohlbefinden und die Lebensqualität sowie die betriebliche Produktivität gefördert werden können (Ilmarinen, J., 2011). Dabei sollte der Ansatzpunkt für Intervention und Prävention nicht erst bei älteren Beschäftigten liegen, sondern aufgrund einer ebenfalls vorliegenden Variation der Arbeitsfähigkeit bei jüngeren Erwerbstätigen frühzeitig ansetzen (Freude & Pech, 2005; Ilmarinen, J., 2009).

Die Aufrechterhaltung und Förderung der Arbeitsfähigkeit für den Erhalt der Teilhabe am Erwerbsleben steht allerdings vor verschiedenen Herausforderungen, welche vor dem Hintergrund des demografischen Wandels und der Veränderungen in der Arbeitswelt entstanden sind (Freude & Pech, 2005).

So hat sich laut Statistischem Bundesamt (2015) die Lebenserwartung der (arbeitenden) Bevölkerung bei Geburt seit Ende des 19. Jahrhunderts ungefähr verdoppelt und wird aufgrund der Entwicklungen im Gesundheitswesen voraussichtlich weiter steigen, während die Geburtenrate stark zurückgegangen ist und derzeit stagniert. Unter anderem aus diesen Gründen stieg der Anteil der Erwerbstätigen zwischen 60 bis 64 Jahren in den letzten 10 Jahren (Stand 2014) von 28 auf 52%. Der Anteil der 65- bis 69-Jährigen in der Erwerbstätigkeit hat sich sogar mit einer Zunahme von 6 auf 14% im gleichen Zeitraum mehr als verdoppelt (Statistisches Bundesamt, 2015). Durch die stufenweise Anhebung seit 2012 liegt das gesetzliche Renteneintrittsalter ab dem Geburtsjahrgang 1964 bei derzeit 67 Jahren. Aufgrund dieser Verlängerung der Lebensarbeitszeit ist eine Zunahme des Durchschnittsalters der Erwerbstätigen zu erwarten. Der Altersanstieg der Erwerbstätigen geht allerdings mit einem Anstieg der krankheitsbedingten Fehltag und Leistungseinbußen einher (Badura, Walter & Hehlmann, 2010). Die drei häufigsten Diagnosegruppen für den Anteil an Arbeitsunfähigkeitstagen der Bevölkerung sind dabei Krankheiten des Muskel-Skelettsystems, psychische Störungen und Verhaltensstörungen sowie Krankheiten des Atmungssystems (Knieps & Pfaff, 2015; Rebscher, 2016). Insbesondere für die psychischen Störungen ist sowohl generell als auch in Abhängigkeit mit dem Alter ein Anstieg der Arbeitsunfähigkeitstage und -fälle in den letzten 10 Jahren zu verzeichnen (Rebscher, 2016). Infolge von Arbeitsunfähigkeiten entstanden im Jahr 2013 volkswirtschaftliche Kosten in Höhe von 59 Milliarden Euro durch Lohnkosten und 103 Milliarden Euro durch den Verlust von Arbeitsproduktivität (Nöllenheidt & Brenscheidt, 2015). Die durch Arbeitsunfähigkeit und Frührente gestiegenen Kosten für das Gesundheits- und Sozialversicherungssystem werden auf lange Sicht aufgrund des demografischen Wandels immer weniger durch Sozialabgaben aufgefangen, wodurch weitere Reformen in den Systemen notwendig und zu erwarten sind (Kerschbaumer & Schroeder, 2005).



Des Weiteren hat sich die Arbeitstätigkeit selbst in den letzten Jahrzehnten aufgrund der Digitalisierung, Globalisierung, Entgrenzung und atypischer Beschäftigungsformen mit hoher Geschwindigkeit in eine neue Richtung entwickelt (Bundesministerium für Arbeit und Soziales, 2015). Diese Veränderungen bieten der Gesellschaft Möglichkeiten der persönlichen und organisationalen Entfaltung, stellen sie gleichzeitig vor Herausforderungen und beinhalten Risiken. Die psychischen Anforderungen an Erwerbstätige steigen sowohl qualitativ als auch quantitativ (Badura et al., 2010). Beschäftigte müssen immer schneller auf wechselnde und simultane Anforderungen reagieren und sind aufgrund der technischen Vernetzung oft auch außerhalb der Arbeitszeit erreichbar, wodurch eine Trennung von Arbeit und Freizeit erschwert wird. Diese und andere Faktoren der Arbeitsbedingungen können einen Einfluss auf die Arbeitsfähigkeit haben (Seitsamo, Tuomi, Ilmarinen, J. & Gould, 2008).

Vor dem Hintergrund des demografischen Wandels sowie dem Primat einer humanen und gesunden Arbeit ist es somit von hoher Relevanz, die Arbeitsfähigkeit von Erwerbstätigen frühzeitig zu überprüfen, durch präventive Maßnahmen aufrechtzuerhalten und durch Interventionen zu fördern (Freude & Pech, 2005). Zur Erfassung der Arbeitsfähigkeit wird häufig der Work Ability Index (WAI) von Tuomi, Ilmarinen, J., Jahkola, Katajarinne und Tulkki (1998) herangezogen. Im WAI wird die Arbeitsfähigkeit als multidimensionales Konstrukt aufgefasst, welches individuelle Arbeitsbedingungen, mentale Ressourcen und Gesundheitsbedingungen berücksichtigt. Das Instrument selbst wurde jedoch unter der Annahme der Eindimensionalität als ungewichteter Summenindex entwickelt. Zu den psychometrischen Eigenschaften dieses Verfahrens liegen bislang heterogene Ergebnisse vor, welche eine Interpretation des gemessenen Konstrukts der Arbeitsfähigkeit erschweren. So konnte die Annahme der einfaktoriellen Struktur des WAI in verschiedenen Studien nicht gestützt werden. Diese Befundlage wird im Weiteren noch ausgeführt und bildet die primäre Motivation zur Durchführung der vorliegenden Untersuchung. Die Hauptziele der Arbeit liegen in der Konstruktvalidierung des WAI unter Berücksichtigung seiner Faktorenstruktur, Messinvarianz und konvergenten Validität. Desweiteren sollen eine Empfehlung für eine gewichtete Berechnung der individuellen Indexwerte sowie Referenzwerte des WAI präsentiert werden.

Die folgenden Kapitel beginnen mit einer theoriebezogenen Einführung in die Thematik der Arbeitsfähigkeit und der möglichen Operationalisierung des Konstrukts. Anhand der Diskussion bisheriger Befunde zu den psychometrischen Eigenschaften des WAI erfolgt die Herleitung der Fragestellungen. Anschließend werden die eingesetzten empirischen Methoden und die statistischen Analyseverfahren beschrieben. Nach Darstellung der Ergebnisse zur Überprüfung der Hypothesen erfolgt eine kritische Diskussion und Interpretation der Befunde.

## 2 Theorie

Um sich ernähren zu können, ein Dach über dem Kopf und Kleidung am Körper zu haben und seinen persönlichen Lebensstandard in Form von Hobbies, Reisen und Freizeitaktivitäten aufrechterhalten zu können, sind für die Mehrheit der Menschen ausreichend finanzielle Mittel notwendig. Diese Mittel werden in der Regel in Form von Lohn oder Gehalt durch Nachgehen einer Erwerbsarbeit erworben, welche somit eine Lebensgrundlage bildet. Bei Verschlechterung der Arbeitsfähigkeit, Ausfall durch Arbeitsunfähigkeit und einem potentiell anschließenden Verlust des Arbeitsplatzes verschlechtert sich mit hoher Wahrscheinlichkeit die finanzielle Lage sowohl für die Person als auch für die Wirtschaft. Neben der Existenzsicherung hat die Arbeit auch eine weitere Bedeutung für den Menschen. So werden Aktivität und Kompetenz, Zeitstrukturierung, Kooperation und Kontakt, soziale Anerkennung sowie persönliche Identität als psychosoziale Funktionen der Arbeit genannt (Semmer & Meier, 2014).

In den folgenden Abschnitten werden, ausgehend vom Begriff der Arbeit, verschiedene Formen der Erwerbstätigkeit vorgestellt. Anschließend erfolgt eine Einführung zum Konzept der Arbeitsfähigkeit und dessen Operationalisierung. Dabei wird insbesondere die Entwicklung des WAI von Tuomi et al. (1998) vertiefend vorgestellt, um daraus folgend die Fragestellungen und Hypothesen vor dem Hintergrund des aktuellen Forschungsstandes zum WAI herzuleiten.

### 2.1 Der Begriff der Arbeit und ihre Bedingungen

Wie Voß (2010) feststellt, hat der umstrittene Versuch einer streng umrissenen Definition des Begriffs Arbeit eine lange Historie und unterliegt einem stetigen Wandel. In der Enzyklopädie der Philosophie und Wissenschaftstheorie (2004) wird zwischen einem allgemeinen und einem engen Arbeitsbegriff unterschieden. Im allgemeinen Sinne wird die Arbeit als eine rein zweckgebundene produzierende Tätigkeit verstanden, ohne dem arbeitenden Menschen Möglichkeiten zur Selbstverwirklichung zu bieten. Die Selbstverwirklichung fände demnach lediglich in der arbeitsfreien Zeit statt, was zu einer Entfremdung von der Arbeit führe. Selbstverwirklichung und Arbeit werden somit als Gegenspieler betrachtet: Der zunehmende Wunsch nach Selbstverwirklichung zöge eine Verringerung der Arbeit nach sich. Erst unter Beachtung einer humanistischen Konzeption der Arbeit wird die Selbstverwirklichung ebenfalls als Ziel der Arbeit angesehen. Für den Arbeitsbegriff im engeren Sinne werden innerhalb der Zweckdienlichkeit einer Tätigkeit zusätzlich negativ behaftete Aspekte der Mühsal bzw. Unerwünschtheit herangezogen, wodurch eine Beeinträchtigung der Bedürfnisse der Arbeitenden durch die Arbeit stattfindet. Für die vorliegende Arbeit wird der Begriff der Erwerbsarbeit verstanden als eine auf das Einkommen und „auf ein wirtschaftliches oder organisationales Ziel gerichtete planmäßige menschliche Tätigkeit, bei der sowohl körperliche als auch geistige Kräfte eingesetzt werden“ (Nerdinger, Blickle & Schaper, 2014, S. 6). Im Zuge dieser Tätigkeiten werden zum einen Produkte oder Dienstleistungen im Interesse des Arbeitgebers oder Kunden erstellt. Zum anderen werden persönliche materielle und psychische Bedürfnisse, z. B. Selbstverwirklichung, Teilhabe und Weiterentwicklung, befriedigt. Zu beachten ist der wechselseitige Einfluss von Person, Tätigkeit und Umwelt. Nicht nur die Umwelt wird durch Arbeitstätigkeiten von Personen verändert, die veränderte Umwelt

prägt auch den Menschen selbst (Nerdinger et al., 2014). Die Sichtweise der humanen Arbeitsgestaltung darf daher nicht außer Acht gelassen werden:

Human ist eine Arbeit dann, wenn sie menschengerecht und menschenwürdig ausgeführt werden kann und damit die physische und psychische Gesundheit weder kurz [sic] noch langfristig beeinträchtigt. Sie sollte an die Bedürfnisse und Qualifikationen des arbeitenden Menschen angepasst sein und eine Entwicklung der Persönlichkeit sowie eine Entfaltung von Potentialen und Kompetenzen fördern. Ziel ist, Arbeit sicher, gesund und effektiv zu gestalten. (Gesellschaft für Arbeitswissenschaft e. V., 2016)

Abzugrenzen ist diese Definition der Arbeit somit z. B. von häuslicher oder künstlerischer Arbeit, die nicht dem Zweck der Einkommenssicherung dient. In hochindustrialisierten Ländern wie Deutschland ist in den letzten Jahrzehnten ein Strukturwandel in der Arbeitswelt zu verzeichnen. So entwickelt sich die Gesellschaft weg von einer produzierenden materiellen hin zu einer Dienstleistungs- und Wissensgesellschaft mit abnehmenden körperlichen und zunehmenden psychischen Anforderungen an die Erwerbstätigen (Badura et al., 2010). Erwerbsarbeit unterliegt verschiedenen Beschäftigungs- und Arbeitsbedingungen, die jedoch in der Literatur nicht streng voneinander abgegrenzt werden. Nachfolgend wird eine mögliche Einteilung der Bedingungen dargestellt.

Beschäftigungsbedingungen betreffen Inhalte des bestehenden Arbeitsvertrags bezogen auf Arbeits- bzw. Beschäftigungsverhältnisse, Arbeitszeit, Urlaubsregelungen und Lohngestaltung. Beschäftigungsverhältnisse können unterteilt werden in das sogenannte Normalarbeitsverhältnis, ausgehend von einer unbefristeten Vollzeitbeschäftigung, sowie in atypische Beschäftigungsverhältnisse wie befristete oder Leiharbeitsverhältnisse, Teilzeitarbeit, geringfügige Beschäftigung und Selbstständigkeit (Hoffmann & Walwei, 1998). Letztere haben in der Vergangenheit zunehmend an Bedeutung gewonnen. Diese Veränderungen in der Verbreitung der Erwerbsformen werden von Eichhorst (2015) der zunehmenden externen Flexibilität zugeschrieben. So sank zwischen 1992 und 2014 der Anteil der Erwerbstätigen im Normalarbeitsverhältnis von 45 auf 40%, während der Anteil der atypischen Beschäftigung von 20 auf über 30% gestiegen ist (Eichhorst, Tobsch & Wehner, 2016). Bei der geschlechtergetrennten Betrachtung der Arbeitsverhältnisse ist erkennbar, dass der Anteil der atypischen Arbeitsverhältnisse den der normalen Arbeitsverhältnisse bei den Frauen übersteigt, während die Männer deutlich weniger in atypischen Arbeitsverhältnissen tätig sind. Bezogen auf das Alter zeigt sich insbesondere für die jüngeren Erwerbstätigen zwischen 15 und 24 Jahren im untersuchten Zeitraum eine Verlagerung der Beschäftigung von knapp 18 auf rund 36% hin zu atypischen Beschäftigungsverhältnissen. Als Gründe für die Veränderungen in den Erwerbsformen nennt Eichhorst (2015) den Wandel hin zu einer Dienstleistungsgesellschaft, den Wettbewerbsdruck durch die fortschreitende Globalisierung, Reformen in der Politik ebenso wie den steigenden Frauenanteil auf dem Arbeitsmarkt. Auch innerhalb der Normalarbeitsverhältnisse haben Entwicklungen stattgefunden, die Eichhorst (2015) der internen Flexibilität zuschreibt. Darunter fallen z. B. flexible Arbeitszeitregelungen und -muster wie Gleitzeit, Schicht- und Wochenendarbeit sowie die Vermischung von Arbeitszeit und Freizeit durch Telearbeit und digitale Vernetzung.

Arbeitsbedingungen betreffen die Inhalte und Abläufe der Arbeitsaufgabe an sich ebenso wie Umgebungsfaktoren am Arbeitsplatz, Arbeitsmittel und soziale Beziehungen zu Kollegen oder Führungskräften. Zu ersteren sind beispielhaft die Quantität

und Qualität, die Sinnhaftigkeit und Abwechslung der Arbeitsaufgabe zu nennen. Umgebungsfaktoren am Arbeitsplatz können sowohl physische Eigenschaften des Raumes sein, wie z. B. Licht, Lärm und Größe, aber auch das ausreichende Vorhandensein notwendiger Arbeitsmittel. Von hoher Bedeutung sind auch die sozialen Beziehungen zu Kollegen und Führungskräften. Das kann die Teamarbeit betreffen, die Kommunikation untereinander, den Führungsstil und Rückmeldungen aber auch Unterstützung durch Kollegen und Fairness.

Um nun Erwerbstätige vor dem Hintergrund der älter werdenden Gesellschaft langfristig und gesund in der Arbeitsumgebung zu behalten, ist eine Vermeidung von Arbeitsunfähigkeit unter Beachtung der Arbeitsfähigkeit notwendig. Die Entwicklungen in der Arbeitsunfähigkeit der Erwerbsgesellschaft in Deutschland wurden bereits in der Einleitung angesprochen. Im folgenden Kapitel soll nun das Konzept der Arbeitsfähigkeit vorgestellt werden, wobei auch eine Abgrenzung von der Arbeitsunfähigkeit und der Beschäftigungsfähigkeit erfolgt.

## **2.2 Die Entwicklung des Konzepts der Arbeitsfähigkeit**

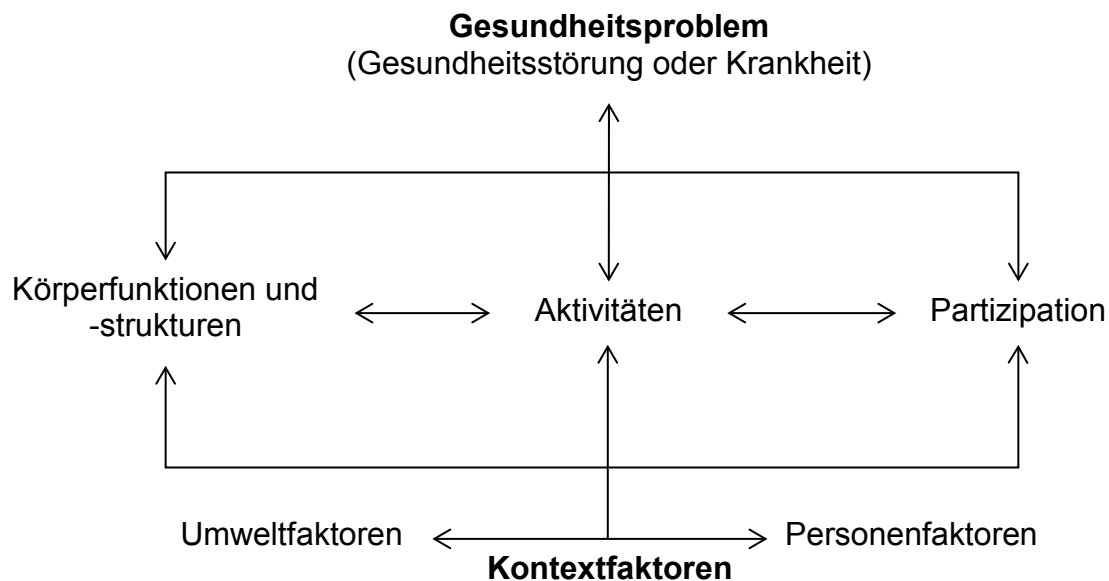
Das Konzept der Arbeitsfähigkeit wird in der wissenschaftlichen Forschung u. a. vor dem Hintergrund der Einflussfaktoren auf und der Konsequenzen aus der mehr oder weniger vorhandenen Arbeitsfähigkeit betrachtet. Im betrieblichen Kontext werden die Ergebnisse der Forschung im Rahmen des Betrieblichen Gesundheitsmanagements (BGM) für den Erhalt oder die Förderung der Arbeitsfähigkeit umgesetzt. Lederer, Loisel, Rivard und Champagne (2014) sowie Tengland (2011) geben in ihren jeweiligen Arbeiten einen Überblick über verschiedene Charakteristika und Definitionen von Arbeitsfähigkeit. Sie kommen zu dem Schluss, dass das Konzept der Arbeitsfähigkeit bis dahin nicht zufriedenstellend und übereinstimmend definiert wurde. Der Begriff der Arbeitsfähigkeit unterliegt dabei einem Wandel von einer rein medizinischen Sicht hin zu einer multidimensionalen Betrachtung (Ilmarinen, J. et al., 2008). So erfolgt die begriffliche Bestimmung in diesem Kapitel entlang des Begriffs der Funktionsfähigkeit bis hin zur Arbeitsfähigkeit.

### **2.2.1 Funktionsfähigkeit**

Die Leistungs- bzw. Funktionsfähigkeit wird in der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) definiert als funktionale Gesundheit, welche zur Bewältigung des Alltages einer Person notwendig ist (World Health Organization [WHO], 2001). Sie umfasst die Konzepte der Körperfunktionen und -strukturen, der Aktivitäten (Handlungen) und der Partizipation an Lebensbereichen (Schuntermann, 2005). Die Körperstrukturen und körperlichen Funktionen einer Person, welche ebenfalls geistige und seelische Funktionen beinhalten, sollten allgemein anerkannten Normen entsprechen. Die Person sollte all jene Aktivitäten in allen subjektiv wichtigen Lebensbereichen vollziehen können, die jemandem ohne gesundheitliche Beeinträchtigung zuzumuten sind. Die funktionale Gesundheit ist nach der ICF eine Funktion eines vorliegenden Gesundheitsproblems und variabler Kontextfaktoren, welche in Umwelt- und Personenfaktoren unterschieden werden können. Umweltfaktoren umfassen dabei u. a. Produkte und Technologien, soziale Unterstützung und Beziehungen sowie Einstellungen, Werte und Überzeugungen anderer Personen und der Gesellschaft. Personenfaktoren können z. B. das Alter und Geschlecht, der Charakter oder die Erfahrung sein. Zusammen bilden all diese Aspekte das in der Abbildung 2.1 dargestellte bio-psycho-soziale Modell der ICF,

welches die Wechselwirkung des Gesundheitsproblems und der Kontextfaktoren auf die Konzepte der Körperfunktionen und -strukturen, der Aktivitäten und der Teilhabe an Lebensbereichen beschreibt (Schuntermann, 2005).

Jeder einzelne Aspekt des Modells kann somit einen (negativen) Einfluss auf alle anderen Aspekte haben, was Sekundärprozesse nach sich ziehen kann. Ein Sportunfall einer Person (Körperfunktion) kann z. B. zu einer chronischen Erkrankung (Gesundheitsproblem) führen, welche die Teilhabe am Erwerbsleben und die Teilnahme an Aktivitäten beeinträchtigt.



**Abb. 2.1** Das bio-psycho-soziale Modell der Komponenten der Gesundheit (nach Schuntermann, 2005).

Zu den Lebensbereichen der Aktivitäten und Partizipation, also der menschlichen Daseinsentfaltung, gehören Bedeutende Lebensbereiche als eines der Hauptkapitel im ICF, welche u. a. auch die Arbeit und Beschäftigung beinhalten. Um Handlungen durchführen zu können, muss eine Person generell und objektiv leistungsfähig sein. Im Erwerbsleben setzt sich die Leistungsfähigkeit aus vier Teilkonstrukten zusammen: körperliche, geistige, seelische und soziale Leistungsfähigkeit (Schuntermann, 2005). Dabei sollten ebenfalls Kontextfaktoren betrachtet werden, welche gemäß der Handlungstheorie von Nordenfelt (2008) sowohl Gegebenheiten (äußere Möglichkeiten zur Handlung) als auch die Handlungsbereitschaft der Person umfassen. Hergeleitet aus der ICF wäre eine Person demnach dann arbeitsfähig, wenn sie unter Berücksichtigung des Kontextes körperlich, geistig, seelisch und sozial dazu in der Lage und dazu bereit ist, ihrer Arbeitstätigkeit nachzugehen und somit am Erwerbsleben teilzunehmen.

Kritisch anzumerken ist, dass die ICF nur gesundheitsbezogene Beeinträchtigungen behandelt. Der Kontext wird zwar grundsätzlich mit einbezogen, jedoch werden Beeinträchtigungen, welche ausschließlich durch Faktoren außerhalb der Gesundheit ihren Ursprung haben, hier nicht betrachtet (Nordenfelt, 2008).

## 2.2.2 Arbeitsfähigkeit

Um eine umfassende Beschreibung des Konzepts der Arbeitsfähigkeit darzustellen, ist Tengland (2011) der Frage nachgegangen, ob das Vorliegen von Gesundheit ausreichend für die Arbeitsfähigkeit ist. Er kommt zu dem Schluss, dass Gesundheit allein keine notwendige Voraussetzung für die Arbeitsfähigkeit ist, denn auch mit Gesundheitseinschränkungen sind Menschen in der Lage, ihrer Arbeitstätigkeit nachzugehen. Ebenfalls ist eine vorhandene Gesundheit nicht generell ausreichend für die Arbeitsfähigkeit, da je nach Arbeitstätigkeit zumindest bestimmte Grundfähigkeiten wie lesen, schreiben oder zählen erforderlich sind.

Als weitere in der Literatur genannten Merkmale der Arbeitsfähigkeit werden bei Tengland (2011) Basisfähigkeiten und berufliche Kompetenzen, Motivation, der Charakter und weitere Qualifikationen diskutiert, welche ebenfalls bei Nordenfelt (2008) bei der Konzeption der Arbeitsfähigkeit betrachtet werden. So führt Tengland (2011) weiter aus, dass auch bei Vorhandensein all dieser Merkmale und somit dem Vorliegen der Arbeitsfähigkeit eine Person nicht in der Lage sein kann, die Arbeitstätigkeit durchzuführen, da ggf. die Arbeitsvoraussetzungen zu hoch und beanspruchend sein können oder die Arbeitsumwelt aufgrund technischer Gegebenheiten kein Arbeiten ermöglicht. Insgesamt kommt Tengland (2011) zu dem Ergebnis, dass zwei Konzepte der Arbeitsfähigkeit notwendig sind: eine spezifische und eine generelle Arbeitsfähigkeit. Das Konzept der spezifischen Arbeitsfähigkeit bezieht sich auf die Fähigkeit einer Person, nach einem Ausfall am Arbeitsplatz ihre bisherige Arbeitstätigkeit wieder mit gleicher Qualität wie jemand anderes mit diesem Beruf aufzunehmen, wohingegen das Konzept der generellen Arbeitsfähigkeit auf die Aufnahme irgendeiner Arbeitstätigkeit bezogen ist, welche Personen üblicherweise nach einer Trainingszeit bewältigen können. Nach Tengland (2011) können Personen im Sinne beider Definitionen, nur im Sinne der zweiten Definition oder gar nicht arbeitsfähig sein.

Die für die vorliegende Arbeit relevante Definition von Arbeitsfähigkeit stammt in ihren Ursprüngen bereits aus Forschungen in den 80er Jahren, welche durch die finnische Regierung angestoßen wurden (Ilmarinen, J. & Tuomi, 2004). Für die Regierung stellte sich im Kontext des demografischen Wandels die Frage, ob es weiterhin sinnvoll ist, das Renteneintrittsalter berufsgruppenabhängig zu gestalten. Zum damaligen Diskussionszeitpunkt reichte die Spanne des Renteneintrittsalters von 53 Jahre für Tiefseetaucher bis 63 Jahre für die Verwaltungsbranche (Ilmarinen, J., 2009). Definiert wurde und wird die Arbeitsfähigkeit durch das Forschungsteam im Finnischen Institut für Arbeitsmedizin (Finnish Institute of Occupational Health – FIOH) als das Vermögen einer Person, die ihr gestellten Arbeitsaufgaben unter Berücksichtigung der Gesundheit, der Arbeitsbedingungen und der mentalen Ressourcen zu bewältigen, was einer Weiterentwicklung der rein medizinischen Betrachtung entspricht (Ilmarinen, J., 2004; Ilmarinen, J. et al., 2008). Dieses Balancemodell der Arbeitsfähigkeit sieht die Beanspruchung innerhalb einer Person als eine Folge von Belastungen am Arbeitsplatz unter Berücksichtigung der abpuffernden individuellen Ressourcen an (Ilmarinen, J., 2009). Aus einer guten Balance kann Gesundheit, Arbeitsfähigkeit und berufliches Wohlbefinden resultieren. Aus einer Imbalance, welche sich sowohl in zu hoher als auch zu niedriger Beanspruchung ausdrücken kann, können dagegen arbeitsbedingte Erkrankungen und Überlastung folgen (Ilmarinen, J. et al., 2008). Diese Balance unterliegt einem stetigen Wandel durch Veränderungen in der Person (z. B. das steigende Alter) und Veränderungen in der Arbeitswelt (z. B. durch veränderte Technologien) und muss daher kontinuierlich überwacht werden (Ilmarinen, J., 2009).

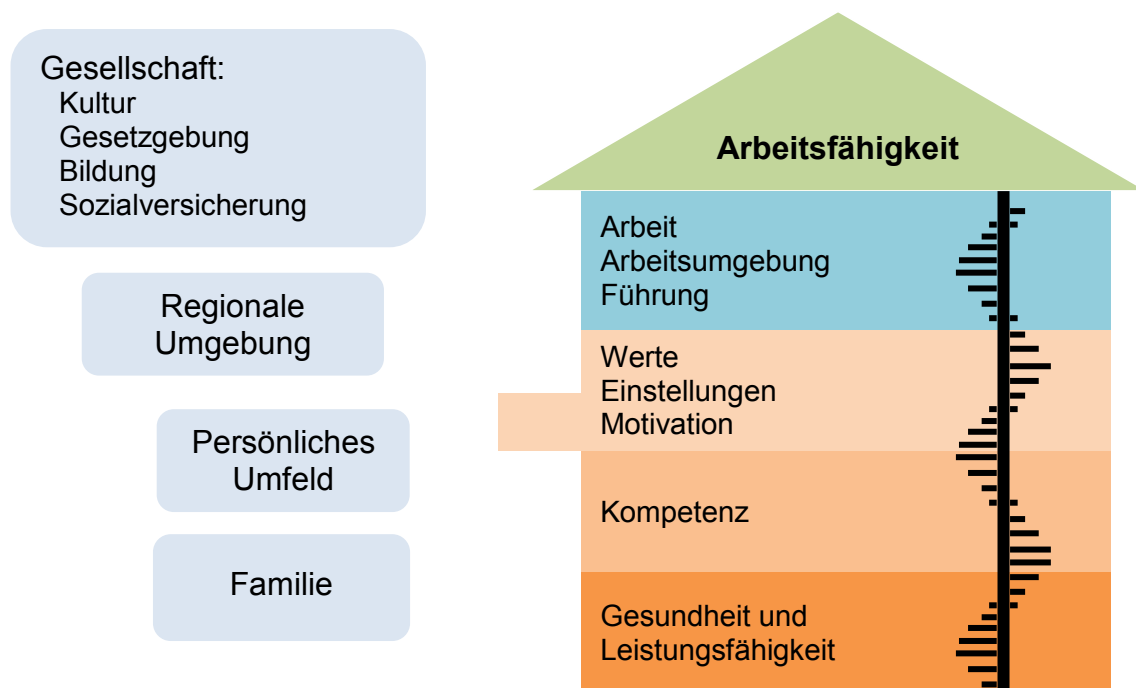
Integrative Modelle der Arbeitsfähigkeit schließen als Einflussfaktoren zusätzlich die Arbeit an sich und die Umwelt außerhalb der Arbeit ein (Ilmarinen, J. et al., 2008). Kombiniert aus dem Balance-Modell und integrativen Konzepten ist das multidimensionale Arbeitsfähigkeits-Modell entstanden, welches aus den zwischen Arbeit und Person balancierten und untereinander interagierenden Dimensionen Bewältigung bei der Arbeit, Kontrolle über die eigene Arbeit und Partizipation in der Arbeitsgemeinschaft zusammengesetzt ist (Järvikoski, Härkäpää & Mannila, 2001; zitiert nach Ilmarinen, J. et al., 2008, S. 16). Das der vorliegenden Arbeit zugrundeliegende Arbeitsfähigkeitsmodell ist ebenfalls multidimensional und hat sich aus den oben genannten Forschungen des FIOH um Juhani Ilmarinen auf Grundlage des Balancemodells entwickelt. Es umfasst jedoch zudem das außerberufliche Umfeld bestehend aus der Familie und weiteren engen Beziehungen, ebenso wie der Gesellschaft, und wird aktuell als Haus der Arbeitsfähigkeit abgebildet, welches in Kapitel 2.2.3 näher erläutert wird (Ilmarinen, J. et al., 2008). Nach diesem Modell ist die Arbeitsfähigkeit multifaktoriell bedingt und wird heute beschrieben als „die Summe von Faktoren, die eine Frau oder einen Mann in einer bestimmten Situation in die Lage versetzen, eine gestellte Aufgabe erfolgreich zu bewältigen“ (Ilmarinen, J. & Tempel, 2002, S. 166).

Die Arbeitsfähigkeit entsteht somit in einem Wechselspiel zwischen der Person und einer konkreten Arbeitssituation. Sie ist dadurch von der Beschäftigungsfähigkeit abzugrenzen, welche eine generelle andauernde Fähigkeit zur Arbeit auch unter wechselnden Situations- und Arbeitsmarktbedingungen darstellt (Prümper & Richenhagen, 2011). Vielmehr ist die Arbeitsfähigkeit eine Teilmenge und Voraussetzung der Beschäftigungsfähigkeit (Freude & Pech, 2005), welche bei Brussig und Knuth (2009) in die qualifikations- und tätigkeitsbezogene Komponente, die marktbezogene Komponente und die soziale Stabilität aufgeteilt wird. Die erste und die dritte Komponente haben dabei in Teilen entsprechende Verknüpfungspunkte zu den Dimensionen der Arbeitsfähigkeit nach J. Ilmarinen et al. (2008).

Arbeitsfähigkeit kann auch nicht automatisch als der entgegengesetzte Pol von Arbeitsunfähigkeit angesehen werden. Zu beachten sind dabei die unterschiedlichen Perspektiven der Wissenschaft, des Sozialrechts und der Betriebspraxis. Für den Arbeitgeber ist sein Arbeitnehmer genau dann arbeitsunfähig, wenn er rechtzeitig eine entsprechende Bescheinigung von seinem Arzt vorlegt (Prümper & Richenhagen, 2011). Diese Arbeitsunfähigkeitsbescheinigung entsteht aus der sozialrechtlichen Perspektive durch §2 Abs. 1 der Richtlinien über die Beurteilung der Arbeitsunfähigkeit und die Maßnahmen zur stufenweisen Wiedereingliederung vom 17. Dezember 2015 (Gemeinsamer Bundesausschuss, 2015). Eine Arbeitsunfähigkeit besteht demnach dann, wenn die Person aufgrund einer körperlichen, geistigen oder seelischen Erkrankung nicht in der Lage ist, ihre Arbeit fortzuführen bzw. eine Verschlimmerung der Erkrankung die Folge wäre oder eine Arbeitsunfähigkeit als Folge aus der Ausübung der Arbeitstätigkeit trotz Erkrankung entstehen könnte. Somit kann die Arbeitsunfähigkeit sozialrechtlich in den diskreten Ausprägungen „liegt vor“ und „liegt nicht vor“ eingeteilt werden, ohne dass Abstufungen zwischen geringer und starker Arbeitsunfähigkeit berücksichtigt werden. Tengland (2011) allerdings betrachtet in seiner wissenschaftlichen Sichtweise die Arbeitsfähigkeit auf einem Kontinuum mit den Polen „voll arbeitsfähig“ und „komplett arbeitsunfähig“. Daraus folgend liegt bei einer psychologisch diagnostizierten sehr niedrigen Arbeitsfähigkeit bzw. hohen Arbeitsunfähigkeit nicht zwangsläufig auch eine Arbeitsunfähigkeit aus sozialrechtlicher Perspektive vor.

### 2.2.3 Das Haus der Arbeitsfähigkeit

Um die Dimensionen der Arbeitsfähigkeit anschaulich darzustellen und gleichzeitig potentielle Ansatzpunkte zur Förderung der Arbeitsfähigkeit mittels BGM den Betrieben zu vermitteln, wurde das Haus der Arbeitsfähigkeit konzipiert (Ilmarinen, J. et al., 2008). Jede Etage kann dabei als einzelner interventionsbezogener Ansatzpunkt für ein gesünderes und längeres Arbeiten gesehen werden. Optimal wäre jedoch nach J. Ilmarinen (2011) eine umfassende gleichzeitige Betrachtung aller Etagen. Wie in Abbildung 2.2 zu erkennen ist, besteht die aktuellste Darstellung des Hauses aus vier verbundenen und aufeinander aufbauenden Stockwerken, einem Dach und der umgebenden Umwelt.



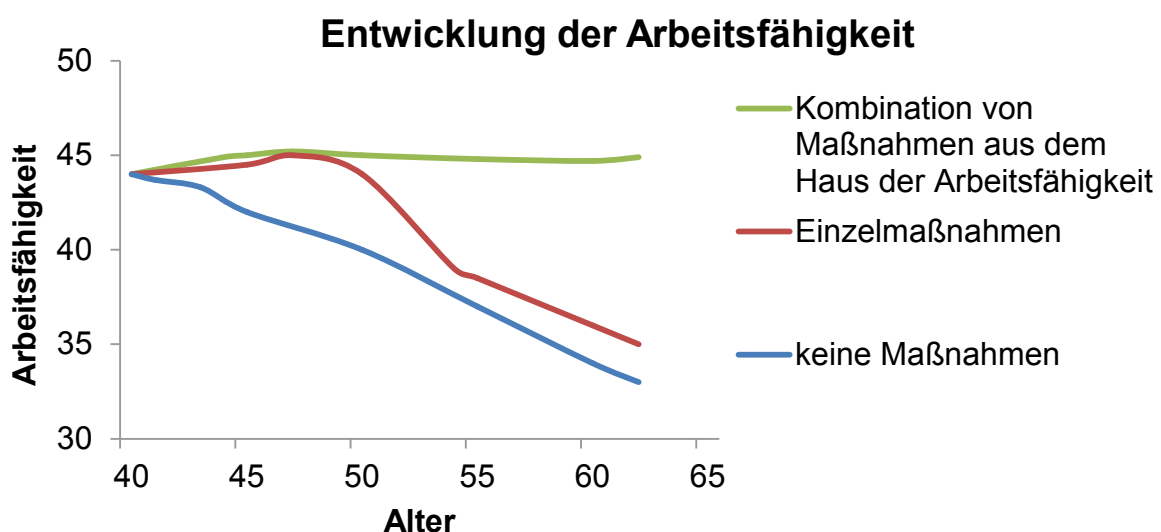
**Abb. 2.2** Haus der Arbeitsfähigkeit (nach Tempel und Ilmarinen, 2013).

Die Beschreibung des Hauses der Arbeitsfähigkeit erfolgt umfassend durch Tempel und J. Ilmarinen (2013) sowie J. Ilmarinen et al. (2008) und wird hier anhand dieser Literatur grob skizziert. Die ersten drei Etagen bilden die individuellen Ressourcen, die vierte Etage repräsentiert die Arbeit der Person und der Balkon stellt die Verbindung zur Umwelt dar. Die erste Etage gilt als das Fundament des Hauses und beinhaltet die individuelle Gesundheit, Krankheit und Leistungsfähigkeit, wobei die Gesundheit aus einem Gleichgewicht physischer, psychischer und sozialer Fähigkeiten resultiert. Das Vorhandensein einer Krankheit ist für die Arbeitsfähigkeit weniger ausschlaggebend als Beeinträchtigungen durch diese Krankheit am Arbeitsplatz. Vielmehr kann eine unterstützende Arbeitsgestaltung und ein BGM auf dieser Etage ein Ansatzpunkt zur Förderung der Arbeitsfähigkeit sein. In der zweiten Etage spielt die individuelle Kompetenz eine Rolle für die Arbeitsfähigkeit, welche sich aus Fachkompetenzen wie Fähigkeiten und Wissen, Methodenkompetenz und Sozialkompetenz zusammensetzt. Diese Kompetenzen unterliegen im Lebensverlauf einem stetigen Wandel und können durch lebenslanges Lernen aktiv beeinflusst werden. Dazu zählt auf betrieblicher Ebene das Angebot von Weiterbildungsmaßnahmen für alle



Altersgruppen und Beschäftigungsverhältnisse. Als interne Ressourcen in der Arbeitswelt haben Werte, Einstellungen und Motivation in der dritten Etage Relevanz für die Arbeitsfähigkeit. Diese werden durch Anforderungen aus der Arbeitswelt positiv oder negativ beeinflusst, haben selbst einen wechselseitigen Einfluss auf das Betriebsklima und stehen dadurch in enger Verbindung mit der vierten Etage. In der vierten und größten Etage erlangt die Arbeit mit allen Bedingungen, Inhalten, Anforderungen und Führungsqualitäten die bedeutendste Rolle für die Arbeitsfähigkeit. So richten sich die Etagen eins bis drei an den Anforderungen der vierten Etage aus. Besteht ein Ungleichgewicht zwischen den unteren Etagen und der vierten Etage, so wirkt sich dies negativ auf die Arbeitsfähigkeit aus. Die Arbeitsfähigkeit wird wie bereits beschrieben nicht losgelöst vom Kontext betrachtet. Die Familie, das persönliche Umfeld sowie die regionale Umgebung und die Gesellschaft können sowohl als Rückhalt für die Arbeitsfähigkeit wirken als auch Risiken bergen. So kann die Familie als soziale Unterstützung angesehen werden, jedoch können auch familiäre Konflikte in das Arbeitsleben hineingetragen werden, welche sich wiederum auf die Arbeitsfähigkeit auswirken. In der Gesellschaft werden über Kultur, Gesetzgebung, Bildung und Sozialversicherungen indirekt Anforderungen an die Arbeitsfähigkeit gestellt, aber auch Möglichkeiten zur Erreichung einer langen und guten Arbeitsfähigkeit formuliert.

Die Dimensionen der Arbeitsfähigkeit unterliegen Veränderungen innerhalb der Lebensspanne einerseits durch das individuelle Alter, andererseits durch Entwicklungen in der Arbeitswelt (Ilmarinen, J. et al., 2008). Eine durchschnittliche Entwicklung der Arbeitsfähigkeit im Unternehmen, gemessen mit dem WAI von Tuomi et al. (1998), in Abhängigkeit von durchgeführten Maßnahmen zur Förderung oder zum Erhalt der Arbeitsfähigkeit ist beispielhaft in Abbildung 2.3 dargestellt (Richenhagen, 2011). Ohne entsprechende Maßnahmen nimmt die Arbeitsfähigkeit mit dem Alter ab (blaue Linie). Werden lediglich Maßnahmen auf einzelnen Etagen des Hauses der Arbeitsfähigkeit durchgeführt, sind zwar positive Effekte zu erwarten, welche jedoch nicht nachhaltig sind (rote Kurve). Erst bei Interventionen mit einer ganzheitlichen Betrachtung aller Etagen kann die Arbeitsfähigkeit auch bis ins höhere Erwerbsalter stabil gehalten werden (grüne Linie).



**Abb. 2.3** Idealtypische Darstellung der Entwicklung der Arbeitsfähigkeit (modifiziert nach Richenhagen, 2011).

Obwohl die Auswirkungen von beeinträchtigenden Faktoren auf die Arbeitsfähigkeit erst ab dem 45. Lebensjahr bei den Erwerbstätigen sichtbar werden, sollte die Erhaltung und Verbesserung in allen Phasen des Arbeitslebens erfolgen. Dies zahlt sich nicht nur während des Arbeitslebens für die Arbeitnehmer- und Arbeitgeberseite aus, sondern auch nach dem Renteneintritt in Bezug auf eine bessere Lebensqualität und geringeren Kosten für das Gesundheitswesen (Tuomi et al., 1998).

Die empirische Fundierung der Einflussfaktoren auf die Arbeitsfähigkeit erfolgte während der Weiterentwicklung des Konzepts anhand umfangreicher, meist finnischer Studien, welche individuelle und arbeitsbezogene Determinanten von Arbeitsfähigkeit ermittelten (Gould, Ilmarinen, J., Järvisalo & Koskinen, 2008; van den Berg, Elders, de Zwart & Burdorf, 2009). Bei den individuellen Einflussfaktoren der Arbeitsfähigkeit fanden sich in diesen Studien z. B. Zusammenhänge mit dem Alter, dem Bildungsstand, mentalen Problemen, Krankheiten wie Depression und Rückenschmerz, schwierigen Lebensumständen außerhalb der Arbeit, Motivation und der körperlichen Fitness. Zusammenhänge der Arbeitsfähigkeit mit hohen psychischen, physischen und physikalischen Arbeitsanforderungen, Wissen und Fertigkeiten, Einstellungen, Arbeitszufriedenheit und fehlender Autonomie sind Beispiele für arbeitsbezogene Einflussfaktoren. Weiterhin fanden sich Unterschiede in der Arbeitsfähigkeit zwischen verschiedenen Berufsgruppen (Gould & Polvinen, 2008) und den Geschlechtern (Gould, Polvinen & Seitsamo, 2008). In aktuelleren Studien spielt ebenfalls das Arbeitsengagement als mediierender Faktor zwischen arbeits- und personenbezogenen Bedingungen und Arbeitsfähigkeit eine Rolle (Airila et al., 2014; Rongen, Robroek, Schaufeli & Burdorf, 2014). In den finnischen Längsschnittstudien zur Arbeitsfähigkeit aus den 80er und 90er Jahren wurden außerdem Zusammenhänge zwischen der erhobenen Arbeitsfähigkeit und dem vorzeitigen Berufsausstieg sowie der Mortalität elf Jahre später gefunden (Tuomi et al., 1997). Diese Befunde wurden durch Salonen, Arola, Nygard, Huhtala und Koivisto (2003) in einer weiteren 11-jährigen Längsschnittstudie zwischen 1989 und 2000 repliziert. Alavinia, de Boer, van Duivenbooden, Frings-Dresen und Burdorf (2009) konnten in einer 23-monatigen Längsschnittstudie den Erhalt einer Invaliditätsrente durch eine mäßige oder niedrige Arbeitsfähigkeit voraussagen.

In den meisten der bisher genannten Studien wurde die Arbeitsfähigkeit mit dem WAI erhoben. Dieser und ein Auszug weiterer Verfahren zur Erfassung der Arbeits- und Funktionsfähigkeit werden in den folgenden Abschnitten beschrieben. Der WAI erhält aufgrund der Thematik der vorliegenden Arbeit eine umfangreichere Betrachtung in einem gesonderten Kapitel.

### **2.3 Operationalisierung von Arbeits- und Funktionsfähigkeit**

Um die individuelle Arbeitsfähigkeit zu bewerten, gibt es verschiedene Ansatzpunkte. So kann man zwischen subjektiven und objektiven Erhebungen aus verschiedenen Perspektiven unterscheiden. Zum einen kann die arbeitende Person selbst gebeten werden, die eigene Arbeitsfähigkeit einzuschätzen. Zum anderen kann die Arbeitsfähigkeit durch (Arbeits)Mediziner oder externe Experten begutachtet werden. Sie kann anhand der individuellen Gesundheit und Funktionsfähigkeit im medizinischen Sinne oder anhand von Arbeitsergebnissen wie Leistung oder Produktivität bewertet werden. In einem systematischen Review zum Vergleich verschiedener Instrumente zur Messung von Interventionseffekten auf die Arbeitsfähigkeit (Amler, Felder, Mau, Merkesdal & Schöffski, 2015) wurden anhand von Schlagwörtern in den Themenbereichen Arbeitsplatz, Instrument und Arbeits(un)fähigkeit, Präsentismus, Absentis-

mus und Produktivität die am häufigsten genannten Instrumente untersucht. Zu diesen zählten u. a. der Work Productivity and Activity Impairment Questionnaire (WPAI) von Reilly, Zbrozek und Dukes (1993) in verschiedenen Ausführungen, der Work Limitations Questionnaire (WLQ) von Lerner et al. (2001), der Health and Work Performance Questionnaire (HPQ) von Kessler et al. (2003) und der WAI (Tuomi et al., 1998), welcher bei anschließenden Experteninterviews den höchsten Bekanntheitsgrad aufwies. Von den gefundenen Instrumenten erfasste lediglich der WAI die Arbeitsfähigkeit direkt. Die übrigen Fragebögen beschränken sich auf Absentismus/Präsentismus und Produktivität in Abhängigkeit von vorliegenden Gesundheitseinschränkungen. Die Autorengruppe kommt letztendlich zu dem Schluss, dass aufgrund von Mängeln in den ausgewerteten Kriterien keines der gefundenen Instrumente als ausschließliche Erhebungsmethode zu empfehlen, jedoch eine gemeinsame Verwendung des WAI und des WPAI für die zukünftige Vergleichbarkeit von Studien anzustreben ist (Amler et al., 2015).

Neben den bereits genannten Fragebögen existieren weitere Instrumente zur Operationalisierung der Arbeits- und Funktionsfähigkeit. Basierend auf der ICF (WHO, 2001) wurde die Norwegian Function Assessment Scale (NFAS) zur Beurteilung der Notwendigkeit von Sozialleistungen und Anpassung der Arbeitsanforderungen zur Wiedereingliederung von krankgeschriebenen Personen entwickelt (Brage, Fleten, Knudsrød, Reiso & Ryen, 2004; Østerås et al., 2007). Die NFAS erfasst anhand einer Selbstbeurteilung die Durchführbarkeit relevanter Handlungen für die Arbeit und Freizeitaktivitäten aus den ICF-Komponenten Aktivitäten und Teilhabe in Abhängigkeit vom gegenwärtigen Gesundheitszustand. Unterschieden werden dabei die mentale und die physische Funktionsfähigkeit. Erfolgreich validiert (Brage et al., 2004) wurde die NFAS u. a. am Short-Form-36 Health Survey (SF-36). Letzterer Fragebogen erfasst die selbstberichtete gesundheitsbezogene Lebensqualität für die körperliche und psychische Gesundheit über acht Dimensionen: körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen, allgemeine Gesundheit, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und mentale Gesundheit (Ware, 2000; Ware, Snow, Kosinski & Gandek, 1993). Der SF-36 liegt außerdem in verschiedenen Kurzversionen (SF-12 und SF-8) vor und wird in internationalen Studien standardisiert in verschiedenen Übersetzungen genutzt (Ware, 2000; Ware, Kosinski, Dewey & Gandek, 2001; Ware, Kosinski & Keller, 1996).

Als ein Oberbegriff für standardisierte Instrumente zur Evaluation der funktionellen körperlichen Fähigkeiten für die Arbeit gilt die Functional Capacity Evaluation (FCE), von der verschiedene Einsatzformen zur Entwicklung und Evaluation von Rehabilitationsmaßnahmen oder zur Vorhersage der Rückkehr zur Arbeit verfügbar sind (Chen, 2007). Bei den Definitionen und Rahmenbedingungen zur Anwendung der verschiedenen FCE-Instrumente herrscht in der bisherigen Literatur hauptsächlich Uneinigkeit (Soer, van der Schans, Groothoff, Geertzen & Reneman, 2008). In einer Delphi-Studie von Soer et al. (2008) beinhaltet die Definition mit der höchsten Zustimmung die körperlichen Funktionen und Strukturen, Umweltbedingungen, persönliche Faktoren und den Gesundheitsstatus als zu betrachtende individuelle Faktoren für die Empfehlung zur Teilhabe am Erwerbsleben. Eine Anlehnung an die Komponenten der ICF (WHO, 2001) ist hier erkennbar.

Diese Aufzählung stellt lediglich eine unvollständige Darstellung der verschiedenen Messinstrumente zur Arbeits- und Funktionsfähigkeit dar, welche einer großen Heterogenität bezüglich der zugrundeliegenden Rahmen und Perspektiven unterliegen. Im folgenden Kapitel wird nun der WAI vorgestellt, welcher in den vorangegangenen

Darstellungen schon mehrfach erwähnt wurde und eine zentrale Rolle für die vorliegende Arbeit spielt.

## 2.4 Work Ability Index (WAI)

Der WAI wurde im Rahmen von finnischen Studien des FIOH durch die Arbeitsgruppe von Toumi et al. (1985) zur Beantwortung der Frage nach einem berufsgruppenabhängig gestaffelten Renteneintrittsalter und in Ermangelung eines geeigneten Instruments zur subjektiven Erfassung der Arbeitsfähigkeit entwickelt. Im deutschen Sprachraum ist der frei verfügbare WAI auch bekannt unter den Namen Arbeitsfähigkeitsindex oder Arbeitsbewältigungsindex (BAuA, 2013). Den Rahmen für die Operationalisierung bildete die aktuelle und zukünftige Bewältigung der Arbeit in Abhängigkeit von den körperlichen und geistigen Arbeitsanforderungen, der Gesundheit und den eigenen Ressourcen. Dazu wurden aus einer Sammlung von Fragen zur Arbeit, Arbeitsfähigkeit, Gesundheit und psychischen Kraftreserven unter Berücksichtigung der Korrelationen der Items untereinander und einer Kreuzklassifizierung die zehn relevanten Items zur Bildung des Arbeitsfähigkeitsindex ausgewählt und in dessen Berechnung integriert (Toumi et al., 1985). Der Gesamtindex wird anhand einer vorgeschriebenen Punktevergabe berechnet und fasst sieben WAI-Dimensionen zusammen (WAI-Netzwerk, 2015):

- (1) Derzeitige Arbeitsfähigkeit im Vergleich zu der besten je erreichten Arbeitsfähigkeit
- (2) Arbeitsfähigkeit im Vergleich zu den Anforderungen der Arbeitstätigkeit
- (3) Anzahl der aktuellen vom Arzt diagnostizierten Krankheiten
- (4) Geschätzte Beeinträchtigung der Arbeitsleistung durch die Krankheiten
- (5) Krankenstandstage in den vergangenen 12 Monaten
- (6) Einschätzung der eigenen Arbeitsfähigkeit in zwei Jahren
- (7) Psychische Leistungsreserven.

Validiert wurde der WAI anhand von Vergleichen mit objektiven klinischen Untersuchungen (Eskelinen, Kohvakka, Merisalo, Hurri & Wägar, 1991) sowie durch Längsschnittuntersuchungen zum Zusammenhang von Arbeitsfähigkeit mit dem Erhalt von Berufsunfähigkeitsrente und Sterblichkeit in einem Zeitraum von 11 Jahren (Tuomi et al., 1997). Es stellte sich bei den Untersuchungen heraus, dass gerade die nicht explizit gesundheitsbezogenen Dimensionen (6), (2), (4) und (1) die höchste Aussagekraft (in dieser Reihenfolge) für die Prädiktivität der Arbeitsfähigkeit hatten (Ilmarinen, J. & Tuomi, 2004). Wie bereits in Kapitel 2.2.3 erwähnt, konnten die Befunde zur Berufsunfähigkeitsrente und Sterblichkeit durch Alavinia et al. (2009) und Salonen et al. (2003) repliziert werden. In späteren Studien mit übersetzten Versionen des WAI wurde die Konvergenzvalidität u. a. am SF-36 (Abdolalizadeh et al., 2012; Alexopoulos, Merikoulias, Gnardellis & Jelastopulu, 2013; Radkiewicz & Widerszal-Bazyl, 2005) zufriedenstellend überprüft.

Eingesetzt werden kann der mit lediglich zehn Fragen recht kurze WAI-Fragebogen als Screening-Instrument im betrieblichen Kontext u. a. zur Selbstauskunft bei Beschäftigten oder im Dialog mit dem Betriebsarzt sowie zu wissenschaftlichen Zwecken bei Untersuchungen zur Arbeitsfähigkeit der Bevölkerung. Dies kann sowohl anhand von Ist-Analysen im Querschnittsdesign, in deren Folge Interventionsmaßnahmen angeregt werden, als auch in Form von Verlaufsanalysen im Längsschnitt-design in Folge von Interventionsmaßnahmen erfolgen (BAuA, 2013). Häufig wird er

zusammen mit weiteren Erhebungsinstrumenten eingesetzt, da er lediglich Teilaspekte der Arbeitsfähigkeitsdimensionen gemäß dem Haus der Arbeitsfähigkeit abbildet (WAI-Netzwerk, 2015).

Der WAI wurde in den 90er Jahren zu einer zweiten Version weiterentwickelt und wird heute in mindestens 26 Übersetzungen weltweit zu Forschungszwecken eingesetzt (Ilmarinen, J., 2009; Tuomi et al., 1998). Es existiert außerdem eine Kurzversion, bei der die Krankheitsliste für die Dimension (3) von 51 Krankheiten auf 14 Krankheitsgruppen reduziert wurde (Nübling, Hasselhorn, Seitsamo & Ilmarinen, J., 2005). Des Weiteren kann die Dimension (1) Derzeitige Arbeitsfähigkeit im Vergleich zu der besten je erreichten Arbeitsfähigkeit als Einzelitem im Work Ability Score (WAS) erfasst werden (Gould, Koskinen, Seitsamo, Tuomi, Polvinen & Sainio, 2008). Eine Vergleichsstudie bestätigte die konvergente Validität des WAS gegenüber dem WAI, wodurch sich eine Anwendung des WAS für umfangreichere Befragungen aufgrund seiner Kürze empfiehlt (El Fassi et al., 2013). Mit dem jüngst entwickelten WAI 2.0, bestehend aus dem work ability – personal radar und einem anschließenden Dialog (Betriebs-Radar), soll der bereits genannte Mangel der nicht umfangreichen Betrachtung des Hauses der Arbeitsfähigkeit im WAI ausgebessert werden (Ilmarinen, V., Ilmarinen, J., Huuhtanen, Louhevaara & Näsman, 2015). Der Fokus liegt weniger auf Erkrankungen und gesundheitsbezogenen Beeinträchtigungen, sondern neben den persönlichen Ressourcen auf dem Arbeitskontext und der Verbindung zwischen Arbeit und Freizeit. Jedoch bedarf es zur Nutzung des WAI 2.0 einer kostenpflichtigen Schulung zum Lizenzerwerb für Deutschland, Österreich sowie die Schweiz. Betriebe müssen für den Einsatz der Methode ebenfalls Lizenzgebühren leisten (Ilmarinen, J., 2016). Es wird sich erst in der Zukunft zeigen, inwieweit dieses Instrument aufgrund seiner Kosten in den Betrieben angenommen wird.

Die Indexwerte des WAI in der klassischen zweiten Version (Tuomi et al., 1998) können Werte zwischen 7 und 49 Punkte erreichen, wobei niedrigere Punkte eine schlechtere Arbeitsfähigkeit anzeigen. Eine Kategorisierung der Arbeitsfähigkeit erfolgte in den 80er Jahren derart, dass die 15% der Beschäftigten mit den schlechtesten Indexwerten in die niedrigste Kategorie fielen und die 15% der Beschäftigten mit den besten Indexwerten in die höchste Kategorie. Die übrigen 70% wurden durch den Median in zwei weitere Kategorien aufgeteilt. Als Ergebnis ergaben sich die in Tabelle 2.1 dargestellten vier Kategorien der Arbeitsfähigkeit mit vier abgeleiteten Zielen von Maßnahmen (Tuomi et al., 1998). Die Aktualität dieser Einteilung sollte jedoch aufgrund der Veränderungen in der Arbeitswelt an einer aktuellen Stichprobe überprüft werden. Des Weiteren kann die Arbeitsfähigkeit auch als kontinuierliche Variable untersucht werden (WAI-Netzwerk, 2015).

**Tab. 2.1** Kategorisierung der Arbeitsfähigkeit und Ableitung von Maßnahmen nach Tuomi et al. (1998)

Anteil	Punkte	Arbeitsfähigkeit	Ziele von Maßnahmen
15%	7-27	Schlecht	Arbeitsfähigkeit wiederherstellen
35%	28-36	Mittelmäßig	Arbeitsfähigkeit verbessern
35%	37-43	Gut	Arbeitsfähigkeit unterstützen
15%	44-49	Sehr gut	Arbeitsfähigkeit erhalten

In Deutschland steht den interessierten Nutzern das WAI-Netzwerk als Informationsportal zur Arbeitsfähigkeit und zum WAI zur Verfügung (<http://www.arbeitsfaehig.com/de/wai-netzwerk-35.html>). Die Gründung des Netzwerks wurde im Jahr 2003 durch die BAuA initiiert und die Betreuung fand bis 2015 durch die Bergische Universität Wuppertal statt. Seit dem Frühjahr 2016 wird die Arbeit dieses Netzwerks durch das Institut für Arbeitsfähigkeit fortgeführt. Ziel ist die Förderung der Anwendung des WAI in deutschen Betrieben. Bisher ist es den Nutzern des WAI erst bei anonymisierter Freigabe der eigenen erhobenen WAI-Daten möglich, Zugriff auf eine nationale Referenzdatenbank zu erhalten, deren Daten jedoch nicht als Normdaten zu behandeln sind (WAI-Netzwerk, 2015).

## **2.5 Empirischer Forschungsstand und Herleitung der Fragestellungen**

Neben den Vorteilen der Kürze, Verständlichkeit und freien Verfügbarkeit des Fragebogens und der erprobten Einsätze in europäischen betrieblichen Kontexten (BAuA, 2002) erfuhr der WAI in der Vergangenheit Kritik aus den Reihen der Arbeitsmediziner. So wird u. a. angeführt, dass dieser lediglich die Beanspruchungsfolgen bei Erwerbstätigen erfasst, während die individuellen Arbeitsbedingungen unberücksichtigt bleiben (Elsner, 2005). Der WAI wurde jedoch als subjektives Screening-Instrument konzipiert, welches durch das Antwortverhalten der Teilnehmenden einen bestehenden Handlungsbedarf in individuellen oder arbeitsbezogenen Bereichen aufzeigen kann. Daher sind dem WAI weitergehende Gespräche und Untersuchungen zu den Quellen der Beeinträchtigungen anzuschließen (Bieneck, Sedlatschek, Kuhn, Freude & Pech, 2005).

Kritisch zu betrachten ist jedoch die lediglich auf Kreuzklassifizierungen, Korrelationsanalysen und inhaltlichen Überlegungen beruhende Annahme der Eindimensionalität des Arbeitsfähigkeits-Gesamtindex (Tuomi et al., 1985) sowie die vorgegebene Punktevergabe zur Berechnung des WAI, welche in der Literatur lediglich festgelegt, jedoch nicht weiter begründet wird (Ilmarinen, J. & Tuomi, 2004). Die sieben WAI-Dimensionen, welche aus zehn Items resultieren, bilden dabei keine Faktoren gemäß der Testtheorie ab (WAI-Netzwerk, 2015). So kamen verschiedene Reliabilitätsstudien während der Untersuchung der psychometrischen Eigenschaften des WAI zu unterschiedlichen Ergebnissen bezüglich dessen Faktorenstruktur. Martus, Jakob, Rose, Seibt und Freude (2010) fanden in ihren Analysen an einer deutschen Stichprobe mittels konfirmatorischer Faktorenanalysen, dass ein Modell, bei dem die sieben Dimensionen des WAI auf zwei korrelierenden Faktoren luden, bessere Fit-Indizes lieferte als das eindimensionale Modell. Die WAI-Dimensionen (1), (2) und (7) wurden dabei als Faktor der subjektiv eingeschätzten Arbeitsfähigkeit und Ressourcen interpretiert, die WAI-Dimensionen (3) und (5) als gesundheitsbezogener Faktor. Die WAI-Dimensionen (4) und (6) zeigten dagegen uneinheitliche Muster je nach Berufsgruppe. Eine Schwäche der Studie von Martus et al. (2010) ist die aus drei Einzelstudien gebildete ad-hoc-Stichprobe bestehend aus 324 weiblichen Bürokräften, Lehrerinnen und Lehrern sowie Erzieherinnen (Martus, Freude, Rose, Seibt & Jakob, 2011). Analysen im Rahmen des Second German Sociomedical Panel of Employees (Bethge, Radoschewski & Gutenbrunner, 2012) an 1 036 Erwerbstätigen hingegen stützen die Annahme einer einfaktoriellen Struktur des WAI, wobei mehrfaktorielle Modelle nicht explizit geprüft wurden. In einer früheren psychometrischen Untersuchung des WAI an rund 40 000 Krankenschwestern aus verschiedenen europäischen Ländern fanden sich dagegen länderspezifische Unterschiede in der Fakto-

renstruktur (Radkiewicz & Widerszal-Bazyl, 2005). So wurde anhand von Hauptkomponentenanalysen für Stichproben aus Deutschland und Finnland eine einfaktorische Struktur gefunden, während in den übrigen teilnehmenden Ländern eine übereinstimmende zweifaktorielle Struktur gefunden wurde. Die WAI-Dimensionen (1), (2), (6) und (7) wurden in dieser Studie als subjektive, die Dimensionen (3), (4) und (5) als objektive Komponenten der Arbeitsfähigkeit interpretiert. Auch weitere Untersuchungen mit übersetzten WAI-Versionen fanden uneinheitliche Faktorenstrukturen. In einer griechischen Stichprobe fanden sich zwei Faktoren, welche ähnlich zu den Ergebnissen von Martus et al. (2010) interpretiert wurden (Alexopoulos et al., 2013), während für iranische, brasilianische und argentinische Stichproben drei Faktoren des WAI gefunden wurden (Abdolalizadeh et al., 2012; Martinez, Latorre & Fischer, 2009; Peralta, Godoi Vasconcelos, Härter Griep & Miller, 2012). Tabelle 2.2 gibt eine Übersicht über die Studien mit der gefundenen Anzahl an Faktoren und der Verteilung der WAI-Dimensionen.

**Tab. 2.2** Befunde zur Faktorenstruktur des Work Ability Index

Autoren	Land	Faktoren	Faktor 1	Faktor 2	Faktor 3
Bethge et al. (2012)	D	1	(1) - (7)		
Radkiewicz & Widerszal-Bazyl (2005)	D & F Rest-Europa	1 2	(1) - (7) (1), (2), (6), (7)	(3), (4), (5)	
Martus et al. (2010)	D	2	(1), (2), (4)*, (6)*, (7)	(3), (4)*, (5), (6)*	
Alexopoulos et al. (2013)	G	2	(1), (2), (6)*, (7)	(3), (4), (5), (6)*	
Abdolalizadeh et al. (2012)	I	3	(1), (2), (6)	(3), (4), (5)	(7)
Martinez et al. (2009)	B	3	(1), (2)	(3), (4), (5), (6)	(7)
Peralta et al. (2012)	A	3	(1), (2), (7)	(5), (6)	(3), (4),

Anmerkungen. D = Deutschland, F = Finnland, G = Griechenland, I = Iran, B = Brasilien, A = Argentinien.

\*Doppelladung.

Die Heterogenität der bisherigen Studienergebnisse kann verschiedene Ursprünge haben. Sie ist als Resultat kultureller Unterschiede, sprachlicher Mängel in den Übersetzungen, nicht vergleichbarer Stichproben mit teilweise geringer Güte oder unterschiedlicher Analysemethoden denkbar. So wurden in den voran genannten Analysen, soweit im Methodenteil erläutert, statistische Analysen (Hauptkomponentenanalysen in IBM SPSS, konfirmatorische Faktorenanalysen in IBM SPSS AMOS)

angewandt, welche den Items des WAI ein metrisches Skalenniveau zuschreiben (Abdolalizadeh et al., 2012; Martus et al., 2010). Dies kann jedoch nicht für alle Items des WAI angenommen werden. Bei diesem Vorgehen beruhen die durchgeführten Faktorenanalysen auf für ordinale Variablen nicht angemessene Produkt-Moment-Korrelationen, obwohl die Berechnung von polychorischen Korrelationen dem vorzuziehen ist (Eid, Gollwitzer & Schmitt, 2013).

Um den in Deutschland weitgehend anerkannten und im BGM häufig im Zusammenhang mit dem Haus der Arbeitsfähigkeit eingesetzten WAI korrekt interpretieren zu können, soll in der vorliegenden Arbeit die Faktorenstruktur des WAI anhand von Strukturgleichungsmodellen unter Berücksichtigung des Messniveaus überprüft werden. Für die Analysen liegen bereits Daten einer repräsentativen Stichprobe der Erwerbsbevölkerung in Deutschland aus der „Studie zur Mentalen Gesundheit bei der Arbeit“ (S-MGA) vor. Die detaillierte Beschreibung der Studie erfolgt in Kapitel 3.1. Angelehnt wird die Überprüfung an die Untersuchung von Martus et al. (2010), so dass neben dem Modell mit einer einfaktoriellen Struktur zwei weitere genestete Modelle mit einer zweifaktoriellen Struktur confirmatorisch überprüft werden. Diese beiden Modelle unterscheiden sich dadurch, dass wie bei Martus et al. (2010) in einem Modell Doppelladungen für die Dimensionen (4) und (6) auf beide Faktoren zugelassen werden. Im Hinblick auf eine angenommene diskriminante Validität beider Faktoren auf die Arbeitsfähigkeit ist das Zulassen von Doppelladungen jedoch nur schwer interpretierbar. Darüber hinaus ergaben die Analysen von Martus et al. (2010), dass das Item der WAI-Dimension (4) in der Gesamtstichprobe auf dem zweiten Faktor deutlich höher lädt als auf dem ersten Faktor. Dagegen lädt das Item der WAI-Dimension (6) höher auf dem ersten Faktor als auf dem zweiten Faktor. Aus diesen Gründen wird ein viertes Modell überprüft, bei dem die WAI-Dimensionen (1), (2), (6) und (7) sowie die WAI-Dimensionen (3), (4) und (5) durch zwei Faktoren erklärt werden, wie es in den Studien von Radkiewicz und Widerszal-Bazyl (2005) außerhalb Deutschlands und Finnlands gefunden wurde. Grundsätzlich wird angenommen, dass die eindimensionale Struktur des WAI für eine repräsentative Erwerbstätigenstichprobe nicht bestätigt werden kann. Vielmehr wird erwartet, dass ein zweifaktorielles Modell mit Einfachstruktur, d. h. ohne Doppelladungen, zu einer besseren Modellgüte führt. Aufgrund der Befunde der genannten Studien wird dabei angenommen, dass das vierte Modell gemäß Radkiewicz und Widerszal-Bazyl (2005) am Besten auf die Daten passt. Die beiden Faktoren repräsentieren dabei zum einen die subjektiv eingeschätzte Arbeitsfähigkeit und Ressourcen und zum anderen die Gesundheitsbedingungen.

**Hypothese 1a:** Für eine repräsentative Erwerbstätigenstichprobe aus Deutschland wird die Annahme der zweifaktoriellen Struktur des WAI gestützt.

**Hypothese 1b:** Von den geprüften zweifaktoriellen Modellen zeigt das Modell gemäß Radkiewicz und Widerszal-Bazyl (2005) für die vorliegenden Daten eine bessere Modellpassung als die übrigen Modelle.

Die Abhängigkeit des Alters und des Geschlechts auf die mit dem WAI gemessene Arbeitsfähigkeit wurde bereits in den Längsschnittuntersuchungen von 1981-1992 betrachtet (Ilmarinen, J., Tuomi & Klockars, 1997). In den Studien von Gould, Polvinen und Seitsamo (2008) wurden das Absinken der gemessenen Arbeitsfähigkeit mit dem Alter und Unterschiede im WAI zwischen weiblichen und männlichen Erwerbstätigen bestätigt. Die ungünstigeren Ergebnisse im WAI der Frauen wurden dabei hauptsächlich durch eine größere Anzahl von Erkrankungen und krankheitsbe-



dingten Fehltagen hervorgerufen. Diese Ergebnisse spiegeln sich auch im Anstieg der AU-Tage mit dem Alter und einer höheren Anzahl von AU-Tagen für Frauen wider (Badura et al., 2010; Rebscher, 2016; Knieps & Pfaff, 2015). Auch Martus et al. (2011) fanden einen negativen Einfluss des Alters auf den WAI, welcher sich jedoch nicht für alle untersuchten Berufsgruppen zeigte.

Um Vergleiche von Testwerten zwischen Subgruppen vornehmen zu können, gilt als Voraussetzung, dass in den jeweiligen Subgruppen mit dem Testinstrument das gleiche Konstrukt gemessen wird. Diese Voraussetzung wird Messinvarianz genannt und lässt sich in vier hierarchisch abhängige Formen unterteilen: konfigurale, metrische, skalare und strikte faktorielle Messinvarianz (Christ & Schlüter, 2012). Konfigurale Messinvarianz bezieht sich dabei auf die Gleichheit der Faktorstruktur und Ladungsmuster der manifesten und latenten Variablen zwischen den Substichproben und ist die am wenigsten restriktive Form der Messinvarianz. Metrische Messinvarianz ist restriktiver als die konfigurale Messinvarianz und liegt vor, wenn die Faktorladungen zwischen den Subgruppen identisch sind. Skalare Messinvarianz wiederum ist restriktiver als die metrische Messinvarianz und impliziert zusätzlich die Gleichheit der Intercepts bzw. Konstanten (Höhe der Ausprägung der manifesten Variablen, wenn die latente Variable null ist) in den Subgruppen. Die Überprüfung der strikten faktoriellen Messinvarianz (Meredith, 1993) ist für den Vergleich zwischen Substichproben optional und bezieht sich auf die Gleichheit der Residuen der manifesten Variablen. Erst wenn wenigstens partielle skalare Messinvarianz zwischen den Subgruppen vorliegt, sind Vergleiche von Testwerten anhand von latenten Mittelwerten sinnvoll interpretierbar (Byrne, Shavelson & Muthén, 1989).

In der Literatur sind bisher keine Untersuchungen der Messinvarianz für den WAI bekannt. Da unklar ist, ob die verschiedenen Subgruppen die Items des WAI unterschiedlich wahrnehmen und bewerten (z. B. aufgrund unterschiedlicher körperlicher Voraussetzungen zwischen Männern und Frauen sowie zwischen Jüngeren und Älteren), wird für spätere Mittelwertvergleiche in der vorliegenden Arbeit die Messinvarianz des WAI für das Geschlecht und für drei Altersgruppen (31- bis 40-Jährige, 41- bis 50-Jährige und 51- bis 60-Jährige) überprüft. Grundsätzlich wird davon ausgegangen, dass der WAI in allen Subgruppen das gleiche Konstrukt der Arbeitsfähigkeit misst. Daraus leiten sich folgende zwei Hypothesen ab:

**Hypothese 2:** Für den WAI liegt skalare Messinvarianz zwischen Frauen und Männern vor.

**Hypothese 3:** Für den WAI liegt skalare Messinvarianz zwischen drei Altersgruppen vor.

Eine weitere zentrale Fragestellung ist, ob die Berechnung eines einzelnen Indexwertes der Arbeitsfähigkeit aus der Summation der ungewichteten Subskalen bei einer angenommenen Zwei-Faktoren-Struktur zukünftig angemessen ist oder ob nicht zwei gewichtete Indexwerte basierend auf Faktorwerten eine adäquatere Abbildung des zugrundeliegenden Konstrukts ermöglichen. Die individuellen Indexwerte bilden den Schätzwert für den jeweiligen Faktorwert einer Person, der wiederum die relative Ausprägung eines Personenmerkmals auf dem latenten Faktor abbildet (Grice, 2001). DiStefano, Zhu und Mîndrilă (2009) geben einen Überblick über verschiedene Möglichkeiten zur Berechnung von Faktorwerten, wobei sie zwischen refined (feine) und non-refined (grobe) Methoden unterscheiden. Zu den non-refined Methoden zählt unter anderem der wie bei dem WAI bisher angewandte ungewichtete Summenindex, welcher bei der Berechnung eine Gleichgewichtung aller Items impliziert,

unabhängig von der Ladung der jeweiligen Items auf dem Faktor. Die Auswertung von Martus et al. (2010) zeigte jedoch bereits, dass sich die Höhe der Item-Ladungen innerhalb des jeweiligen Faktors unterscheidet. Somit ist eine ungewichtete Indexbildung auch aufgrund der unterschiedlichen Metrik der Items nicht reliabel (DiStefano et al., 2009). Zu den refined Methoden gehört u. a. die gewichtete Berechnung von Faktorwerten anhand von Faktorwert-Koeffizienten, welche für den WAI im Rahmen dieser Arbeit angewandt werden soll. Die Gewichte sollen für die Anwender des WAI in Deutschland anhand der repräsentativen Stichprobe zur Verfügung gestellt werden. Da diese Koeffizienten für kategoriale Daten nicht direkt in Mplus ausgegeben werden können, muss eine Annäherung erfolgen, welche im Kapitel 3.4.4 spezifiziert wird. Die Gewichte sollen dabei weiterhin aus den Ergebnissen der konfirmatorischen Faktorenanalyse resultieren. Es wird erwartet, dass die annähernde Berechnung der gewichteten Faktorwerte eine adäquate Abbildung der exakten Faktorwerte darstellt. Als Produkt der neugebildeten Berechnung des WAI können getrennt nach Geschlecht und Altersgruppen Normwerttabellen präsentiert und deren Aussagekraft interpretiert werden.

**Hypothese 4:** Die auf Basis der konfirmatorischen Faktorenanalysen gewichteten Summenindizes der zwei Faktoren des WAI bilden dessen exakte Faktorwerte adäquat ab.

Weiterhin stellt sich die Frage, ob der WAI tatsächlich das multidimensionale Konstrukt der Arbeitsfähigkeit erfasst, wie es von den Entwicklern des WAI für das Haus der Arbeitsfähigkeit definiert wurde. Wie bereits bei den Validierungsstudien zum WAI aufgefallen ist, hatten gerade die nicht explizit gesundheitsbezogenen Dimensionen in der Reihenfolge (6), (2), (4) und (1), welche hauptsächlich und direkt die aktuelle und zukünftige subjektive Arbeitsfähigkeit erfragen, die höchste Aussagekraft für die Vorhersage der Berufsunfähigkeit und Mortalität (Ilmarinen, J. & Tuomi, 2004). Die Prädiktivität des WAI für den Bezug von Berufsunfähigkeitsleitungen wurde somit als Rechtfertigung für die Validität des Instruments herangezogen. Weitere Validierungen erfolgten, wie bereits in Kapitel 2.4 erwähnt, am SF-36 zur gesundheitsbezogenen Lebensqualität für die physische und mentale Funktionsfähigkeit und weiteren Instrumenten der physischen und mentalen Gesundheit in verschiedenen Stichproben (Abdolalizadeh et al., 2012; Alexopoulos et al., 2013; Martinez et al., 2009; Radkiewicz & Widerszal-Bazyl, 2005). So zeigten sich sowohl in den Studien zur iranischen, griechischen und brasilianischen Version des WAI als auch in der europäischen Studie an Krankenschwestern signifikant positive Zusammenhänge zwischen der mit dem WAI gemessenen Arbeitsfähigkeit und den Einzeldimensionen des SF-36. Neben der Überprüfung der inneren Faktorenstruktur des WAI soll daher ergänzend die Assoziation mit weiteren externen Indikatoren der Funktionsfähigkeit anhand einer repräsentativen Stichprobe Erwerbstätiger in Deutschland geprüft werden, um Hinweise zur konvergenten Validität der neugebildeten und gewichteten Einzelfaktoren zu erlangen. Dazu bieten sich als Instrumente für die Funktionsfähigkeit am Arbeitsplatz der SF-12 und der NFAS an (siehe Kapitel 2.3), da diese in der S-MGA ebenfalls erfasst wurden. Zusammenhänge des WAI mit dem NFAS wurden zum jetzigen Zeitpunkt noch nicht untersucht. Da der NFAS am SF-36 validiert wurde, können grundsätzlich Zusammenhänge zwischen den Faktoren des WAI mit den Skalen des SF-12 und der NFAS erwartet werden. Allgemein wird angenommen, dass höhere Ausprägungen der Faktoren der Arbeitsfähigkeit im WAI mit einer höheren gesundheitsbezogenen Lebensqualität (SF-12) und geringeren Einschränkungen

der Funktionsfähigkeit (NFAS) einhergehen. Dabei sollte der Zusammenhang jedoch nicht auf eine Redundanz der Methoden hindeuten.

**Hypothese 5a:** Es besteht ein positiver Zusammenhang der Faktoren des WAI mit den Skalen des SF-12.

**Hypothese 5b:** Es besteht ein negativer Zusammenhang der Faktoren des WAI mit den Skalen der NFAS.

Bei einer angenommenen zweifaktoriellen Struktur des WAI wird aufgrund der inhaltlichen Nähe vermutet, dass der gesundheitsbezogene Faktor stärker mit den physischen Skalen des SF-12 und der NFAS korreliert als der subjektive Faktor der Arbeitsfähigkeit und Ressourcen. Dagegen wird erwartet, dass der Faktor der Arbeitsfähigkeit und Ressourcen höher mit den mentalen Skalen des SF-12 und der NFAS korreliert als der gesundheitsbezogene Faktor.

**Hypothese 6a:** Bei einer zweifaktoriellen Struktur des WAI korreliert der gesundheitsbezogene Faktor höher mit der physischen Skala des SF-12 als der Faktor der Arbeitsfähigkeit und Ressourcen.

**Hypothese 6b:** Bei einer zweifaktoriellen Struktur des WAI korreliert der Faktor der Arbeitsfähigkeit und Ressourcen höher mit der mentalen Skala des SF-12 als der gesundheitsbezogene Faktor.

**Hypothese 7a:** Bei einer zweifaktoriellen Struktur des WAI korreliert der gesundheitsbezogene Faktor höher mit der physischen Skala der NFAS als der Faktor der Arbeitsfähigkeit und Ressourcen.

**Hypothese 7b:** Bei einer zweifaktoriellen Struktur des WAI korreliert der Faktor der Arbeitsfähigkeit und Ressourcen höher mit der mentalen Skala der NFAS als der gesundheitsbezogene Faktor.

## 3 Methode

Im folgenden Kapitel werden neben dem Untersuchungsdesign und der Durchführung der Studie ebenfalls die zur Beantwortung der Fragestellungen verwendeten Messinstrumente vorgestellt. Anschließend erfolgen die deskriptive Beschreibung der Stichprobe inklusive des Vorgehens der Datenbereinigung sowie eine Darstellung der angewandten statistischen Methoden.

### 3.1 Untersuchungsdesign und Durchführung

Die Daten für die Analysen der vorliegenden Arbeit stammen aus der ersten Querschnittserhebung der „Studie zur Mentalen Gesundheit bei der Arbeit“ (S-MGA), welche durch die BAuA initiiert und in Zusammenarbeit mit dem Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit und dem Institut für angewandte Sozialwissenschaft GmbH (infas) durchgeführt wurde. Ziel der Studie ist die Entwicklung einer empirischen Basis zur Arbeitsfähigkeit und funktionellen Gesundheit anhand einer Repräsentativerhebung an der Erwerbsbevölkerung in Deutschland mit gleichzeitiger Erfassung der Arbeits- und Beschäftigungsbedingungen als wesentliche Einflussfaktoren. Es sollen weiterhin für die zukünftige Forschung Referenzwerte für die verwendeten Instrumente und deren Validierung bereitgestellt werden. Einen Überblick über die in Zusammenarbeit der BAuA und infas entwickelten Inhalte der Befragung gibt Tabelle 3.1. Eine umfangreiche Darstellung der Planung und Durchführung der Studie und zum vorab durchgeführten Pretest sowie der im Folgenden angerissenen Merkmale der Datenerhebung kann bei Rose et al. (eingereicht) sowie im Methodenbericht von Schröder, Schiel, Schulz und Kleudgen (2015) nachgelesen werden, auf deren Grundlage die Beschreibung der Studie in diesem Kapitel erfolgt. Eine genaue Darstellung der Messinstrumente für die Analysen der vorliegenden Arbeit erfolgt im nachfolgenden Kapitel 3.2.

Die Grundgesamtheit für die Datenerhebung bildeten alle in Deutschland sozialversicherungspflichtig Beschäftigten der Geburtsjahrgänge 1951 bis 1980, die zum 31. Dezember 2010 bei der Bundesagentur für Arbeit gemeldet waren. Ausgeschlossen wurden somit Beamte und Selbstständige. Eingeschlossen waren sowohl Vollzeit- als auch Teilzeitbeschäftigte ebenso wie Minijobber mit einem Monatseinkommen unter 400 Euro. Des Weiteren wurden Erwerbstätige mit Migrationshintergrund nicht ausgeschlossen, solange es keine sprachlichen Barrieren bei den Interviews gab. Die Stichprobenziehung fand anhand einer zweistufigen Auswahl statt. Aus einer Grundgesamtheit von 12 227 Gemeinden wurde zuerst zufällig eine bevölkerungsproportionale und repräsentative Gemeindestichprobe mit 206 Gemeinden gezogen. Anschließend wurde ebenfalls zufällig aus diesen Gemeinden eine Brutto-Stichprobe von 13 590 Adressen gezogen, um eine ausreichende Realisierung der angestrebten 4 500 Interviews zu erreichen.

Vor einer telefonischen Kontaktaufnahme wurde den Zielpersonen ein Anschreiben mit Datenschutzerklärung zugesandt. Neben den Informationen zur Studie und der Adressherkunft wurde im Anschreiben die Kontaktaufnahme durch den Interviewenden angekündigt sowie eine Entschädigung in Höhe von 10 Euro für die Teilnahme in Aussicht gestellt. In der sechsmonatigen Feldphase zwischen November 2011 und Juni 2012 wurden 4 549 Interviews an erwerbstätigen Personen im Alter zwischen 31 und 60 Jahren realisiert, von denen nach Überprüfung durch infas 4 511 auswertbare Fälle übrig blieben. Durchgeführt wurden die Interviews von 243 geschulten Intervie-

wenden anhand einer computergestützten persönlichen Befragung. Bei umfangreicheren Antwortmöglichkeiten wurde den befragten Personen ein Listenheft mit den Antwortvorgaben als Unterstützung vorgelegt. Zusätzlich hatten die befragten Personen einen dreiseitigen Fragebogen mit sensiblen Fragen zur psychosozialen Situation selbstständig auszufüllen, ohne dass der Interviewende die Angaben sehen konnte, um Einflüsse durch den Interviewenden zu minimieren (Rücklaufquote 92%). Die Interviewdauer betrug durchschnittlich 65.6 Minuten mit einer Spannweite von 13.8 bis 163.5 Minuten. Da ein Zeitraum von 11 bis 17 Monaten zwischen dem Stichtag des Erwerbsstatus 31. Dezember 2010 für die Stichprobenziehung und dem Zeitpunkt der Befragung lagen, waren einige der befragten Personen nicht mehr erwerbstätig. Dadurch entfielen einige Fragenblöcke für diese Personen, das Interview selbst wurde jedoch durchgeführt, um eventuelle Zusammenhänge der Nichterwerbstätigkeit mit gesundheitlichen Gründen zu erfassen.

**Tab. 3.1** Inhalte der Studie zur Mentalen Gesundheit bei der Arbeit nach Rose et al. (2017)

Thematik	Indikatoren (Auszug)
Sozio-demografische Angaben	Geschlecht, Alter, Schul- und Berufsausbildung, beruflicher Status, Familienstand, Anzahl Kinder, Einkommen, Migrationsstatus
Beschäftigungsbedingungen	Beschäftigungsstatus, Vertragsverhältnis, Veränderung in der Organisation, Weiterbildungen und Beschäftigungsmaßnahmen
Arbeitsbedingungen	Arbeitszeit/-muster, körperliche Arbeitsbedingungen, psychische Belastungen durch die Arbeit (z. B. quantitative und kognitive Anforderungen, Entscheidungsspielraum, Rollenklarheit, Arbeitsplatzunsicherheit), Mobbing durch Kollegen
Arbeits- und Funktionsfähigkeit	Arbeitsfähigkeit (Work Ability Index – WAI), Funktionsfähigkeit (Norwegian Functioning Assessment Scale – NFAS), gesundheitsbezogene Lebensqualität (Short-Form-12 Health Survey – SF-12)
Motivation und Willenskraft	Selbstwirksamkeit, Arbeitsengagement, Überlegung den Arbeitsplatz/Arbeitgeber zu wechseln, Überlegungen zum Berufsausstieg
Persönliche Faktoren	Soziale Unterstützung, kritische Lebensereignisse, Gesundheitsverhalten
Psychische Gesundheit	Erholungsunfähigkeit, Arbeits- und Lebenszufriedenheit, depressive Symptomatik, Burnout-Symptomatik, positive und negative Affektivität

Die Güte der gezogenen Stichprobe wurde anhand der Strukturmerkmale der befragten Personen im Vergleich zur Grundgesamtheit bestimmt. Diese umfassten das Geschlecht, den Geburtsjahrgang, die Lage der Betriebsstätte (Ost/West), die Ausbildung, geringfügige Beschäftigung, die berufliche Stellung, die Nationalität, das Tagesentgelt und die Berufsklassifikation. Insgesamt ist die Stichprobengüte im Vergleich zur Grundgesamtheit sehr hoch einzuschätzen. In den Repräsentativitätsana-

lysen zwischen der im Register gezogenen Zufallsstichprobe und der im Feld realisierten Stichprobe der Interviewten ergeben sich lediglich geringe Abweichungen in den soziodemografischen Strukturmerkmalen. Somit stellt die vorliegende Stichprobe ein weitgehend unverzerrtes Abbild der Population der sozialversicherungspflichtig Beschäftigten zwischen 31 und 60 Jahren in Deutschland dar.

Die S-MGA ist als Panelstudie angelegt worden, um bei der Zweiterhebung im Jahr 2017 Zusammenhänge zwischen gesundheitlichen Veränderungen der befragten Personen und der von ihnen bei der ersten Erhebung genannten Arbeitsbedingungen und des Gesundheitsverhaltens zu ermitteln. Die Bereitschaft zur wiederholten Befragung und zur Adressspeicherung wurde schriftlich eingeholt (Zustimmung 87%). Bei Vorliegen des schriftlichen Einverständnisses wurden die Daten der Erhebung aus den Interviews mit den Prozessdaten der Bundesagentur für Arbeit zur Untersuchung der vorausgegangenen und nachgelagerten Beschäftigungshistorie gekoppelt (Zustimmung 75%). Dies ermöglicht u. a. eine Analyse der Vorhersagekraft der erhobenen Daten bezogen auf einen vorzeitigen Erwerbsausstieg.

Die Nutzung der erhobenen Daten ist an strenge Datenschutzrichtlinien gebunden und setzt einen Antrag auf wissenschaftliche Nutzung voraus, welcher das Forschungsziel, die Relevanz der Fragestellung für das soziale Sicherungssystem sowie den methodischen Forschungsansatz enthalten muss. Die Analysen erfolgten unter Beachtung der Auflagen zur Datensicherung mittels eines faktisch anonymisierten Datensatzes (Scientific Use File).

## **3.2 Messinstrumente**

Wie bereits im vorangegangenen Kapitel 3.1 beschrieben, wurde in der Studie eine Vielzahl von Variablen mit unterschiedlichen standardisierten Verfahren erfasst. In den folgenden Absätzen werden lediglich jene Verfahren näher beschrieben, welche für die Beantwortung der Fragestellungen in der vorliegenden Arbeit relevant sind. Ergänzt werden diese Verfahren durch Angaben zum soziodemografischen Status, wie z. B. dem Alter, dem Geschlecht und dem Berufs- und Ausbildungsstatus. Im Anhang A können aufgrund des Urheberrechts lediglich die Items des frei verfügbaren WAI dargestellt werden. Ein Überblick über die Skalen und Beispiele von Items werden in den Abschnitten der Messinstrumente gegeben.

### **3.2.1 Work Ability Index (WAI)**

Im Fokus der vorliegenden Arbeit steht als Messinstrument der frei verfügbare deutschsprachige WAI (Tuomi et al., 1998), welcher die Arbeitsfähigkeit anhand von sieben WAI-Dimensionen (siehe Kapitel 2.4) mit jeweils einem bis drei Items in einem grundsätzlich ungewichteten Summenindex abbildet. Die einzige Gewichtung erfolgt bei der Berechnung der Dimension (2) Arbeitsfähigkeit im Vergleich zu den Anforderungen der Arbeitstätigkeit, bei der zwei Gewichtungsfaktoren je nach Art der Tätigkeit (geistig, körperlich oder sowohl geistig als auch körperlich) in die Berechnung eingehen. Eine theoretische Einführung zum WAI als subjektives Screening-Instrument zur Bewältigungsfähigkeit der Arbeitsanforderungen wurde bereits in Kapitel 2.4 gegeben. Für die S-MGA wurde der WAI in einer leicht modifizierten Version der Kurzform mit kurzer Krankheitsliste (Nübling et al., 2005) präsentiert. Der Unterschied der verwendeten Version zur originalen Kurzform des WAI besteht in der Formulierung der Items in der Krankheitsliste für die Dimension (3) Anzahl der aktuellen vom Arzt diagnostizierten Krankheiten. Statt allgemein nach den durch einen

Arzt diagnostizierten Krankheiten oder Verletzungen zu fragen, wurden in einem ersten Schritt die jemals festgestellten Krankheiten erfragt und anschließend differenziert, ob dies innerhalb der letzten vier Wochen oder innerhalb der letzten 12 Monate stattfand oder bereits länger her ist. Relevant für die Analysen waren lediglich die Krankheiten der letzten 12 Monate. Insgesamt sind die Items des WAI sehr heterogen mit ebenfalls diversen Antwortformaten unterschiedlicher Metrik (metrisch und ordinal). Zum Beispiel wird die Dimension (1) Derzeitige Arbeitsfähigkeit im Vergleich zu der besten je erreichten Arbeitsfähigkeit mit dem Item „Wenn Sie Ihre beste, je erreichte Arbeitsfähigkeit mit 10 Punkten bewerten: Wie viele Punkte würden Sie dann für Ihre derzeitige Arbeitsfähigkeit geben? Null bedeutet, dass Sie derzeit arbeitsunfähig sind.“ auf einer 11-stufigen Likert-Skala erfragt (0 - völlig arbeitsunfähig bis 10 - die beste jemals erreichte Arbeitsfähigkeit). Dagegen wird die Dimension (7) Psychische Leistungsreserven anhand von drei Items (z. B. „Haben Sie in der letzten Zeit Ihre täglichen Aufgaben mit Freude erledigt?“) erfasst, welche jeweils ein Antwortformat von 0 - immer bis 4 - niemals aufweisen. Zur Berechnung der Dimension (7) wird aus den drei Antworten die Summe der angekreuzten Werte von 0 bis 12 gebildet, welche gemäß Handbuch (WAI-Netzwerk, 2015) als Grundlage für die Vergabe von 1 bis 4 Punkten dient. Die ausführlichen Items der verwendeten WAI-Version inklusive der Krankheitsliste kann dem Anhang A entnommen werden. Einen ersten Überblick über die Struktur und Berechnung des WAI gibt die Tabelle 3.2. Insgesamt kann beim klassischen WAI bei einer postulierten eindimensionalen Struktur ein Summenwert zwischen 7 und 49 erreicht werden, wobei höhere Werte eine bessere Arbeitsfähigkeit anzeigen.

Hinsichtlich der Gütekriterien kann von einer eingeschränkten Objektivität ausgegangen werden. Die Durchführungsobjektivität ist nur bedingt gegeben, da es unterschiedliche, nicht festgeschriebene Möglichkeiten der Datenerhebung ohne standardisierte Instruktionen gibt. So kann der WAI z. B. als Fragebogen vor einem Gespräch beim Betriebsarzt oder als Interview durch diesen erfasst werden (WAI-Netzwerk, 2015). Die Auswertungsobjektivität ist durch die festgelegte Berechnung des Index gegeben. Die Objektivität der Interpretation ist abhängig vom Hintergrund der Befragung (Forschung oder BGM) und den Zielen des Betriebes.

Ausgehend von einer einfaktoriellen Struktur werden als Maß der Reliabilität interne Konsistenzen (Cronbachs Alpha) für verschiedene Stichproben von  $\alpha = .58$  bis  $\alpha = .77$  (Martus et al., 2010),  $\alpha = .77$  (Abdolalizadeh et al., 2012) und  $\alpha = .83$  (Bethge et al., 2012) berichtet. In den europaweiten Analysen von Radkiewicz und Widerszal-Bazyl (2005) fanden sich interne Konsistenzen von  $\alpha = .54$  für die Slowakei bis zu  $\alpha = .78$  für Deutschland und  $\alpha = .79$  für Finnland (Total  $\alpha = .72$ ). Unter der Annahme einer dreifaktoriellen Struktur berichten Martinez et al. (2009) und Peralta et al. (2012) ebenfalls Cronbachs Alpha als Maß der internen Konsistenz ( $\alpha = .72$  und  $\alpha = .80$ ) des Gesamtindex. Dabei ist zu beachten, dass Cronbachs Alpha lediglich dann ein Maß für die wahre Reliabilität ist, wenn das Modell essentiell tau-äquivalent ist, d. h. die Ladungen der Items innerhalb eines Faktors bei unkorrelierten Residuen gleich sind, was selten geprüft wird. Ist diese Annahme verletzt, ist Cronbachs Alpha als Maß der internen Konsistenz lediglich eine untere Schranke der Reliabilität, sie wird unterschätzt (Steyer & Eid, 2001). Bei der Bewertung der internen Konsistenz sollte außerdem die Heterogenität und Kürze des Instruments beachtet werden. Eine korrekte Interpretation von Cronbachs Alpha setzt weiterhin Eindimensionalität der Skala voraus (Eid et al., 2013). Dadurch ist das Berichten der o. g. Cronbachs Alpha-Werte für den Gesamtindex in den Studien mit gefundener Mehrdimensionalität nur schwer nachvollziehbar.

**Tab. 3.2** Berechnung des Work Ability Index gemäß Handbuch (WAI-Netzwerk, 2015)

Dimension	Anzahl Items	Punkteverteilung der Antworten								
1 Derzeitige Arbeitsfähigkeit im Vergleich zu der besten je erreichten Arbeitsfähigkeit	1	0-10 Punkte								
2 Arbeitsfähigkeit im Vergleich zu den Anforderungen der Arbeitstätigkeit	2	2-10 Punkte Gewichtung der Punkte entsprechend dem Arbeitsinhalt (Formel siehe Manual)								
3 Anzahl der aktuellen vom Arzt diagnostizierten Krankheiten	Liste von 14 Krankheitsgruppen	<table> <tr> <td>≥ 4 Krankheiten</td> <td>= 1 Punkt</td> </tr> <tr> <td>2-3 Krankheiten</td> <td>= 3 Punkte</td> </tr> <tr> <td>1 Krankheit</td> <td>= 5 Punkte</td> </tr> <tr> <td>0 Krankheiten</td> <td>= 7 Punkte</td> </tr> </table>	≥ 4 Krankheiten	= 1 Punkt	2-3 Krankheiten	= 3 Punkte	1 Krankheit	= 5 Punkte	0 Krankheiten	= 7 Punkte
≥ 4 Krankheiten	= 1 Punkt									
2-3 Krankheiten	= 3 Punkte									
1 Krankheit	= 5 Punkte									
0 Krankheiten	= 7 Punkte									
4 Geschätzte Beeinträchtigung der Arbeitsleistung durch die Krankheiten	1	1-6 Punkte								
5 Krankenstandstage in den vergangenen 12 Monaten	1	1-5 Punkte								
6 Einschätzung der eigenen Arbeitsfähigkeit in zwei Jahren	1	1, 4 oder 7 Punkte								
7 Psychische Leistungsreserven	3	Summation der angekreuzten Werte: Summe 0-3 = 1 Punkt Summe 4-6 = 2 Punkte Summe 7-9 = 3 Punkte Summe 10-12 = 4 Punkte								

Hinweise zur Validierung wurden bereits in den Kapiteln 2.4 und 2.5 gegeben und beziehen sich sowohl auf die Kriteriumsvalidität (z. B. Eskelinen et al., 1991; Tuomi et al., 1997), die konvergente Validität (z. B. Abdolizadeh et al., 2012; Alexopoulos et al., 2013) als auch auf die Faktorenstruktur, welche eine der Hauptfragestellungen der vorliegenden Arbeit bildet. Die konvergente Validität wurde u. a. durch die Zusammenhänge des WAI mit den Dimensionen des SF-36 gestützt, welche Werte aufzeigen von  $r = .51$  für die Slowakei bis zu  $r = .67$  für Deutschland und  $r = .72$  für Finnland bezogen auf die allgemeine Gesundheit (Radkiewicz & Widerszal-Bazyl, 2005) sowie von  $r = .25$  bis  $r = .65$  für die einzelnen Gesundheitsdimensionen in den unterschiedlichen Stichproben (Abdolizadeh et al., 2012; Alexopoulos et al., 2013; Martinez et al., 2009). Weitere Zusammenhänge zu anderen Kenngrößen der Arbeits- und Funktionsfähigkeit werden im WAI-Manual zusammengefasst (WAI-Netzwerk, 2005).

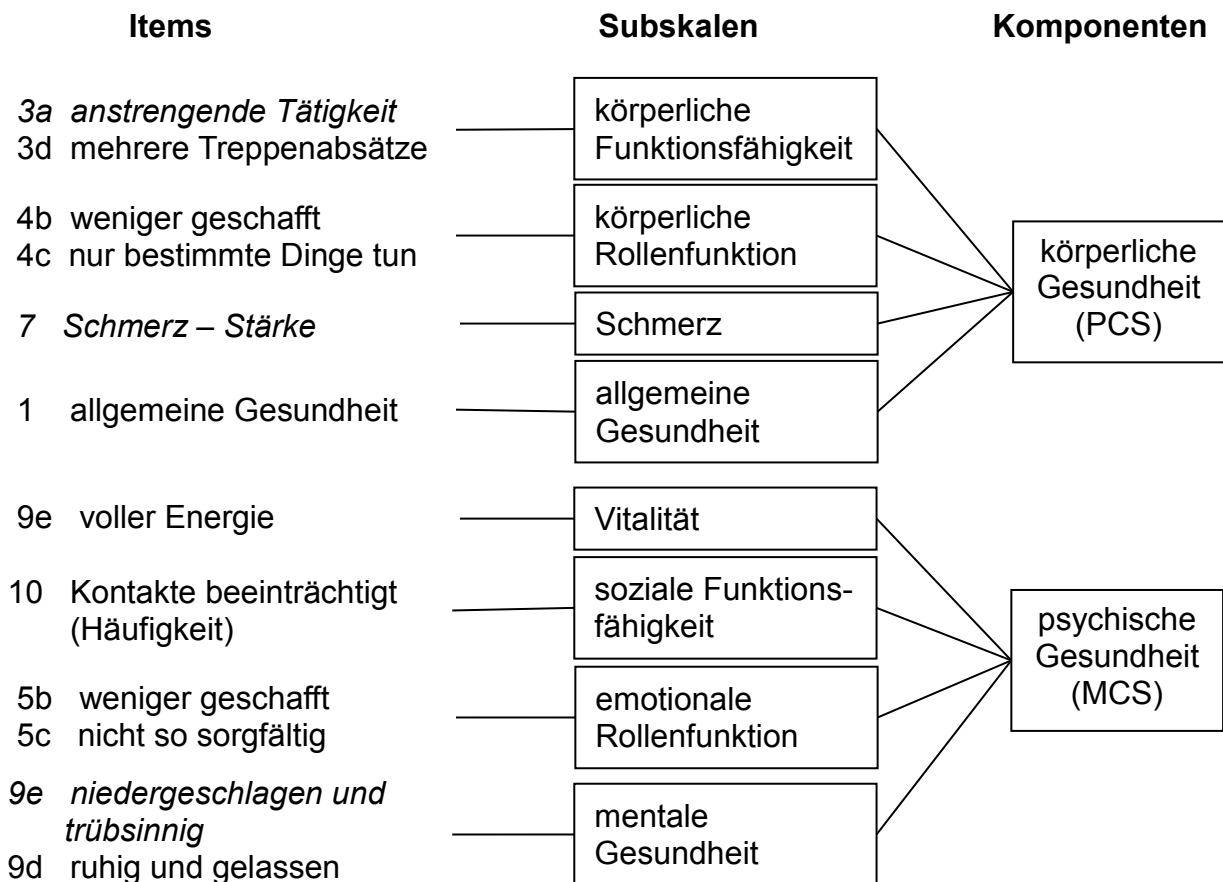


### 3.2.2 Short-Form-12 (SF-12) Health Survey

Zur Überprüfung der konvergenten Validität des WAI wurde u. a. eine deutschsprachige modifizierte Fassung der zweiten Version des SF-12 genutzt, welcher bereits im Kapitel 2.3 dargestellt wurde (Bullinger, Kirchberger & Ware, 1995; Ware et al., 1996; Ware, Kosinski, Turner-Bowker & Gandek, 2002). Das Instrument erfasst die selbstberichtete gesundheitsbezogene Lebensqualität für die physische und psychische Gesundheit über acht Subskalen mit jeweils einem bis zwei Items. Die Auswahl der Items für den SF-12 ergab sich aus Regressionsanalysen der Items des SF-36, bei denen diejenigen Items ausgewählt wurden, welche mindestens 90% der Varianz für die beiden Komponenten im SF-36 erklärten (Ware et al., 1996). Die Kreuzvalidierung des SF-12 erfolgte am SF-36 in verschiedenen Ländern. Dabei werden hohe Korrelationen von  $r = .94$  bis  $r = .97$  für die Komponenten körperliche Gesundheit und psychische Gesundheit zwischen dem SF-12 und dem SF-36 berichtet (Gandek et al., 1998). Die Verwendung des SF-12 wird empfohlen, wenn der Fokus der Forschung auf den zwei übergeordneten Komponenten statt auf den Subskalen liegt und die Stichprobengröße mit  $> 500$  ausreichend groß ist (Gandek et al., 1998).

Die Modifizierung der zweiten Version des SF-12 fand im Rahmen des Sozioökonomischen Panels 2004 statt (Nübling, Andersen & Mühlbacher, 2006) und umfasst neben der Reihenfolge und Formulierung der Fragen auch den Austausch des Items zur Behinderung durch Schmerz gegen ein Item aus dem SF-36 (Ware, 2000) zu starken körperlichen Schmerzen. Im Rahmen der S-MGA fanden zwei weitere Auswechselungen mit Items aus dem SF-36 statt (siehe Abbildung 3.1). Die Komponente körperliche Gesundheit (PCS - Physical Component Summary) beinhaltet die vier Subskalen Allgemeine Gesundheit (ein Item), körperliche Funktionsfähigkeit (zwei Items), körperliche Rollenfunktion (zwei Items) sowie körperliche Schmerzen (ein Item). Die zweite Komponente psychische Gesundheit (MCS - Mental Component Summary) umfasst die vier Subskalen Vitalität (ein Item), soziale Funktionsfähigkeit (ein Item), emotionale Rollenfunktion (zwei Items) und mentale Gesundheit (zwei Items). Die Antworten werden jeweils auf einer 3- oder 5-stufigen Ratingskala erfasst.

Die Berechnung der Subskalen und Komponenten erfolgt nach den Vorgaben von Nübling et al. (2006). Nachdem die Subskalen auf eine Metrik im Range von 0-100 transformiert und gemittelt wurden, erfolgte eine z-Standardisierung anhand der Stichprobenkennwerte der S-MGA für die einzelnen Subskalen sowie ein „norm-based scoring“, woraus sich standardisierte Werte mit einem Mittelwert von 50 ( $SD = 10$ ) ergeben. Niedrigere Werte indizieren eine geringere gesundheitsbezogene Lebensqualität. Die Komponenten PCS und MCS werden durch die gewichteten und summierten Faktorwerte mit jeweils allen acht standardisierten Subskalen aggregiert. Die Gewichte entstammen dabei einer Hauptkomponentenanalyse mit Varimax-Rotation und werden als Faktorwert-Koeffizienten anhand der Regressionsmethode geschätzt (Bühner, 2011). Für die modifizierte Version des SF-12 liegen zum jetzigen Zeitpunkt keine Daten zur Reliabilität und Validität vor. Studien zur Originalversion des SF-12 beschreiben den Fragebogen als reliabel und valide für die überprüften Stichproben. So werden interne Konsistenzen von  $\alpha = .89$  und  $\alpha = .86$  (Cronbachs Alpha) sowie  $\alpha = .88$  und  $\alpha = .82$  (Mosier Alpha; Mosier, 1943) für PCS und MCS berichtet sowie eine prädiktive Validität für die zukünftige physische und psychische Gesundheit (Cheak-Zamora, Wyrwich & McBride, 2009; Ware et al., 2002).



**Abb. 3.1** Struktur der Items und Komponenten des Short-Form-12 Health Survey in der Studie (kursive Items ausgewechselt gegenüber der Originalversion von Ware et al., 1996).

### 3.2.3 Norwegian Function Assessment Scale (NFAS)

Als zweites Instrument zur Überprüfung der konvergenten Validität wurde die auf der ICF basierende NFAS verwendet, welche ebenfalls bereits in Kapitel 2.3 kurz vorgestellt wurde (Brage et al., 2004; Østerås et al., 2007). Zu diesem Zweck wurde für die S-MGA in einem mehrstufigen Verfahren eine deutsche Übersetzung erstellt. Anhand von 39 Items werden Angaben zur körperlichen und mentalen Funktionsfähigkeit im Arbeitsleben und bezogen auf Freizeitaktivitäten erfasst, welche sich auf sieben Dimensionen abbilden lassen (Brage et al., 2004): Gehen/Stehen mit sieben Items (z. B. „Hatten Sie in der letzten Woche Schwierigkeiten Treppen zu steigen?“), Halten/Aufheben mit acht Items (z. B. „...Auto zu fahren?“), Heben/Tragen mit sechs Items (z. B. „...Einkaufstaschen mit den Händen zu tragen?“), Sitzen mit drei Items (z. B. „...auf einem Küchenstuhl zu sitzen?“), Bewältigen mit sieben Items (z. B. „...in einer Gruppe zu arbeiten?“), Zusammenarbeit/Kommunikation mit sechs Items (z. B. „...schriftliche Mitteilungen zu verstehen?“) und Sinnesorgane mit zwei Items (z. B. „...Fernsehen zu schauen?“). Antworten konnten auf einer 5-stufigen Likert-Skala (1 - keine Schwierigkeiten bis 5 - konnte ich nicht) gegeben werden, welche hinsichtlich der Datenqualität, der internen Konsistenz und der diskriminanten Validität der ursprünglich 4-stufigen Likert-Skala vorzuziehen ist (Østerås et al., 2008). Der Ge-

samtscore der NFAS errechnet sich aus dem Mittelwert aller beantworteten Items. Zusätzlich ist es möglich, die ersten vier und die letzten drei Dimensionen in die jeweils körperliche (PD) und mentale Domäne (MD) einzuordnen. Die Werte für diese Domänen werden ebenfalls durch Mittelwertbildung der beantworteten Items berechnet. Niedrigere Werte indizieren dabei eine bessere Funktionsfähigkeit. Da es möglich sein kann, dass bestimmte Aktivitäten auf einige Personen nicht zutreffen, werden keine Anforderungen an die Mindestmenge der beantworteten Fragen gestellt. Die Validierung des Originalinstruments mit einer norwegischen Stichprobe erfolgte am SF-36 und zeigte signifikante Zusammenhänge mit den entsprechenden Komponenten (Brage et al., 2004). Für die einzelnen Dimensionen werden, obwohl eine der Dimensionen lediglich zwei Items hat, interne Konsistenzen von  $\alpha = .67$  (Sitzen) bis  $\alpha = .91$  (Gehen/Stehen) und für den Gesamtwert  $\alpha = .95$  berichtet, wobei wiederum zu beachten ist, dass Cronbachs Alpha von der Itemzahl abhängig ist (Eid et al., 2013). Die Überprüfung der deutschen Version erfolgt im Rahmen eines weiteren Projektes bei der BAuA anhand der S-MGA-Stichprobe (Jankowiak & Kersten, 2016). Aufgrund der Analysen in diesem Projekt wird für die Auswertung der Konvergenzvalidität des WAI eine reduzierte Anzahl von Variablen herangezogen, welche sich ausschließlich auf arbeitsbezogene Items beschränken (NFAS-w). So lassen sich insgesamt 22 Items auf den folgenden fünf Dimensionen abbilden: Gehen/Stehen (fünf Items), Heben/Tragen (vier Items), Bewältigen (sieben Items) sowie einer Aufsplitterung der Dimension Zusammenarbeit/Kommunikation in zwei getrennte Dimensionen mit je drei Items. Auch hier ist eine Einordnung der ersten beiden Dimensionen in eine körperliche (w\_PD) sowie der letzten drei Dimensionen in eine mentale Domäne (w\_MD) möglich.

### 3.3 Stichprobenbeschreibung

Zur Auswertung der Fragestellungen lagen 4 511 auswertbare Fälle vor, welche bereits eine Datenüberprüfung durch infas durchlaufen hatten. Diese Überprüfung durch infas orientierte sich an der Interviewdauer, an widersprüchlichen Angaben sowie an schriftlichen Kontrollfragebögen, welche den befragten Personen nach Abschluss des Interviews zugesandt wurden. Von den auswertbaren Fällen wurden aus inhaltlichen Gründen einige Personen ausgeschlossen. So wurden die Daten von 310 Personen ausgeschlossen, welche hauptberuflich nicht mindestens in einem geringfügigen oder unregelmäßigen Arbeitsverhältnis standen. Sie sollten sich jedoch nicht in einer Beschäftigungs- oder Qualifizierungsmaßnahme befinden. Weiterhin ausgeschlossen wurden die Daten von 74 selbstständig oder freiberuflich tätigen Personen sowie von 159 weiteren nicht sozialversicherungspflichtig beschäftigten Personen.

Letztendlich wurden die Daten von  $N = 3\,968$  Personen (51.1% männlich) in die Analysen einbezogen, was 87.2% der realisierten Interviews und 88.0% der durch infas gelieferten Daten entspricht. Eine genaue Angabe zur Altersverteilung ist nicht möglich, da aufgrund des Anonymisierungskonzepts lediglich Daten in 5-Jahresstufen zwischen 31 und 60 Jahren vorlagen. Die größte der sechs Gruppen bilden in der Stichprobe die 46- bis 50-Jährigen mit 21.8%. Für die Berechnungen der Messinvarianz wurde das Alter der Teilnehmenden zusätzlich in drei Altersgruppen (31-40, 41-50 und 51-60 Jahre) zusammengefasst. Hier bilden mit 41.8% die 41- bis 50-Jährigen die größte Gruppe. Die meisten Personen (40.5%) gaben als höchsten Schulabschluss die mittlere Reife (Realschulabschluss, Fachschulreife, 10. Klasse Polytechnische Oberschule) an, gefolgt von 32.2% der Personen mit

(Fach-)Hochschulreife (Abitur, Fachoberschule, Erweiterte Oberschule). Als höchsten Ausbildungsabschluss gaben 50.7% den Abschluss einer betrieblichen Berufsausbildung an. Einen Hochschulabschluss gaben 12.2% an und 5.3% hatten keinen oder einen anderen Ausbildungsabschluss. Mit 72.6% ist der größte Teil der Stichprobe mit mindestens 35 Stunden pro Woche in Vollzeit erwerbstätig. Einen zeitlich unbefristeten Arbeitsvertrag gaben 93.0% der Personen an. Der größte Teil der Stichprobe ist in Deutschland geboren (89.4%). Eine detaillierte Darstellung der soziodemografischen Daten der Stichprobe kann der Tabelle 1 im Anhang 2 entnommen werden.

### 3.4 Datenauswertung und statistische Methoden

Die Voranalysen und die Datenauswertung zur Überprüfung der Hypothesen fanden mit Hilfe der Statistikprogramme IBM SPSS Statistics (Version 23) und Mplus (Version 7.4; Muthén & Muthén, 1998-2012) statt. Bei den Voranalysen wurden neben der Erstellung der deskriptiven Ergebnisse außerdem die Voraussetzungen für die Durchführung der konfirmatorischen Faktorenanalysen geprüft. Laut Kline (2011) und Bühner (2011) betrifft dies vor allem die Kollinearität, Ausreißer, multivariate Normalverteilung, Stichprobengröße sowie fehlende Werte.

#### 3.4.1 Prüfung der Voraussetzungen

Grundsätzlich ist eine multivariate Normalverteilung Voraussetzung für die Durchführung von konfirmatorischen Faktorenanalysen mittels der standardmäßig angewandten Maximum-Likelihood-Methode. Nach Inspektion der Histogramme und Q-Q-Plots für die analysierten Items der sieben WAI-Dimensionen (nachfolgend WAI1 bis WAI7) konnte für kein Item von einer univariaten Normalverteilung ausgegangen werden, welche die Voraussetzung für eine multivariate Normalverteilung ist (Kline, 2011). Alle sieben Items wiesen zudem eine negative Schiefe (linksschief) von bis zu -1.93 (SD = 0.04) auf. Der von SPSS ausgegebene Test auf Normalverteilung (Kolmogorov-Smirnov-Test) konnte lediglich für die beiden metrisch skalierten Items WAI1 und WAI2 interpretiert werden (Eid et al., 2013). Er deutete ebenfalls auf die Verletzung der univariaten Normalverteilungsannahme für beide Items hin ( $p < .001$ ), wobei bedacht werden sollte, dass dieser Test bei großen Stichproben äußerst sensitiv ist und schnell signifikant wird (Field, 2013). Aufgrund der fehlenden multivariaten Normalverteilung und der ordinalen Metrik der Items WAI3 bis WAI7 wurde für die konfirmatorischen Faktorenanalysen auf ein robustes Schätzverfahren zurückgegriffen. Das WLSMV-Schätzverfahren (mean- and variance-adjusted weighted least squares) gehört zu der Gruppe der Asymptotically Distribution-Free Methoden und ist in Mplus bei der Analyse kategorialer Variablen als Grundeinstellung gegeben (Muthén & Muthén, 1998-2012). Auch inhaltlich kann für die Items des WAI weder von einer Normalverteilung noch von der Abwesenheit von Ausreißern ausgegangen werden, da (noch) erwerbstätige Personen befragt wurden. Daher wurden die anhand der Boxplots analysierten Ausreißer aus inhaltlichen Gründen nicht ausgeschlossen.

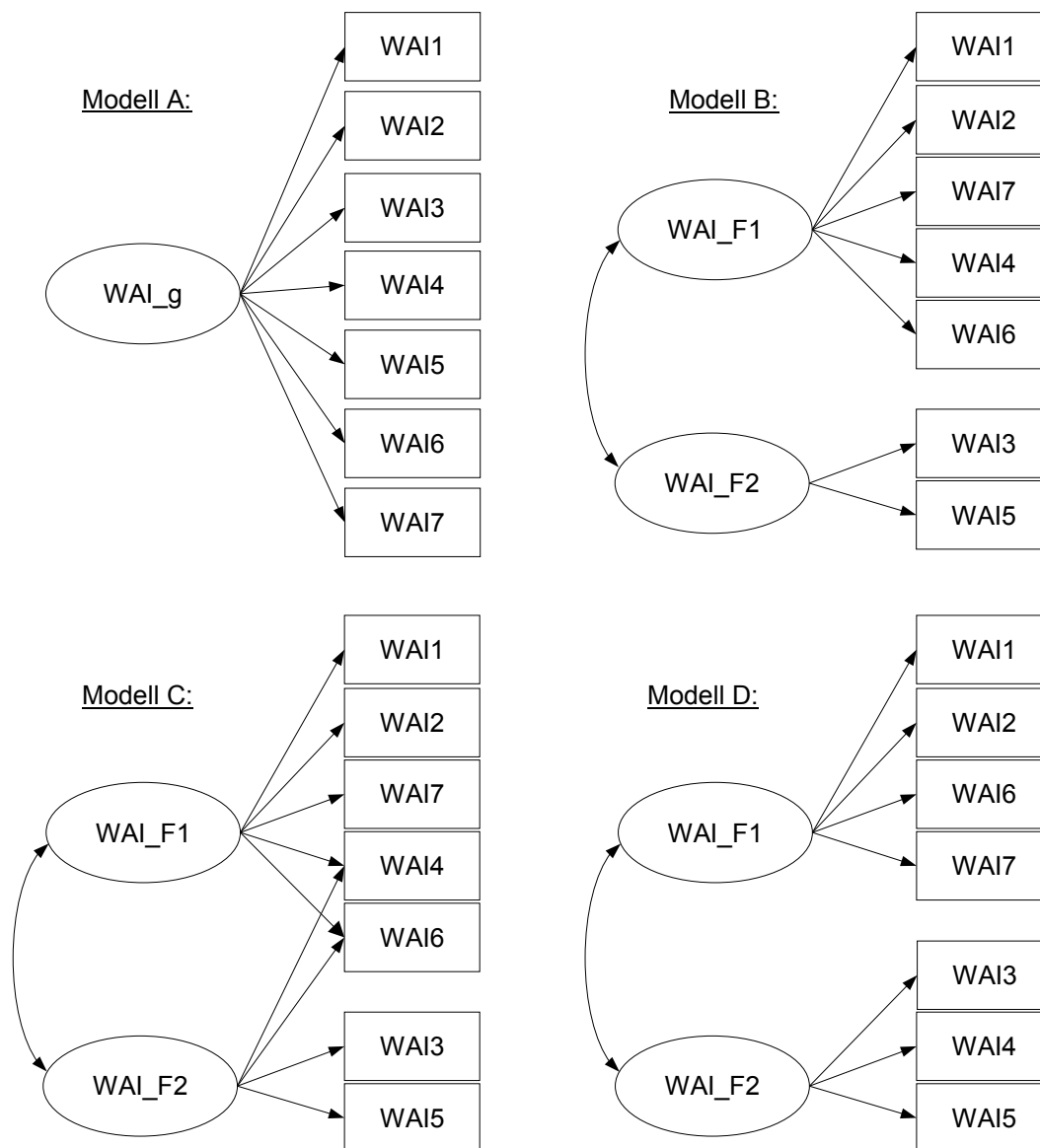
Die Überprüfung der Kollinearität erfolgte anhand der Betrachtung der Korrelationen der Items untereinander sowie der Indikatoren  $R^2_{smc}$  als quadrierte multiple Korrelation jedes einzelnen Items mit den übrigen Items und dem Varianz-Inflations-Faktor (VIF) gemäß Bühner (2011) und Kline (2011). Für die Analyse der bivariaten Zusammenhänge zwischen den Items wird eine polychorische Korrelation ( $r_{pol}$ ) zugrun-

de gelegt. Dabei wird unterstellt, dass eine beobachtete Variable mit ordinaler Metrik aus einer metrischen und normalverteilten Variable resultiert, was durch die Metrik des Items nicht abzubilden war (Eid et al., 2013). Sämtliche bivariaten Korrelationen der Items übertraten nicht die Schwelle von  $r_{\text{pol}} > .85$  (Bühner, 2011). Weder  $R^2_{\text{smc}}$  noch der VIF überstiegen die von Kline (2011) angegebenen Grenzen von  $R^2_{\text{smc}} > .90$  und  $\text{VIF} > 10$ . Die höchsten Werte lieferte der WAI1 als abhängige Variable mit  $R^2_{\text{smc}} = .44$  und  $\text{VIF} = 1.79$ .

Die Stichprobengröße übertrifft mit  $N = 3\,968$  Personen die von Bühner (2011) angestrebte Mindestgröße für konfirmatorische Faktorenanalysen von  $N > 250$  bei weitem. Im Gegensatz zu nicht robusten WLS-Verfahren zeigt das robuste WLSMV-Verfahren sowohl für sehr große ( $N = 1\,000$ ) als auch für kleinere ( $N = 100$ ) Stichproben zuverlässige Schätzungen (Flora & Curran, 2004). Fehlende Werte lagen für die interessierenden Variablen lediglich in geringem Maße vor. So würden bei einem listenweisen Fallausschluss bei der Analyse der Items des WAI lediglich 2.5% ( $n = 98$ ) der Probanden ausgeschlossen werden. Für die Analyse der übrigen Skalen würden bedeutend weniger Probanden (maximal 0.4%) ausgeschlossen werden. Daher wurde bei Berechnungen mit Mplus der Fallausschluss paarweise durchgeführt. Für die übrigen Analysen in SPSS wurden die Personen mit fehlenden Daten durch listenweisen Fallausschluss nicht beachtet, da die Ausschlussquote weniger als 5% beträgt (Graham, 2009).

### 3.4.2 Konfirmatorische Faktorenanalysen

Die Durchführung der konfirmatorischen Faktorenanalysen zur Modellüberprüfung gemäß der Hypothesen 1a und 1b orientierte sich an den Vorgaben von Geiser (2011) und Kline (2011). Ziel der durchgeführten Analysen war es, aus vorher festgelegten spezifizierten Modellen dasjenige Modell zu benennen, welches unter empirischen und theoretischen Gesichtspunkten am besten auf die erhobenen Daten passt. Die Spezifikation der Modelle erfolgte angelehnt an Martus et al. (2010) sowie Radkiewicz und Widerszal-Bazyl (2005). Daraus sind vier zu überprüfende und identifizierte Modelle entstanden, deren Strukturen in Abbildung 3.2 veranschaulicht sind. Das Modell A entstammt der ursprünglichen Annahme bei der Entwicklung des WAI und beinhaltet lediglich einen Faktor WAI\_g als latente Variable, welcher die Ausprägungen der sieben manifesten Variablen (Items WAI1 bis WAI7) erklärt. Die Modelle B und C entsprechen den Modellen C und D bei Martus et al. (2010). Im Modell B werden jeweils die Ausprägungen der Items WAI1, WAI2, WAI4, WAI6 und WAI7 sowie WAI3 und WAI5 durch zwei korrelierende Faktoren WAI\_F1 und WAI\_F2 erklärt. Im Modell C werden zusätzlich Doppelladungen der Items WAI4 und WAI6 auf beide Faktoren zugelassen. Modell D ist eine Modifikation von Modell B, in dem das Item WAI4 auf dem Faktor WAI\_F2 statt auf dem Faktor WAI\_F1 lädt. Es wird angenommen, dass die Ausprägungen der Items WAI1, WAI2, WAI6 und WAI7 durch einen Faktor der subjektiv eingeschätzten Arbeitsfähigkeit und Ressourcen (WAI\_F1) sowie die Items WAI3, WAI4 und WAI5 durch einen gesundheitsbezogenen und größtenteils objektiven Faktor (WAI\_F2) erklärt werden. Die Ladungen für alle Items wurden frei geschätzt, wodurch zur Identifizierbarkeit die Metrik des Faktors durch Fixierung der Faktorvarianz auf 1 festgelegt wurde (Geiser, 2011). Für ordinale Variablen wird in Mplus die polychorische Korrelationsmatrix statt einer Kovarianzmatrix zur Schätzung mit dem WLSMV-Verfahren herangezogen (Muthén & Muthén, 1998-2012).



**Abb. 3.2** Konfirmatorisch zu prüfende Modelle.

Zur Überprüfung der Modellgüte, d. h. ob das geprüfte Modell auf die empirischen Daten passt, wurden eine Reihe von Indikatoren herangezogen. Zum einen wurde standardmäßig der  $\chi^2$ -Test (Chi<sup>2</sup>-Test) berichtet, welcher einen Signifikanztest zur Überprüfung der Modellpassung darstellt (Kline, 2011). Ist der  $\chi^2$ -Test signifikant, wird die Nullhypothese der exakten Modellpassung abgelehnt. Allerdings ist dieser Test sowohl abhängig von der Stichprobengröße als auch sensitiv gegenüber Verletzungen der multivariaten Normalverteilungsannahme (Schermelleh-Engel, Moosbrugger & Müller, 2003). So wird die Teststatistik bei sehr großen Stichproben auch bei gut passenden Modellen signifikant und die Nullhypothese abgelehnt (Kline, 2011). Daher wurden neben dem  $\chi^2$ -Test ebenfalls der Root-Mean-Square-Error of Approximation (RMSEA), der Comparative-Fit-Index (CFI) und der Tucker-Lewis-Index (TLI) betrachtet, welche bei der WLSMV-Schätzung in Mplus mit ausgegeben wurden. Sowohl der RMSEA als auch der CFI und der TLI ziehen zwar zur Berechnung das  $\chi^2$  heran, jedoch unter Berücksichtigung der Stichprobengröße bzw. der

Freiheitsgrade. In der Literatur werden verschiedene Richtwerte, die mehr oder weniger streng sind, für die Fit-Indizes herangezogen (Bühner, 2011; Hu & Bentler, 1999; Schermelleh-Engel et al., 2003). Meist zitiert werden Hu und Bentler (1999), wobei Marsh, Hau und Wen (2004) und Schermelleh-Engel et al. (2003) davor warnen, diese sogenannten „Faustregeln“/„Goldenen Regeln“ der Cut-Off Werte zu streng zu betrachten. Gründe dafür liegen in der Anfälligkeit gegenüber Missspezifikationen, der Stichprobengröße, den Verteilungseigenschaften und der Schätzverfahren (Hu & Bentler, 1998). So sollen die Fit-Indizes laut Marsh et al. (2004) eher dazu herangezogen werden, um verschiedene Modelle untereinander zu vergleichen.

Nichtsdestotrotz wurden für den Vergleich der spezifizierten Modelle folgende Richtwerte zur Beurteilung der Fit-Indizes herangezogen (nach Schermelleh-Engel et al., 2003): Ein RMSEA  $\leq$  .05 spricht für eine gute Modellpassung, ein RMSEA  $>$  .05 und  $\leq$  .08 für eine akzeptable Modellpassung. Zusätzlich gibt Mplus für den RMSEA ein 90%-Konfidenzintervall (KI) und einen p-Wert für die Nullhypothese RMSEA  $\leq$  .05 aus. Das KI gibt an, in welchem Bereich der RMSEA mit 90-prozentiger Wahrscheinlichkeit in der Population liegt (Christ & Schlüter, 2012). Ist der p-Wert signifikant, muss die Hypothese eines guten Modellfits verworfen werden. Sowohl der CFI als auch der TLI sollten für eine akzeptable Modellpassung einen Wert von  $\geq$  .95 und für eine gute Modellpassung einen Wert von  $\geq$  .97 aufweisen. Von einer empirisch geleiteten Verbesserung der Modelle mit Hilfe von Modifikationsindizes wird abgesehen, da sie dem konfirmatorischen Ansatz der Arbeit widerspricht. Zum Vergleich der Passung zwischen den spezifizierten und genesteten Modellen wurde aufgrund des WLSMV-Schätzverfahrens die von Mplus bereitgestellte Option DIFFTEST angewandt, da eine Berechnung des  $\chi^2$ -Differenztests auf die herkömmliche Weise anhand der Differenzen der  $\chi^2$ -Werte und Freiheitsgrade bei diesem Schätzverfahren nicht möglich ist (Muthén & Muthén, 1998-2012).

Für die korrekte Interpretation der Reliabilität mittels Cronbachs Alpha als Maß der internen Konsistenz wurde im Anschluss an die Modellüberprüfungen die Annahme der essentiellen tau-Äquivalenz in Mplus überprüft (Brown, 2006; Eid et al., 2013). Dabei wird die Güte eines Modells mit gleichen Faktorladungen bei unkorrelierten Fehlern innerhalb eines Faktors überprüft. Die Berechnung von Cronbachs Alpha basiert auf der Kovarianzmatrix der Items für metrische Variablen, während für ordinale Daten die polychorische Korrelationsmatrix relevant ist. Diese dient als Grundlage für die Berechnung des ordinalen Alphas ( $\alpha_{\text{pol}}$ ) für die Faktoren des WAI (Gadermann, Guhn & Zumbo, 2012). Für die Vergleichbarkeit mit vorherigen und zukünftigen Studien wird zusätzlich das klassische Cronbachs Alpha berichtet.

### 3.4.3 Multiple Gruppenvergleiche

Die Überprüfung der Messinvarianz des WAI gemäß Hypothese 2 für das Geschlecht und Hypothese 3 für die drei Altersgruppen (31-40, 41-50 und 51-60 Jahre) wurde in Mplus nach dem Step-Up-Ansatz für multiple Gruppenvergleiche anhand von konfirmatorischen Faktorenanalysen in vier Schritten durchgeführt (Brown, 2006; Christ & Schlüter, 2012). Orientiert wurde sich dabei an dem spezifizierten Modell, welches sich in den konfirmatorischen Faktorenanalysen nach theoretischen und empirischen Gesichtspunkten als plausibel herausgestellt hat. So wurde, nach einer Überprüfung der oben genannten Modell-Fits für jede einzelne Gruppe im ersten Schritt, im am wenig restriktivsten Modell die konfigurale Messinvarianz mit freier Schätzung der Ladungen sowie der Intercepts für metrische bzw. Thresholds (Schwellen) für kate-

goriale Variablen geprüft. Aufgrund von Voreinstellungen in Mplus wurden dabei die Mittelwerte der latenten Faktoren in beiden Gruppen auf 0 fixiert. Sind die Modell-Fits in diesem Schritt akzeptabel, kann von konfigurationaler Messinvarianz ausgegangen werden und das Modell wird als Basismodell definiert. Sind die Modell-Fits nicht akzeptabel, sind keine weiteren Analysen notwendig und Mittelwertvergleiche zwischen den Gruppen sind nicht zulässig (Brown, 2006). Im dritten Schritt wurden zur Überprüfung der metrischen Messinvarianz die Ladungen für die jeweiligen Gruppen gleichgesetzt, die Intercepts bzw. Thresholds wurden weiterhin frei geschätzt. So entstand ein restriktiveres Modell mit mehr Freiheitsgraden. Sind die Modell-Fits akzeptabel und ist der  $\chi^2$ -Differenztest per DIFFTEST-Option im Vergleich mit dem Baseline-Modell nicht signifikant, kann von metrischer Messinvarianz ausgegangen und mit dem vierten Schritt fortgesetzt werden. Da jedoch auch hier der  $\chi^2$ -Differenztest sensitiv gegenüber großen Stichproben ist, wurde zusätzlich die Veränderung des CFI betrachtet ( $\Delta$ CFI). Wenn der CFI sich nicht um mehr als -.01 verschlechtert, kann von Invarianz der Modelle ausgegangen werden (Cheung & Rensvold, 2002). Weiterhin wurde betrachtet, ob der RMSEA vom restriktiveren Modell im 90%-KI des weniger restriktiven Modells liegt, was ebenfalls für die Invarianz der Modelle spricht (Little, Card, Slegers & Ledford, 2007). Zeigen sich jedoch Verschlechterungen bei den Parametern der Modellgüte, sollten schrittweise anhand der Modifikationsindizes diejenigen Ladungen freigesetzt werden, welche das  $\chi^2$  signifikant verbessern (Brown, 2006). Anschließend wird wiederum im Modellvergleich überprüft, ob die Modellgüte im Vergleich zum weniger restriktiven konfiguralen Modell nicht mehr verschieden ist. Dann kann man von partieller metrischer Messinvarianz sprechen, wobei nach Byrne et al. (1989) jedoch mindestens zwei Indikatoren je Faktor gleichgesetzt bleiben sollten. Im vierten Schritt wurden zur Überprüfung der skalaren Messinvarianz zusätzlich die Intercepts bzw. Thresholds in den Gruppen gleichgesetzt, die Gleichsetzung der latenten Mittelwerte aufgehoben und die Modelle analog zu Schritt drei verglichen und ggf. durch Freisetzen von Parametern angepasst. Im Fall partieller metrischer Messinvarianz bleiben auch im vierten Schritt die Ladungen der identifizierten nicht invarianten Indikatoren freigeschätzt. Erst wenn mindestens partielle skalare Messinvarianz zwischen den Gruppen vorliegt, sind Mittelwertvergleiche zulässig (Byrne et al., 1989). Die Ergebnisse der Mittelwertvergleiche durch die in Mplus ausgegebenen z-Werte werden in die Effektstärker  $r$  umgerechnet (Rosenthal & DiMatteo, 2001) und nach Cohen (1988) interpretiert.

#### 3.4.4 Gewichtung der Items

Ausgehend von den Ergebnissen der konfirmatorischen Faktorenanalysen unter der Annahme einer zweifaktoriellen Struktur des WAI sollen zwei gewichtete Summenindizes berechnet werden. In Kapitel 2.5 wurden die refined und non-refined Methoden zur Berechnung von Faktorwerten vorgestellt. Die genannten Nachteile des ungewichteten Summenindex implizieren eine Anwendung von refined Methoden durch Gewichtung der individuellen Indexwerte. Dazu gehört u. a. die gewichtete Berechnung der Faktorwerte durch die Regressionsmethode von Thurstone (1935; zitiert nach DiStefano et al., 2009, S. 4), welche nach Grice (2001) im Vergleich mit anderen Methoden der Berechnung von Faktorwerten die höchste Validität aufzeigte. Die Gewichtung individueller Indexwerte erfolgt dabei durch die sogenannten Faktorwertkoeffizienten (factor score coefficients). Dieses Vorgehen ist bei explorativen Faktorenanalysen in SPSS oder bei konfirmatorischen Faktorenanalysen in Mplus für metrische Variablen möglich. Für kategoriale Variablen können in Mplus lediglich die



Faktorwerte berechnet und ausgegeben werden, jedoch nicht die Faktorwert-Koeffizienten (B. Muthén, persönl. Mitteilung, 08.06.2016), welche für eine Darstellung einer anwendbaren Berechnungsformel relevant sind. Aus diesem Grund wurden die Gewichte für die Formel zur Summenindexbildung durch eine Annäherung mittels multipler Regressionsanalysen in SPSS berechnet, bei der die in Mplus berechneten Faktorwerte ( $FS_{Mplus}$ ) als Kriterium und die sieben WAI-Items als Prädiktoren für beide Faktoren eingegangen sind. Die daraus resultierenden standardisierten Regressionskoeffizienten ergeben eine Annäherung an die Faktorwert-Koeffizienten, welche zur Berechnung neuer Faktorwerte ( $F1_{Reg}$  und  $F2_{Reg}$ ) mit den z-standardisierten Item-Ausprägungen multipliziert wurden. Jedoch ist zu beachten, dass für die ordinalen Variablen in der linearen multiplen Regressionsanalyse ein metrisches Skalenniveau angenommen wird. Die korrekte Berechnung durch Dummy-Variablen wäre aufgrund der Anzahl der ordinalen Items allerdings nicht praktikabel.

Zur Überprüfung der Hypothese 4 wurde die Validität der neu berechneten Faktorwerte ( $F_{Reg}$ ) gegenüber zwei weiteren Möglichkeiten zur Berechnung von Faktorwert-Koeffizienten analysiert. Zum einen wurden in SPSS die Faktorwerte ( $F_{PCA}$ ) durch eine Hauptkomponentenanalyse mit zwei zu extrahierenden Faktoren und Varimax-Rotation berechnet, bei der ebenfalls von einem metrischen Skalenniveau der Items ausgegangen wird (Bühner, 2011). Zum anderen wurden in Mplus die Faktorwerte ( $F_{Mplus\_metr}$ ) unter der Annahme von metrischen Variablen mit Maximum-Likelihood-Schätzverfahren für alle WAI-Items berechnet. Der Zusammenhang der drei berechneten Faktorwerte ( $F_{Reg}$ ,  $F_{PCA}$  und  $F_{Mplus\_metr}$ ) mit dem exakten Faktorwert ( $FS_{Mplus}$ ) wurde mittels Zusammenhangsanalysen für beide Faktoren untersucht. Dabei wurde der Kendall-tau-b-Koeffizient analysiert, da für diese Zusammenhänge nicht die absolute Ausprägung des Faktorwerts, sondern der Unterschied in den Rängen gleicher Personen relevant ist.

Anhand der berechneten Faktorwerte werden getrennt für Männer und Frauen Normwerttabellen für die drei Altersgruppen dargestellt, um Anwendern des WAI repräsentative Vergleichswerte zu liefern. Aufgrund der fehlenden Normalverteilung der Daten und der guten Interpretierbarkeit werden für die erfassten und transformierten Rohwerte deren Prozentränge als Normwertskala berechnet. Dabei bedeutet ein individueller Prozentrang von z. B. 35, dass 35% der Personen aus der Eichstichprobe einen gleichen oder schlechteren Wert erreicht haben als diejenige Person mit dem entsprechenden Prozentrang. Anhand des ordinalen Alphas ( $\alpha_{pol}$ ), der Standardabweichung (SD) und des daraus resultierenden Standardmessfehlers ( $SE_M$ ) kann für beide Faktoren auf Basis des Testwerts ein 95%-KI berechnet werden, welches mit 95-prozentiger Wahrscheinlichkeit den wahren Wert einer Person umschließt. Dies erleichtert die Interpretation der Ergebnisse des WAI in der Einzelfalldiagnostik (Bühner, 2011).

### 3.4.5 Konvergenzvalidierung

Im Rahmen der Arbeit sollte weiterhin der Zusammenhang des WAI mit weiteren externen Indikatoren der Funktionsfähigkeit gemäß den Hypothesen 5a-b, 6a-b sowie 7a-b überprüft werden. Die Konvergenzvalidierung erfolgte anhand von Zusammenhangsanalysen in SPSS (Produkt-Moment-Korrelation). Dabei wurden die in den vorherigen Schritten berechneten und gewichteten Faktorwerte der Faktoren WAI\_F1 und WAI\_F2 jeweils mit der physischen (PCS) und der mentalen (MCS) Skala des SF-12 sowie den physischen ( $w_{PD}$ ) und mentalen ( $w_{MD}$ ) Domänen des arbeits-

bezogenen NFAS-w in Zusammenhang gesetzt. Auf eine Minderungskorrektur der berechneten Korrelationen durch die Reliabilität der eingesetzten Tests wird an dieser Stelle verzichtet, da den physischen und mentalen Skalen des SF-12 und der NFAS heterogene Subskalen zugrunde liegen (Bühner, 2011).

## 4 Ergebnisse

Im folgenden Abschnitt werden die Resultate der durchgeführten Analysen zur Überprüfung der Hypothesen berichtet. Im ersten Teil erfolgt die Präsentation der Ergebnisse zur Modellgüte sowie der Schätzer der spezifizierten Modelle. Anschließend werden die Ergebnisse der Überprüfung der Messinvarianz jenes Modells, welches sowohl nach empirischen als auch nach theoretischen Überlegungen am besten auf die Daten passt, für das Geschlecht sowie für die drei Altersgruppen dargestellt, anhand derer ein alternativer Vorschlag für eine gewichtete Berechnung der Arbeitsfähigkeitsfaktoren präsentiert wird. Zuletzt erfolgt eine Darstellung der Zusammenhänge der Arbeitsfähigkeitsfaktoren mit weiteren Indikatoren der berufsbezogenen Funktionsfähigkeit.

### 4.1 Faktorielle Validität des WAI

Zur Überprüfung der Hypothesen 1a und 1b bezüglich der Faktorenstruktur des WAI wurden vier zuvor spezifizierte Modelle mittels konfirmatorischer Faktorenanalysen überprüft. In Tabelle 4.1 sind die Mittelwerte und Standardabweichungen sowie die polychorische Korrelationsmatrix mit den Standardfehlern der Variablen für die Items des WAI sowie für die drei Altersgruppen, das Geschlecht und den WAI Gesamtwert nach klassischer Berechnung als ungewichteter Summenindex dargestellt. Auf Tests zur Signifikanzprüfung wird an dieser Stelle in Mplus verzichtet. Informationen zur Aussagekraft der Höhe des Zusammenhangs können dem angegebenen Standardfehler des polychorischen Korrelationskoeffizienten entnommen werden.

Wie bereits in Kapitel 3.4.1 erwähnt, zeigen die Items des WAI untereinander positive Zusammenhänge von  $r_{\text{pol}} = .22$  (SE = .02) zwischen dem WAI5 und dem WAI7 bis  $r_{\text{pol}} = .59$  (SE = .01) zwischen dem WAI1 und dem WAI2. Das Alter zeigt negative Zusammenhänge von  $r_{\text{pol}} = -.02$  (SE = .02; WAI5) bis  $r_{\text{pol}} = -.19$  (SE = .02; WAI2 und WAI6) auf: Höhere Altersgruppen haben niedrigere Werte auf den Items als jüngere Altersgruppen. Für das Geschlecht zeigen sich lediglich mit den Items WAI3 ( $r_{\text{pol}} = -.13$ , SE = .02) und WAI7 ( $r_{\text{pol}} = .13$ , SE = .02) substantielle Zusammenhänge. Frauen haben somit geringere Punktwerte auf diesem Item und somit eine höhere Anzahl an Krankheiten im vergangenen Jahr diagnostiziert bekommen als Männer. Außerdem berichten sie über mehr psychische Reserven.

Bei der Berechnung des WAI auf seine herkömmliche Weise als ungewichteter Summenscore für einen Faktor (M = 40.22, SD = 6.20) zeigen sich positive Zusammenhänge von  $r_{\text{pol}} = .54$  (SE = .01; WAI5) bis  $r_{\text{pol}} = .76$  (SE = .01; WAI1) mit allen Items des WAI. Der Gesamtwert zeigt einen negativen Zusammenhang mit dem Alter ( $r_{\text{pol}} = -.18$ , SE = .02). Mit der Zugehörigkeit zu einer älteren Altersgruppe sinkt der Gesamtwert für den WAI. Für das Geschlecht liegt kein substantieller Zusammenhang mit dem WAI-Gesamtwert vor ( $r_{\text{pol}} = -.02$ , SE = .02). Für das polychorische ordinale Alpha ergibt sich ein  $\alpha_{\text{pol}} = .82$ , während das klassische Cronbachs Alpha auf Basis der Pearson-Korrelation ein  $\alpha = .75$  als angenommenes Maß der internen Konsistenz der Skala erreicht.

Eine Zusammenfassung der Parameter der Modellgüte für die vier geprüften Modelle ist in Tabelle 4.2 dargestellt. Für alle spezifizierten Modelle war der  $\chi^2$ -Test aufgrund der Stichprobengröße signifikant.

**Tab. 4.1** Mittelwerte, Standardabweichungen und polychorische Inter-Item-Korrelationen (Standardfehler in Klammern) des Work Ability Index

Item	M	SD	1	2	3	4	5	6	7	8	9
1 Altersgruppen	-	-	-								
2 Geschlecht	-	-	.06 (.02)	-							
3 WAI Gesamtwert	40.22	6.20	-.18 (.02)	-.02 (.02)							
4 WAI1: selbsteingeschätzte Arbeitsfähigkeit	8.02	1.78	-.13 (.02)	.02 (.02)	.76 (.01)						
5 WAI2: Arbeitsfähigkeit (bezogen auf die Arbeitstätigkeit)	8.41	1.37	-.19 (.02)	.01 (.02)	.73 (.01)	.59 (.01)					
6 WAI3: Anzahl Krankheiten (Jahr)	4.84	1.95	-.13 (.02)	-.13 (.02)	.67 (.01)	.29 (.02)	.29 (.02)				
7 WAI4: Beeinträchtigung der Arbeitsleistung durch Krankheiten	5.32	0.95	-.16 (.02)	-.00 (.02)	.73 (.01)	.52 (.01)	.51 (.01)	.49 (.01)			
8 WAI5: Krankenstandstage (Jahr)	3.98	1.01	-.02 (.02)	-.02 (.02)	.54 (.01)	.29 (.02)	.26 (.02)	.39 (.02)	.39 (.02)		
9 WAI6: Ausübung der derzeitigen Arbeit in den nächsten Jahren	6.20	1.63	-.19 (.02)	-.01 (.03)	.71 (.01)	.37 (.02)	.47 (.02)	.30 (.02)	.51 (.02)	.25 (.02)	
10 WAI7: Psychische Leistungsreserven	3.36	0.70	-.06 (.02)	.13 (.02)	.59 (.01)	.43 (.01)	.48 (.01)	.26 (.02)	.36 (.02)	.22 (.02)	.46 (.02)

Anmerkungen. N = 3 968. M = Mittelwert, SD = Standardabweichung, Standardfehler in Klammern. Geschlecht: 1 = männlich, 2 = weiblich. Altersgruppen: 1 = 31-40 Jahre, 2 = 41-50 Jahre, 3 = 51-60 Jahre.

**Modell A:** Die standardisierten Ladungen aller sieben WAI-Items sind für den WAI-Gesamtfaktor signifikant von Null verschieden ( $\lambda_{WAI5} = .46$  bis  $\lambda_{WAI4} = .77$ ). Dennoch lässt sich das einfaktorische Modell nicht zufriedenstellend durch die empirischen Daten abbilden. Neben dem  $\chi^2$ -Test ( $\chi^2 = 632.25$ ,  $df = 14$ ,  $p < .001$ ) implizieren auch die übrigen Fit-Indizes eine schlechte Passung (CFI = .92; TLI = .88; RMSEA = .11 [.10; .11]). Somit misst der WAI kein eindimensionales Konstrukt und sowohl das ordinale  $\alpha_{pol} = .82$  als auch Cronbachs  $\alpha = .75$  sind als untere Schranke der Reliabilität nicht zuverlässig zu interpretieren.

**Modell B:** Die standardisierten Ladungen für den Faktor WAI\_F1 ( $\lambda_{WAI7} = .60$  bis  $\lambda_{WAI4} = .78$ ) und den Faktor WAI\_F2 ( $\lambda_{WAI3} = .67$  und  $\lambda_{WAI5} = .57$ ) im zweifaktoriellen Modell, bei dem die Items WAI3 und WAI5 auf dem WAI\_F2 und die übrigen Items auf dem WAI\_F1 laden, sind hier ebenfalls signifikant von Null verschieden. Die beiden Faktoren zeigen mit  $r_{pol} = .73$  einen positiven Zusammenhang. Die Modellgüte ist sowohl für den  $\chi^2$ -Test ( $\chi^2 = 469.53$ ,  $df = 13$ ,  $p < .001$ ) als auch für die deskriptiven Fit-Indizes (CFI = .94; TLI = .91; RMSEA = .09 [.09; .10]) nicht im akzeptablen Bereich. Der  $\chi^2$ -Differenztest im Vergleich mit Modell A zeigt eine signifikante Verbesserung ( $\chi^2_{diff} = 121.57$ ,  $df = 1$ ,  $p < .001$ ) des weniger restriktiven Modells B und spiegelt sich deskriptiv in leicht verbesserten Werten der Fit-Indizes wider.

**Modell C:** Das dritte Modell entspricht grundsätzlich Modell B, jedoch wurden Doppelladungen der Items WAI4 und WAI6 auf beide Faktoren zugelassen. Die standardisierten Ladungen reichen von  $\lambda_{WAI4} = .37$  bis  $\lambda_{WAI2} = .77$  für den WAI\_F1 und  $\lambda_{WAI6} = .19$  bis  $\lambda_{WAI3} = .66$  für den WAI\_F2 mit einer jeweils signifikanten Abweichung von Null und einem positiven Zusammenhang beider Faktoren ( $r_{pol} = .61$ ). Das Item WAI4 lädt dabei auf den zweiten Faktor höher als auf den ersten (Faktor 1:  $\lambda_{WAI4} = .37$ ; Faktor 2:  $\lambda_{WAI4} = .50$ ) während das Item WAI6 auf den ersten Faktor höher lädt (Faktor 1:  $\lambda_{WAI6} = .49$ ; Faktor 2:  $\lambda_{WAI6} = .19$ ). Bei signifikantem  $\chi^2$ -Test ( $\chi^2 = 161.95$ ,  $df = 11$ ,  $p < .001$ ) liegen der CFI = .98 in einem guten und der RMSEA = .06 [.05; .07] sowie der TLI = .96 in einem akzeptablen Bereich. Der Modellvergleich mit dem restriktiveren Modell B ist signifikant von Null verschieden ( $\chi^2_{diff} = 228.51$ ,  $df = 2$ ,  $p < .001$ ) und spricht dafür, dass Modell C besser als Modell B auf die empirischen Daten passt.

**Modell D:** Im vierten Modell laden die Items WAI3, WAI4 und WAI5 auf dem WAI\_F2 und die übrigen auf dem WAI\_F1, ohne Doppelladungen zuzulassen. Bei einem positiven Zusammenhang beider Faktoren ( $r_{pol} = .77$ ) weichen die standardisierten Ladungen sowohl für den ersten Faktor ( $\lambda_{WAI7} = .61$  bis  $\lambda_{WAI2} = .76$ ) als auch für den zweiten Faktor ( $\lambda_{WAI5} = .50$  bis  $\lambda_{WAI4} = .88$ ) signifikant von Null ab. Neben dem signifikanten  $\chi^2$ -Test ( $\chi^2 = 289.60$ ,  $df = 13$ ,  $p < .001$ ) zeigen die deskriptiven Fit-Indizes mit einem CFI = .97 und einem RMSEA = .07 [.07; .08] eine akzeptable Modellgüte an. Der TLI = .94 liegt dagegen knapp in einem nicht akzeptablen Bereich. Im Modellvergleich mit dem weniger restriktiven Modell C zeigt sich eine signifikante Verschlechterung der Modellgüte ( $\chi^2_{diff} = 101.75$ ,  $df = 2$ ,  $p < .001$ ) von Modell D. Ein  $\chi^2$ -Differenztest zwischen Modell B und Modell D ist nicht möglich, da sie nicht genestet sind. Jedoch weisen die deskriptiven Fit-Indizes darauf hin, dass Modell D besser als Modell B auf die empirischen Daten passt.

Wenngleich die Fit-Indizes des dritten Modells C insgesamt auf eine bessere Passung auf die empirischen Daten hindeuten, soll im Folgenden das sparsamere Modell D weiter betrachtet werden. Dafür sprechen die Ladungsmuster der Items WAI4 und WAI6, die akzeptablen Fit-Indizes CFI und RMSEA sowie die bessere inhaltliche Interpretierbarkeit für das Modell D. Alle weiteren Berechnungen und Ergebnisse der Arbeit beziehen sich daher auf dieses Modell. In Tabelle 4.3 werden die Muster der

standardisierten Ladungen für alle vier Modelle inklusive der Standardfehler präsentiert.

**Tab. 4.2** Parameter der Modellgüte (Fit-Indizes) der konfirmatorischen Faktorenanalysen

Modell	df	$\chi^2$	$\chi^2_{diff}$ Modell- vergleich mit	CFI	TLI	RMSEA	90%-KI RMSEA
A 1-Faktor-Modell	14	632.25**	-	.92	.88	.11**	[.10; .11]
B 2-Faktor-Modell (Items WAI3 und WAI5 auf Faktor WAI_F2)	13	469.53**	A 121.57* df = 1	.94	.91	.09**	[.09; .10]
C 2-Faktor-Modell (Doppelladungen der Items WAI4 und WAI6)	11	161.95**	B 228.51* df = 2	.98	.96	.06*	[.05; .07]
D 2-Faktor-Modell (Items WAI3, WAI4 und WAI5 auf Faktor WAI_F2)	13	289.60**	C 101.75* df = 2	.97	.94	.07**	[.07; .08]

Anmerkungen. N = 3 968. df = Zahl der Freiheitsgrade;  $\chi^2$  = Chi<sup>2</sup>-Test;  $\chi^2_{diff}$  = Chi<sup>2</sup>-Differenztest per DIFFTEST-Option; CFI = Comparative-Fit-Index, TLI = Tucker-Lewis-Index; RMSEA = Root-Mean-Square-Error of Approximation; KI = Konfidenzintervall.

\*p < .05, \*\*p < .001.

Aufgrund der akzeptablen Fit-Indizes von Modell D kann man von einer Eindimensionalität der beiden Faktoren jeweils für sich gesehen sprechen. Die Annahme der tau-Äquivalenz des Modells kann für den ersten Faktor im Vergleich zu Modell D ohne die Äquivalenz-Restriktion nicht bestätigt werden ( $\chi^2 = 1\,463.09$ , df = 16, p < .001;  $\chi^2_{diff} = 736.63$ , df = 3, p < .001; CFI = .82; TLI = .76; RMSEA = .15 [.14; .16]). Eine weitere Testung für beide Faktoren zusammen entfällt somit. Sowohl das ordinale  $\alpha_{pol\_1} = .78$  für WAI\_F1 und  $\alpha_{pol\_2} = .69$  für WAI\_F2 als auch Cronbachs  $\alpha_1 = .70$  für WAI\_F1 und  $\alpha_2 = .66$  für WAI\_F2 können somit lediglich als untere Grenze der Reliabilität interpretiert werden.

Zusammenfassend lässt sich festhalten, dass die in der Hypothese 1a angenommene Zweidimensionalität des WAI gestützt werden kann. Das Modell D, welches angelehnt an der Studie von Radkiewicz und Widerszal-Bazyl (2005) spezifiziert wurde, zeigt zwar nicht wie in Hypothese 1b erwartet die beste Modellgüte bezogen auf die Fit-Indizes, doch ist es im Vergleich zu Modell C durch seine Einfachstruktur besser interpretierbar.

**Tab. 4.3** Ladungsmuster der Items des Work Ability Index auf die Faktoren der vier getesteten Modelle A-D

Item	Modell	A		B		C		D	
		WAI_g	WAI_F1	WAI_F2	WAI_F1	WAI_F2	WAI_F1	WAI_F2	
		$\lambda$ (SE)	$\lambda$ (SE)	$\lambda$ (SE)	$\lambda$ (SE)	$\lambda$ (SE)	$\lambda$ (SE)	$\lambda$ (SE)	
WAI1									
selbsteingeschätzte Arbeitsfähigkeit		.68 (.01)	.69 (.01)		.73 (.01)		.71 (.01)		
WAI2									
Arbeitsfähigkeit (bezogen auf die Arbeitstätigkeit)		.72 (.01)	.72 (.01)		.77 (.01)		.76 (.01)		
WAI3: Anzahl Krankheiten (Jahr)		.53 (.02)		.67 (.02)		.66 (.02)		.57 (.02)	
WAI4: Beeinträchtigung der Arbeitsleistung durch Krankheiten		.77 (.01)	.78 (.01)		.37 (.03)	.50 (.03)		.88 (.01)	
WAI5: Krankenstandstage (Jahr)		.46 (.02)		.57 (.02)		.57 (.02)		.50 (.02)	
WAI6: Ausübung der derzeitigen Arbeit in den nächsten Jahren		.63 (.02)	.63 (.02)		.49 (.03)	.19 (.03)		.64 (.02)	
WAI7: Psychische Leistungsreserven		.59 (.01)	.60 (.01)		.61 (.01)			.61 (.01)	
Korrelation der Faktoren		-		.73		.61		.77	

Anmerkungen. N = 3 968. SE = Standardfehler. Alle standardisierten Schätzer der Ladungen ( $\lambda$ ) sind statistisch signifikant mit  $p < .001$ .

## 4.2 Messinvarianz des WAI

In den folgenden Abschnitten werden die Ergebnisse der multiplen Gruppenvergleiche zur Überprüfung der Messinvarianz des Modells zwischen den spezifizierten Gruppen berichtet. Gemäß der Hypothesen 2 und 3 wird angenommen, dass sowohl für das Geschlecht (weiblich und männlich) als auch für die drei Altersgruppen (31-40, 41-50 und 51-60 Jahre) skalare Messinvarianz vorliegt. In sämtlichen Schritten war der  $\chi^2$ -Test aufgrund der Stichprobengröße signifikant, weshalb das Hauptaugenmerk bei der Beurteilung der Modellgüte auf den deskriptiven Fit-Indizes liegt.

### 4.2.1 Geschlecht

Im ersten Schritt wurde das zweifaktorielle Modell D für beide Geschlechtergruppen getrennt geprüft. Die Fit-Indizes liegen für die Männer ( $\chi^2 = 125.22$ ,  $df = 13$ ,  $p < .001$ ; CFI = .97; TLI = .96; RMSEA = .07 [.06; .08]) und bis auf den TLI auch für die Frauen ( $\chi^2 = 184.82$ ,  $df = 13$ ,  $p < .001$ ; CFI = .96; TLI = .93; RMSEA = .08 [.07; .09]) im akzeptablen Bereich. In beiden Gruppen weichen die frei geschätzten standardisierten Ladungen der jeweiligen Faktoren signifikant von Null ab.

Die Schätzung des Basis-Modells im zweiten Schritt zur Beurteilung der konfiguralen Messinvarianz ergibt außer für den TLI akzeptable Fit-Indizes ( $\chi^2 = 311.00$ ,  $df = 26$ ,  $p < .001$ ; CFI = .96; TLI = .94; RMSEA = .07 [.07; .08]). Es wurden die Ladungen, Intercepts und Thresholds in beiden Gruppen frei geschätzt, die latenten Mittelwerte wurden zu Identifikationszwecken in beiden Gruppen auf 0 fixiert. Somit kann die Annahme der gleichen Faktorenstruktur in beiden Gruppen gestützt werden.

Die Gleichsetzung der Faktorladungen für beide Gruppen im dritten Schritt führt zu sieben zusätzlichen Freiheitsgraden. Die Modellgüte dieses restriktiveren Modells verbessert sich im Vergleich zum Basis-Modell ( $\chi^2 = 184.05$ ,  $df = 33$ ,  $p < .001$ ; CFI = .98; TLI = .98; RMSEA = .05 [.04; .06]) mit Fit-Indizes im guten Bereich. Der Modellvergleich mit dem Basis-Modell ist nicht signifikant ( $\chi^2_{diff} = 7.80$ ,  $df = 7$ ,  $p = .35$ ),  $\Delta CFI$  zeigt einen Anstieg von .02 und der RMSEA liegt sogar noch unterhalb des 90%-KIs des Basis-Modells. Somit kann von metrischer Messinvarianz ausgegangen werden, d. h. die Ladungen der Items auf beiden Faktoren sind für Männer und Frauen identisch.

Im letzten und vierten Schritt wurden die Intercepts und Thresholds in beiden Gruppen gleichgesetzt und die Gleichsetzung der latenten Mittelwerte für beide Gruppen aufgehoben. Insgesamt ist die Modellgüte weiterhin zufriedenstellend mit akzeptablen bis guten Fit-Indizes ( $\chi^2 = 255.56$ ,  $df = 45$ ,  $p < .001$ ; CFI = .97; TLI = .98; RMSEA = .05 [.04; .05]). Der  $\chi^2$ -Differenztest weist zwar eine signifikante Abweichung vom Modell der metrischen Messinvarianz auf ( $\chi^2_{diff} = 93.10$ ,  $df = 12$ ,  $p < .001$ ), jedoch verringerte sich der CFI lediglich um  $\Delta CFI = -.01$ . Der RMSEA des skalaren Modells liegt außerdem im 90%-KI des weniger restriktiven Modells. Dies bedeutet, dass sich die Intercepts und Thresholds in beiden Gruppen nicht voneinander unterscheiden und von skalarer Messinvarianz ausgegangen werden kann. Latente Mittelwertvergleiche sind für beide Faktoren zwischen Männern und Frauen somit zulässig.

In Tabelle 4.4 sind die Ergebnisse der hierarchischen Vergleiche zusammengefasst dargestellt.

Aufgrund der im letzten Schritt in Mplus festgelegten freien Schätzung der latenten Mittelwerte beider Faktoren für die Frauen bei gleichzeitiger Fixierung dieser bei den Männern auf 0 können die Mittelwerte direkt verglichen werden. Die Männer dienen somit hier als Referenzgruppe und für die Frauen wird die Differenz zum Mittelwert



der Männer berechnet. So ist der Mittelwert der Frauen auf dem ersten Faktor der subjektiv eingeschätzten Arbeitsfähigkeit und Ressourcen um 0.12 Skalenwerte (SE = .04,  $z = 2.82$ ,  $p = .01$ ) höher als bei den Männern. Der zweite Faktor der Gesundheitsbedingungen ist dagegen um -0.14 (SE = .04,  $z = -3.22$ ,  $p = .001$ ) Skalenwerte niedriger als bei den Männern. Zur Beurteilung der Signifikanz wurden beide  $z$ -Werte nach Rosenthal und DiMatteo (2001) in die Effektstärken  $r = .05$  für den ersten Faktor und  $r = -.05$  für den zweiten Faktor umgerechnet. Nach Cohen (1988) liegt demnach kein Effekt des Geschlechts für beide Faktoren vor.

**Tab. 4.4** Ergebnisse der Prüfung der Messinvarianz für das Geschlecht

	$\chi^2$	df	$\chi^2_{diff}$	$\Delta df$	CFI	$\Delta CFI$	TLI	RMSEA	90%-KI
Modelle der Einzelgruppen									
Männer (n = 2 029)	125.22**	13			.97		.96	.07	[.06; .08]
Frauen (n = 1 939)	184.82**	13			.96		.93	.08	[.07; .09]
Messinvarianz									
konfigurale MI	311.00**	26			.96		.94	.07	[.07; .08]
metrische MI	184.05**	33	7.80	7	.98	.02	.98	.05	[.04; .06]
skalare MI	255.56**	45	93.10**	12	.97	-.01	.98	.05	[.04; .05]

Anmerkungen. N = 3 968. MI = Messinvarianz; df = Zahl der Freiheitsgrade;  $\Delta df$  = Differenz der Freiheitsgrade;  $\chi^2$  = Chi<sup>2</sup>-Test;  $\chi^2_{diff}$  = Chi<sup>2</sup>-Differenztest per DIFFTEST-Option; CFI = Comparative-Fit-Index;  $\Delta CFI$  = Veränderung im CFI; TLI = Tucker-Lewis-Index; RMSEA = Root-Mean-Square-Error of Approximation; KI = Konfidenzintervall.

\*\* $p < .001$ .

#### 4.2.2 Altersgruppen

Die Überprüfung des Modells erfolgte zunächst für die drei Altersgruppen getrennt. In allen drei Gruppen wiesen die meisten Fit-Indizes akzeptable Werte auf. Für die 31- bis 40-Jährigen ( $\chi^2 = 75.25$ ,  $df = 13$ ,  $p < .001$ ; CFI = .96; TLI = .93; RMSEA = .07 [.06; .09]) und für die 41- bis 50-Jährigen ( $\chi^2 = 139.60$ ,  $df = 13$ ,  $p < .001$ ; CFI = .96; TLI = .94; RMSEA = .08 [.07; .09]) erreicht lediglich der TLI keine zufriedenstellenden Werte. Für die 51- bis 60-Jährigen ( $\chi^2 = 112.07$ ,  $df = 13$ ,  $p < .001$ ; CFI = .97; TLI = .95; RMSEA = .08 [.06; .09]) liegen alle Fit-Indizes im akzeptablen Bereich. Für alle drei Altersgruppen weichen die frei geschätzten standardisierten Ladungen signifikant von Null ab.

Die Modellgüte für das Basis-Modell mit frei geschätzten Ladungen, Intercepts und Thresholds für alle drei Gruppen liegt wiederum außer für den TLI im akzeptablen Bereich ( $\chi^2 = 321.89$ ,  $df = 39$ ,  $p < .001$ ; CFI = .96; TLI = .94; RMSEA = .07 [.07; .08]). Die Annahme der äquivalenten Faktorstruktur (konfigurale Messinvarianz) in allen drei Gruppen kann somit gestützt werden.

Im dritten Schritt wurden die Ladungen der Items auf die Faktoren in allen drei Gruppen gleichgesetzt. Diese Gleichsetzung führt in allen drei Gruppen weiterhin zu einer signifikanten Abweichung der Ladungen von Null. Für dieses restriktivere Modell verbessert sich die Modellgüte der deskriptiven Fit-Indizes im Vergleich zum Basis-Modell aus dem vorhergehenden Schritt ( $\chi^2 = 227.27$ ,  $df = 53$ ,  $p < .001$ ; CFI = .98; TLI = .97; RMSEA = .05 [.04; .06]). Bei signifikantem  $\chi^2$ -Differenztest ( $\chi^2_{diff} = 31.50$ ,  $df = 14$ ,  $p = .01$ ) steigt der CFI um  $\Delta CFI = .01$  und der RMSEA liegt unterhalb des KIs

des weniger restriktiven Modells. Die Annahme der metrischen Messinvarianz mit gleichen Faktorladungen in allen drei Altersgruppen kann somit gestützt werden.

Die Fit-Indizes des Modells der skalaren Messinvarianz mit gleichgesetzten Intercepts und Thresholds weisen auf eine gute Passung hin ( $\chi^2 = 311.70$ ,  $df = 77$ ,  $p < .001$ ; CFI = .97; TLI = .98; RMSEA = .05 [.04; .05]). Wiederum ist der  $\chi^2$ -Differenztest signifikant ( $\chi^2_{diff} = 104.20$ ,  $df = 24$ ,  $p < .001$ ), der CFI sinkt allerdings lediglich um  $\Delta CFI = -.01$  und der im Vergleich zum weniger restriktiven Modell unveränderte RMSEA liegt in dessen 90%-KI. Somit kann daraus geschlossen werden, dass sich die Intercepts und Thresholds in allen drei Altersgruppen nicht unterscheiden und skalare Messinvarianz vorliegt. Daher sind auch für die drei Altersgruppen Vergleiche der latenten Mittelwerte für beide Faktoren zulässig.

Eine zusammenfassende Darstellung der Ergebnisse der hierarchischen Vergleiche für die Altersgruppen erfolgt in Tabelle 4.5.

**Tab. 4.5** Ergebnisse der Prüfung der Messinvarianz für die drei Altersgruppen

	$\chi^2$	df	$\chi^2_{diff}$	$\Delta df$	CFI	$\Delta CFI$	TLI	RMSEA	90%-KI
Modelle der Einzelgruppen									
31-40 Jahre (n = 965)	75.25**	13			.96		.93	.07	[.06; .09]
41-50 Jahre (n = 1 658)	139.60**	13			.96		.94	.08	[.07; .09]
51-60 Jahre (n = 1 345)	112.07**	13			.97		.95	.08	[.06; .09]
Messinvarianz									
konfigurale MI	321.89**	39			.96		.94	.07	[.07; .08]
metrische MI	227.27**	53	31.50*	14	.98	.01 <sup>a</sup>	.97	.05	[.04; .06]
skalare MI	311.70**	77	104.20**	24	.97	-.01	.98	.05	[.04; .05]

Anmerkungen. N = 3 968. MI = Messinvarianz; df = Zahl der Freiheitsgrade;  $\Delta df$  = Differenz der Freiheitsgrade;  $\chi^2$  = Chi<sup>2</sup>-Test;  $\chi^2_{diff}$  = Chi<sup>2</sup>-Differenztest per DIFFTEST-Option; CFI = Comparative-Fit-Index;  $\Delta CFI$  = Veränderung im CFI; TLI = Tucker-Lewis-Index; RMSEA = Root-Mean-Square-Error of Approximation; KI = Konfidenzintervall.

\*\*p < .001, \*p < .01.

<sup>a</sup>Rundungsdifferenz.

Analog zu den Ergebnissen der Unterschiede der latenten Mittelwerte für beide Geschlechter können nun die Ausprägungen der Mittelwerte für die drei Altersgruppen verglichen werden. Die Mittelwerte in der Gruppe der 31- bis 40-Jährigen beider Faktoren bilden die Referenz. Für den ersten Faktor der subjektiv eingeschätzten Arbeitsfähigkeit und Ressourcen verringert sich der Mittelwert für die 41- bis 50-Jährigen um -0.23 Skalenwerte (SE = .05,  $z = -4.31$ ,  $p < .001$ ). Dies entspricht einem  $r = -.08$  und stellt somit keinen Effekt dar. Für die 51- bis 60-Jährigen verringert sich der Mittelwert um -0.52 Skalenwerte (SE = .06,  $z = -9.39$ ,  $p < .001$ ) im Vergleich zu den 31- bis 40-Jährigen. Die Stärke des Effekts ist hier mit  $r = -.20$  klein. Auch für den zweiten Faktor der Gesundheitsbedingungen zeigt sich eine Verringerung mit dem Wechsel in höhere Altersgruppen. So sinkt der Mittelwert im Vergleich zur Referenzgruppe für die 41- bis 50-Jährigen um -0.20 Skalenwerte (SE = .06,  $z = -3.56$ ,  $p < .001$ ) was einem  $r = -.07$  und somit keinem Effekt entspricht. Der Mittelwert der 51- bis 60-Jährigen sinkt im Vergleich zur Referenzgruppe um -0.40 Skalenwerte (SE = .05,  $z = -6.93$ ,  $p < .001$ ). Die Stärke des Effekts ist hier mit  $r = -.14$  klein.

Zusammenfassend können die Annahmen der Hypothesen 2 und 3 gestützt werden. Ausgehend von einer zweifaktoriellen Struktur des WAI erfasst er sowohl im Vergleich zwischen Männern und Frauen als auch zwischen den drei Altersgruppen das gleiche Konstrukt der Arbeitsfähigkeit. Ein Messmodell, in welchem sowohl die Ladungen als auch die Intercepts und Thresholds in den jeweiligen Subgruppen gleichgesetzt werden, weist eine gute Passung auf, wodurch die Annahme der skalaren Messinvarianz in den Subgruppen gestützt werden kann. Auf Grundlage der Ergebnisse der multiplen Gruppenvergleiche erfolgt im nächsten Abschnitt die Entwicklung eines gewichteten Summenindex für beide Faktoren. Da sowohl für das Geschlecht als auch für die Altersgruppen die Faktorladungen äquivalent sind, erfolgt bei der Bildung der Gewichtung keine Unterscheidung für diese Gruppen.

### 4.3 Gewichtete Summenindizes und Normwerttabellen

Aufgrund der zweifaktoriellen Struktur ist eine Berechnung und Interpretation des WAI als ungewichteter Summenindex nicht angemessen. Zur Überprüfung von Hypothese 4 wurden mittels multipler Regressionsanalysen, basierend auf den Ergebnissen der konfirmatorischen Faktorenanalyse für Modell D, zwei gewichtete Summenindizes gebildet, welche mit anderen Methoden zur Berechnung von Faktorwerten verglichen wurden.

Die Ergebnisse der multiplen Regressionsanalyse für die in Mplus ausgegebenen Faktorwerte ( $FS_{Mplus}$ ) je Faktor als Kriterium und die sieben WAI Items als Prädiktoren sind in Tabelle 4.6 zusammengefasst. Die standardisierten Regressionskoeffizienten aller Items werden als Annäherung an die benötigten Faktorwertkoeffizienten interpretiert. Durch die Items werden 99.6% der Varianz im ersten Faktorwert  $FS1_{Mplus}$  ( $M = -0.02$ ,  $SD = 0.87$ ) erklärt ( $F_{(7, 3862)} = 71669.02$ ,  $p < .001$ ). Für den zweiten Faktorwert  $FS2_{Mplus}$  ( $M = -0.04$ ,  $SD = 0.82$ ) werden 96.6% der Varianz durch die Items erklärt ( $F_{(7, 3862)} = 15466.73$ ,  $p < .001$ ). Anhand der standardisierten Regressionskoeffizienten mit Einschluss aller Items für jeweils beide Faktoren wurden die gewichteten Summenindizes  $F1_{Reg}$  ( $M = 0.01$ ,  $SD = 0.99$ ) und  $F2_{Reg}$  ( $M = 0.01$ ,  $SD = 0.98$ ) berechnet.

Ebenfalls in Tabelle 4.6 dargestellt sind die Faktorwert-Koeffizienten, welche für die Faktorwerte  $F1_{PCA}$  ( $M = 0$ ,  $SD = 1$ ) und  $F2_{PCA}$  ( $M = 0$ ,  $SD = 1$ ) der beiden Faktoren aus der Hauptkomponentenanalyse resultieren. Die hier nicht abgebildete, rotierte Komponentenmatrix entspricht in ihrem Ladungsmuster nicht vollständig dem spezifizierten Modell D, da für das Item WAI4 eine Doppelladung auf beide Faktoren geschätzt wird ( $\lambda_1 = .54$  und  $\lambda_2 = .53$ ).

Außerdem sind in Tabelle 4.6 die Faktorwert-Koeffizienten für  $F1_{Mplus\_metr}$  ( $M = 0.00$ ,  $SD = 0.90$ ) und  $F2_{Mplus\_metr}$  ( $M = 0.00$ ,  $SD = 0.87$ ) aus der konfirmatorischen Faktorenanalyse unter der Annahme metrischer Variablen abgebildet. Als Maß für die Güte der Schätzung der Faktorwerte wird der Faktorwert-Determinationskoeffizient für beide Faktoren ausgegeben, welcher sich aus der Korrelation des geschätzten Faktorwerts mit dem latenten Faktor ergibt (Eid et al., 2013). Dieser beträgt für den ersten Faktor  $r_{F1} = .90$  und  $r_{F2} = .87$  für den zweiten Faktor.

Die Ergebnisse der Zusammenhangsanalysen zwischen den berechneten Faktorwerten sind in Tabelle 4.7 dargestellt. Der Zusammenhang für den in Mplus unter Beachtung der ordinalen Variablen berechneten Faktorwert des ersten Faktors  $FS1_{Mplus}$  ist mit dem Faktorwert der Regressionsmethode  $F1_{Reg}$  mit  $r_T = .97$  ( $p < .001$ ) am höchsten, wobei der Unterschied zum Zusammenhang mit  $FS1_{Mplus\_metr}$  ( $r_T = .96$ ,  $p < .001$ ) lediglich marginal ist. Für den zweiten Faktorwert  $FS2_{Mplus}$  ist der Zusammenhang mit

dem Faktorwert  $FS2_{Mplus\_metr}$  unter der Annahme metrischer Variablen in Mplus am höchsten ( $r_T = .95$ ,  $p < .001$ ). Der Unterschied zum Zusammenhang mit  $F2_{Reg}$  ist hier auch gering ( $r_T = .94$ ,  $p < .001$ ).

**Tab. 4.6** Koeffizienten der Regressionsanalyse, Hauptkomponentenanalyse und metrischen konfirmatorischen Faktorenanalyse für die Berechnung der Faktorwerte des Work Ability Index

	$\beta$ (N = 3 870)		FWK <sub>PCA</sub> (N = 3 870)		FWK <sub>Mplus</sub> (N = 3 968)	
	FS1 <sub>Mplus</sub>	FS2 <sub>Mplus</sub>	F1 <sub>PCA</sub>	F2 <sub>PCA</sub>	F1 <sub>Mplus_metr</sub>	F2 <sub>Mplus_metr</sub>
WAI1: selbsteingeschätzte Arbeitsfähigkeit	.320	.099	.307	-.006	.192	.085
WAI2: Arbeitsfähigkeit (bezogen auf die Arbeitstätigkeit)	.428	.199	.359	-.084	.250	.111
WAI3: Anzahl Krankheiten (Jahr)	.083	.246	-.167	.570	.029	.082
WAI4: Beeinträchtigung der Arbeitsleistung durch Krankheiten	.171	.541	.124	.254	.173	.499
WAI5: Krankenstandstage (Jahr)	.050	.157	-.186	.589	.052	.149
WAI6: Ausübung der derzeitigen Arbeit in den nächsten Jahren	.143	.052	.311	-.112	.075	.033
WAI7: Psychische Leistungsreserven	.181	.082	.375	-.194	.211	.094

Anmerkungen. Für eine genauere Anwendung in der Praxis werden drei Nachkommastellen berichtet.  $\beta$  = standardisierte Regressionskoeffizienten mit  $FS1_{Mplus}$  und  $FS2_{Mplus}$  als Kriterien;  $FWK_{PCA}$  = Faktorwertkoeffizienten der Hauptkomponentenanalyse;  $FWK_{Mplus}$  = Faktorwertkoeffizienten der konfirmatorischen Faktorenanalyse in Mplus unter der Annahme metrischer Variablen.

**Tab. 4.7** Zusammenhänge der berechneten Faktorwerte des Work Ability Index

	FS1 <sub>Mplus</sub>	F1 <sub>Reg</sub>	F1 <sub>PCA</sub>	FS2 <sub>Mplus</sub>	F2 <sub>Reg</sub>	F2 <sub>PCA</sub>
F1 <sub>Reg</sub>	.97*					
F1 <sub>PCA</sub>	.69*	.69*				
FS1 <sub>mplus_metr</sub>	.96*	.97*	.69*			
F2 <sub>Reg</sub>				.94*		
F2 <sub>PCA</sub>				.46*	.47*	
FS2 <sub>mplus_metr</sub>				.95*	.95*	.45*

Anmerkungen. N = 3 870. Korrelationskoeffizient Kendall-tau-b. FS<sub>Mplus</sub> = exakter Faktorwert aus Mplus; F<sub>Reg</sub> = gewichteter Summenindizes mit Regressionskoeffizienten als Gewicht; FWK<sub>PCA</sub> = Faktorwertkoeffizienten der explorativen Faktorenanalyse; FWK<sub>Mplus</sub> = Faktorwertkoeffizienten der konfirmatorischen Faktorenanalyse in Mplus unter der Annahme metrischer Variablen. \*p < .001.

Aufgrund der guten Annäherung der Faktorwerte aus der Regressionsmethode können die Regressionsgewichte ( $\beta$ ) aus Tabelle 4.6 zur Berechnung individueller Faktorwerte der beiden Faktoren des WAI herangezogen werden. Für Anwender ergeben sich dadurch folgende zwei Formeln (4.1a) und (4.1b) zur Berechnung, welche eine z-Standardisierung der erhobenen Items voraussetzen. Liegen den Anwendern keine repräsentativen Daten aus der eigenen Stichprobe zur Standardisierung vor, können die Kennwerte der Referenzstichprobe aus der S-MGA verwendet werden (siehe Tabelle 4.1).

$$F1_{\text{Reg}} = (0.320 * z_{\text{WAI1}}) + (0.428 * z_{\text{WAI2}}) + (0.083 * z_{\text{WAI3}}) + (0.171 * z_{\text{WAI4}}) + (0.050 * z_{\text{WAI5}}) + (0.143 * z_{\text{WAI6}}) + (0.181 * z_{\text{WAI7}}) \quad (4.1a)$$

$$F2_{\text{Reg}} = (0.099 * z_{\text{WAI1}}) + (0.199 * z_{\text{WAI2}}) + (0.246 * z_{\text{WAI3}}) + (0.541 * z_{\text{WAI4}}) + (0.157 * z_{\text{WAI5}}) + (0.052 * z_{\text{WAI6}}) + (0.082 * z_{\text{WAI7}}) \quad (4.1b)$$

Für die Erstellung der Normwerttabellen erfolgte eine Transformation der Faktorwerte. Anhand des theoretischen Minimums (Min) und Maximums (Max) der jeweiligen Faktorwerte F1<sub>Reg</sub> und F2<sub>Reg</sub>, welche auch dem empirischen Min und Max entsprechen, wurden die Rohwerte der Faktorwerte anhand der Formeln (4.2a) und (4.2b) einer linearen Transformation unterzogen:

$$\text{WAI\_F1} = 4 + \frac{F1_{\text{Reg}} - \text{Min}}{\text{Max} - \text{Min}} * 27 \quad (4.2a)$$

$$\text{WAI\_F2} = 3 + \frac{F2_{\text{Reg}} - \text{Min}}{\text{Max} - \text{Min}} * 15 \quad (4.2b)$$

So erhielten die beiden Faktoren nach zusätzlicher Rundung in ihrer Metrik das gleiche Min und Max, wie es bei der klassischen Berechnung des WAI für die jeweiligen zusammengefassten Items möglich wäre. Dies dient ausschließlich der besseren Handhabbarkeit individueller Ergebnisse. Eine Überführung klassischer WAI-Berechnungen in die Normwerttabellen ist nicht möglich. Die neuen deskriptiven Ergebnisse der Faktoren sind im Vergleich zu den Faktorwerten in Tabelle 4.8 dargestellt.

**Tab. 4.8** Mittelwerte, Standardabweichungen, Minima und Maxima der ursprünglichen und transformierten Faktorwerte des Work Ability Index

	M	SD	Min	Max
F1 <sub>Reg</sub>	0.01	0.99	-5.55	1.36
F2 <sub>Reg</sub>	0.01	0.98	-5.19	1.26
WAI_F1	25.73	3.88	4	31
WAI_F2	15.08	2.29	3	18

Anmerkungen. N = 3 870. M = Mittelwert, SD = Standardabweichung, Min = Minimum, Max = Maximum.

In Tabelle 4.9 werden die deskriptiven Ergebnisse der beiden Faktoren getrennt für Männer und Frauen sowie für die Altersgruppen dargestellt. Die Unterschiede in den Mittelwerten zwischen den Geschlechtern sowie zwischen den Altersgruppen und die dazugehörigen Effektstärken wurden bereits im Rahmen der Messinvarianz (Kapitel 4.2) berichtet.

**Tab. 4.9** Deskriptive Beschreibung der transformierten Faktorwerte des Work Ability Index getrennt nach Geschlecht und Altersgruppe

Faktor	Geschlecht	Altersgruppe in Jahren	n	M	SD	Min	Max
WAI_F1	männlich	31-40	518	26.58	3.25	13	31
		41-50	811	25.74	3.80	9	31
		51-60	649	24.88	4.21	6	31
	weiblich	31-40	422	26.62	3.72	8	31
		41-50	806	26.00	3.89	4	31
		51-60	664	24.97	3.90	9	31
WAI_F2	männlich	31-40	518	15.59	1.87	7	18
		41-50	811	15.12	2.24	5	18
		51-60	649	14.76	2.49	3	18
	weiblich	31-40	422	15.50	2.26	5	18
		41-50	806	15.16	2.30	3	18
		51-60	664	14.60	2.33	5	18

Anmerkungen. N = 3 870. M = Mittelwert, SD = Standardabweichung, Min = Minimum, Max = Maximum.

Anhand von Formel (4.3) wurden unter Berücksichtigung des ordinalen Alpha ( $\alpha_{\text{pol}}$ ) auf Basis polychorischer Korrelationen als Maß der Reliabilität und der Standardabweichung (SD) des Messwertes zunächst für beide Faktoren die jeweiligen Standardmessfehler ( $SE_M$ ) berechnet:

$$SE_M = SD * \sqrt{1 - \alpha_{\text{pol}}} \quad (4.3)$$

Für WAI\_F1 ergibt sich mit  $SD_1 = 3.88$  und  $\alpha_{\text{pol}_1} = .78$  ein Standardmessfehler von  $SE_{M_1} = 1.82$  und mit  $SD_2 = 2.29$  und  $\alpha_{\text{pol}_2} = .69$  ein Standardmessfehler von  $SE_{M_2} = 1.28$  für WAI\_F2. Für die Berechnung der Konfidenzintervalle beider Faktoren ergeben sich anhand der Standardmessfehler und des zweiseitigen z-Werts für

die 95-prozentige Sicherheitswahrscheinlichkeit von  $z_{\alpha/2} = 1.96$  folgende Formeln (4.4a) und (4.4b), wobei WAI\_F1 und WAI\_F2 die berechneten Faktorwerte einer Person repräsentieren:

$$\begin{aligned} KI_{WAI\_F1} &= WAI\_F1 \pm SE_{M\_1} * z_{\alpha/2} & (4.4a) \\ &= WAI\_F1 \pm 3.57 \end{aligned}$$

$$\begin{aligned} KI_{WAI\_F2} &= WAI\_F2 \pm SE_{M\_2} * z_{\alpha/2} & (4.4b) \\ &= WAI\_F2 \pm 2.50 \end{aligned}$$

Anhand der vorhandenen Werte lassen sich mit der Formel (4.5) zusätzlich die kritischen Differenzen für beide Faktoren berechnen, ab deren Werte sich Testwerte zweier Personen aus einem identischen Test tatsächlich unterscheiden.

$$D_{krit} = z_{\alpha/2} * SD * \sqrt{2 * (1 - \alpha_{pol})} \quad (4.5)$$

Daraus ergeben sich eine  $D_{krit\_1} = 5.04$  für den WAI\_F1 und eine  $D_{krit\_2} = 3.53$  für den WAI\_F2, ab deren Wert sich zwei Testwerte unterscheiden.

Folgt man der kategorialen Einteilung von Tuomi et al. (1998), welche in Kapitel 2.4 dargestellt wurde, ergeben sich für die Gesamtstichprobe der Erwerbstätigen in Deutschland die in Tabelle 4.10 dargestellten Kategorien der Arbeitsfähigkeit getrennt für beide Faktoren.

**Tab. 4.10** Kategorisierung der Arbeitsfähigkeit der neuen Faktoren des Work Ability Index gemäß Tuomi et al. (1998)

Anteil	Punkte WAI_F1	Punkte WAI_F2	Arbeitsfähigkeit	Ziele von Maßnahmen
15%	4-22	3-13	Schlecht	Arbeitsfähigkeit wiederherstellen
35%	23-26	14-16	Mittelmäßig	Arbeitsfähigkeit verbessern
35%	27-30	17	Gut	Arbeitsfähigkeit unterstützen
15%	31	18	Sehr gut	Arbeitsfähigkeit erhalten

In den Tabellen 4.11 und 4.12 werden neben den transformierten Faktorwerten außerdem die Prozentränge getrennt für beide Geschlechter und die jeweiligen Altersgruppen dargestellt.

Zusammenfassend lässt sich die Annahme der Hypothese 4 stützen. Auf Basis der Daten aus der konfirmatorischen Faktorenanalyse ist es anhand einer Annäherung möglich, gewichtete Summenindizes für beide Faktoren zu bilden, welche eine adäquate Abbildung der exakten Faktorwerte der Faktoren des WAI darstellen. Die Gewichtungskoeffizienten sowie die Kennwerte der Items des WAI können von Anwendern für eigene Berechnungen übernommen werden. Die neu berechneten Faktorwerte konnten wiederum in Normwerttabellen für beide Faktoren getrennt nach Geschlecht und Altersgruppen abgebildet werden.

**Tab. 4.11** Normwerttabelle für beide Faktoren des Work Ability Index für **Männer** getrennt nach Altersgruppen

31-40 Jahre				41-50 Jahre				51-60 Jahre			
WAI_F1	PR	WAI_F2	PR	WAI_F1	PR	WAI_F2	PR	WAI_F1	PR	WAI_F2	PR
4 - 12	0	3 - 6	0	4 - 8	0	3 - 4	0	4 - 5	0	3	0.2
13	0.2	7	0.4	9	0.1	5	0.1	6	0.3	4	0.5
14 - 16	0.6	8	0.6	10	0.4	6	0.4	7	0.5	5	0.8
17	1.4	9	0.8	11	0.6	7	0.5	8 - 10	0.8	6	1.5
18	2.5	10	1.4	12 - 14	0.9	8	1.1	11	1.4	7	1.7
19	3.5	11	2.9	15	1.4	9	2.5	12	1.7	8	2.2
20	5.6	12	6.0	16	2.2	10	3.9	13	1.8	9	3.1
21	7.9	13	12.9	17	3.1	11	6.8	14	2.2	10	5.1
22	10.2	14	26.6	18	4.6	12	11.8	15	2.8	11	9.9
23	15.4	15	40.7	19	7.3	13	21.3	16	3.7	12	16.2
24	23.6	16	60.6	20	9.4	14	33.7	17	5.4	13	25.7
25	33.6	17	88.6	21	13.6	15	49.0	18	7.2	14	39.8
26	43.2	18	100.0	22	18.9	16	68.4	19	10.9	15	54.4
27	56.9			23	23.7	17	88.5	20	14.5	16	71.8
28	66.2			24	31.7	18	100.0	21	19.6	17	91.8
29	78.4			25	40.8			22	23.9	18	100.0
30	92.7			26	53.0			23	29.9		
31	100.0			27	62.4			24	39.4		
				28	73.6			25	50.4		
				29	83.6			26	59.5		
				30	93.3			27	69.6		
				31	100.0			28	81.0		
								29	89.1		
								30	96.0		
								31	100.0		

Anmerkungen. N = 1 978. WAI\_F1 und WAI\_F2 = berechnete und transformierte Faktorwerte einer Person; PR = Prozentrang.



**Tab. 4.12** Normwerttabelle für beide Faktoren des Work Ability Index für **Frauen** getrennt nach Altersgruppen

31-40 Jahre				41-50 Jahre				51-60 Jahre			
WAI_F1	PR	WAI_F2	PR	WAI_F1	PR	WAI_F2	PR	WAI_F1	PR	WAI_F2	PR
4 - 7	0	3 - 4	0	4	0.2	3	0.1	4 - 8	0	3 - 4	0
8	0.2	5	0.9	5 - 6	0.4	4	0.4	9	0.2	5	0.2
9	0.7	6	1.4	7 - 10	0.5	5 - 6	0.7	10	0.5	6	0.8
10 - 11	1.2	7 - 8	1.9	11	0.6	7	1.0	11 - 13	0.9	7	1.1
12 - 14	1.4	9	2.6	12	0.7	8	2.0	14	1.7	8	1.8
15	2.1	10	3.8	13	1.0	9	2.5	15	2.4	9	3.2
16	2.6	11	5.5	14	1.4	10	3.5	16	3.0	10	4.8
17	2.8	12	7.6	15	1.9	11	6.5	17	3.9	11	9.2
18	3.1	13	14.0	16	2.5	12	11.2	18	5.6	12	17.2
19	4.0	14	24.6	17	3.3	13	19.9	19	9.3	13	28.3
20	6.2	15	38.2	18	4.0	14	33.3	20	12.5	14	43.5
21	6.6	16	60.7	19	6.1	15	47.0	21	19.6	15	60.2
22	10.4	17	87.4	20	8.6	16	66.9	22	23.8	16	76.4
23	15.2	18	100.0	21	11.8	17	88.7	23	30.9	17	93.1
24	21.8			22	15.3	18	100.0	24	39.0	18	100.0
25	31.3			23	20.8			25	49.2		
26	41.5			24	28.2			26	62.0		
27	50.9			25	37.0			27	72.0		
28	66.1			26	49.4			28	80.6		
29	75.8			27	61.0			29	88.7		
30	91.2			28	71.0			30	95.9		
31	100.0			29	80.5			31	100.0		
				30	92.8						
				31	100.0						

Anmerkungen. N = 1 892. WAI\_F1 und WAI\_F2 = berechnete und transformierte Faktorwerte einer Person; PR = Prozentrang.

#### 4.4 Konvergenzvalidität des WAI

Zur Konvergenzvalidierung der beiden Faktoren des WAI gegenüber den physischen und mentalen Skalen des SF-12 sowie den arbeitsbezogenen physischen und mentalen Skalen des NFAS-w gemäß Hypothesen 5a-b bis 7a-b wurden die Zusammenhänge untereinander überprüft. Für die PCS des SF-12 beträgt Cronbachs  $\alpha = .84$  und für die MCS  $\alpha = .78$  als Maß für die interne Konsistenz. Für die arbeitsbezogene NFAS-w ergibt sich für die w\_PD ein Cronbachs  $\alpha = .91$  und  $\alpha = .86$  für die w\_MD. Die Ergebnisse der Zusammenhangsanalysen sind in Tabelle 4.13 dargestellt. Zum Vergleich werden ebenfalls die Zusammenhänge mit dem klassischen WAI (WAI\_g) als eindimensionalen ungewichteten Summenwert berichtet. Soweit nicht anders angegeben, sind die Unterschiede zwischen den Zusammenhängen, geprüft anhand von z-Tests für Korrelationskoeffizienten aus abhängigen Stichproben, jeweils mit  $p < .05$  signifikant.

**Tab. 4.13** Mittelwerte, Standardabweichungen und Produkt-Moment-Korrelationen zwischen den Faktoren des Work Ability Index (WAI) und den physischen und mentalen Skalen des Short-Form-12 Health Survey (SF-12) und der arbeitsbezogenen Norwegian Function Assessment Scale (NFAS-w)

	M	SD	1	2	3	4	5	6
1 WAI_g	40.22	6.20						
2 WAI_F1	25.73	3.88	.94*					
3 WAI_F2	15.08	2.29	.93*	.88*				
4 SF-12_PCS	50	10	.60*	.56*	.63*			
5 SF-12_MCS	50	10	.43*	.44*	.39*	.00		
6 NFAS-w_PD	1.25	0.53	-.50*	-.46*	-.53*	-.64*	-.16*	
7 NFAS-w_MD	1.20	0.32	-.45*	-.44*	-.44*	-.29*	-.43*	.45*

Anmerkungen. N = 3 861. PD = physische Domäne; MD = mentale Domäne;

PCS = physische Komponente; MCS = mentale Komponente.

\* $p < .001$ .

Zwischen der PD und der MD des NFAS-w zeigt sich ein positiver Zusammenhang von  $r = .45$  ( $p < .001$ ). Je höher die arbeitsbezogene physische Funktionsfähigkeit ist, desto höher ist auch die arbeitsbezogene mentale Funktionsfähigkeit. Zwischen den Subskalen PCS und MCS des SF-12 besteht dagegen kein Zusammenhang ( $r = .00$ ,  $p = .80$ ). Die Skalen des SF-12 stehen ebenfalls im Zusammenhang mit den Skalen der NFAS-w ( $r = -.16$  bis  $r = -.64$ , alle  $p < .001$ ). Je höher die gesundheitsbezogene Lebensqualität für die mentalen und physischen Skalen ausfällt, desto geringere Einschränkungen werden für die arbeitsbezogene Funktionsfähigkeit berichtet. Dabei sind die Zusammenhänge sowohl zwischen den mentalen als auch zwischen den physischen Skalen höher als die Zusammenhänge zwischen den mentalen und physischen Skalen.

Für beide Faktoren des WAI zeigen sich insgesamt hohe und signifikante Zusammenhänge mit den jeweiligen Subskalen des SF-12 und der NFAS-w. Wie erwartet ist der Zusammenhang zwischen der physischen Skala des SF-12 und dem gesundheitsbezogenen Faktor WAI\_F2 ( $r = .63$ ,  $p < .001$ ) höher als mit dem Faktor der Arbeitsfähigkeit und Ressourcen WAI\_F1 ( $r = .56$ ,  $p < .001$ ). Daneben ist der Zusammenhang zwischen der mentalen Skala des SF-12 und dem Faktor WAI\_F1 ( $r = .44$ ,  $p < .001$ ) höher als mit dem WAI\_F2 ( $r = .39$ ,  $p < .001$ ). Betrachtet man jedoch die Zusammenhänge innerhalb der beiden Faktoren des WAI mit den jeweiligen Skalen des SF-12, so ist der Zusammenhang für beide Faktoren jeweils höher mit der PCS als mit der MCS.

Verglichen mit dem eindimensionalen WAI-Summenwert ist zu erkennen, dass der Zusammenhang der PCS mit WAI\_F2 über und mit WAI\_F1 unter dem des WAI\_g ( $r = .60$ ,  $p < .001$ ) liegt. Für die MCS ist der Zusammenhang mit WAI\_F1 größer und mit WAI\_F2 kleiner als mit dem WAI\_g ( $r = .43$ ,  $p < .001$ ). Insgesamt lässt sich sagen, dass je höher die Werte für beide Faktoren des WAI sind, desto besser ist auch die gesundheitsbezogene Lebensqualität bezogen auf die körperliche und mentale Gesundheit.

Bezogen auf den arbeitsbezogenen NFAS-w ist der Zusammenhang zwischen der w\_PD und dem WAI\_F2 ( $r = -.53$ ,  $p < .001$ ) ebenfalls stärker als mit dem WAI\_F1 ( $r = -.46$ ,  $p < .001$ ). Die Zusammenhänge zwischen der w\_MD und dem WAI\_F1 ( $r = -.44$ ,  $p < .001$ ) und dem WAI\_F2 ( $r = -.44$ ,  $p < .001$ ) sind dagegen nicht verschieden ( $z = -0.29$ ,  $p = .39$ ). Innerhalb der Faktoren des WAI ist auch hier der Zusammenhang jeweils stärker mit der physischen Skala als mit der mentalen Skala.

Der Vergleich mit dem klassischen WAI-Summenwert zeigt hier in Bezug auf die physische Domäne, dass der Zusammenhang mit WAI\_F2 über und mit WAI\_F1 unter dem des WAI\_g ( $r = -.50$ ,  $p < .001$ ) liegt. Für die mentale Domäne des NFAS-w zeigt sich hier kein Unterschied zwischen WAI\_g ( $r = -.45$ ,  $p < .001$ ) sowohl mit WAI\_F1 ( $z = 0.62$ ,  $p = .27$ ) als auch mit WAI\_F2 ( $z = 0.93$ ,  $p = .18$ ). Auch hier lässt sich insgesamt sagen, dass höhere Werte auf beiden Faktoren des WAI mit weniger Einschränkungen bei der Durchführbarkeit arbeitsrelevanter Handlungen einhergehen.

Zusammenfassend lassen sich Hinweise der konvergenten Validität der Faktoren WAI\_F1 und WAI\_F2 des WAI gegenüber den Skalen des SF-12 und der NFAS-w finden, wobei nicht alle aufgestellten Hypothesen gestützt werden.

## 5 Diskussion

In den folgenden Abschnitten werden zunächst die Ergebnisse vor dem Hintergrund der gestellten Hypothesen diskutiert und inhaltlich interpretiert. Außerdem erfolgen eine kritische Auseinandersetzung mit den eingesetzten Methoden und eine zusammenfassende Einschätzung der vorliegenden Arbeit mit einem Ausblick für weiteren Forschungsbedarf.

### 5.1 Ergebnisdiskussion

Der WAI ist ein bekanntes und weltweit häufig in der Forschung und im betrieblichen Kontext eingesetztes Instrument zur Erfassung des Konstrukts der Arbeitsfähigkeit (Ilmarinen, J., 2009; Tuomi et al., 1998). In Deutschland wird der WAI ebenfalls im BGM eingesetzt, mit dem Ziel durch Hinweise auf Interventionsbedarf individuelle Arbeitsfähigkeit im Betrieb zu erhalten und fördern (BAuA, 2013). Das übergeordnete Ziel der vorliegenden Arbeit war eine Konstruktvalidierung des WAI durch psychometrische Analysen, welche die Untersuchung der inneren Faktorenstruktur auch unter Gesichtspunkten der Messinvarianz zwischen verschiedenen Subgruppen beinhaltet. Diese Analysen bildeten die Grundlage für eine verbesserte Berechnung und Validierung der individuellen Indexwerte.

#### 5.1.1 Dimensionalität des WAI

Dem WAI liegt eine multidimensionale Perspektive der Arbeitsfähigkeit unter Berücksichtigung der individuellen Arbeitsbedingungen, mentalen Ressourcen und Gesundheit zugrunde. Dennoch wurde der WAI in den 80er Jahren unter der theoretischen Annahme der Eindimensionalität entwickelt (Tuomi et al., 1985). So wird der individuelle Index der Arbeitsfähigkeit als ein ungewichteter Summenwert aus den erfassten Items berechnet. Diese postulierte Eindimensionalität wurde in der Vergangenheit im Rahmen verschiedener Studien bezüglich der psychometrischen Eigenschaften des WAI in Frage gestellt, wobei heterogene Ergebnisse an unterschiedlichen Stichproben mit nicht vergleichbaren Analysemethoden gefunden wurden (z. B. Bethge et al., 2012; Martus et al., 2010; Radkiewicz & Widerszal-Bazyl, 2005). Neben einer Überprüfung der inneren Faktorenstruktur des WAI an einer repräsentativen Erwerbstätigenstichprobe in Deutschland war das Ziel der vorliegenden Arbeit eine angemessenere Anwendung statistischer Analysemethoden unter Berücksichtigung der Metrik der Items. Aufgrund der schiefen Verteilung und der teilweise ordinalen Metrik der Items wurden die konfirmatorischen Faktorenanalysen zur Überprüfung von vier spezifizierten Modellen (siehe Abbildung 5 im Kapitel 3.4.2) in Mplus unter Berücksichtigung eines robusten Schätzverfahrens durchgeführt. Dabei wurde statt der Kovarianzmatrix die polychorische Korrelationsmatrix zugrunde gelegt.

Die Ergebnisse der Modellüberprüfung deuten darauf hin, dass der WAI nicht wie von den Entwicklern impliziert einer eindimensionalen Struktur folgt. Die schlechten Fit-Indizes vom einfaktoriellen Modell A veranschaulichen die Nichtpassung des Modells auf die empirischen Daten der Studie. Eher kann davon ausgegangen werden, dass dem Konstrukt der Arbeitsfähigkeit zwei korrelierende Faktoren zugrunde liegen, welche zum einen durch die subjektive Arbeitsfähigkeit und Ressourcen sowie zum anderen durch die individuellen Gesundheitsbedingungen abgebildet werden. Somit kann die Annahme der Hypothese 1a gestützt werden.

Es wurden weiterhin drei konkurrierende zweifaktorielle Modelle zur Überprüfung der Hypothese 1b getestet. Das Modell B weist im Vergleich zu den übrigen getesteten zweifaktoriellen Modellen eine schlechte Modellgüte auf, woraus geschlossen werden kann, dass der zweite Faktor der Gesundheitsbedingungen nicht allein durch die Items WAI3 und WAI5 erklärt werden kann. Die alleinige Betrachtung der Fit-Indizes würde für die Annahme von Modell C sprechen. Hier werden jedoch Doppelladungen der Items WAI4 und WAI6 zugelassen, welche eine Interpretation des Konstrukts der Arbeitsfähigkeit durch zwei getrennte, wenn auch korrelierende Faktoren erschweren. Auch weisen die Ladungsmuster dieser beiden Items darauf hin, dass das Item WAI4 eher dem zweiten Faktor und das Item WAI6 deutlich dem ersten Faktor zugeordnet werden sollten. Durch die Spezifikation dieser Doppelladungen gehen zwei Freiheitsgrade verloren, wodurch das Modell weniger restriktiv bzw. sparsam ist. Ohne theoretische Begründung ist diese Komplexitätserhöhung auch bei besseren Fit-Indizes nicht haltbar. Allein dem WAI4 kann auch inhaltlich ein geringer Bezug zum ersten Faktor der Arbeitsfähigkeit zugeschrieben werden, da er die Beeinträchtigung der Arbeitsleistung durch Krankheiten erfasst. Der Schwerpunkt dieses Items liegt jedoch auf der individuellen Gesundheit. Dagegen ist das Modell D bei gleicher Anzahl von Freiheitsgraden wie Modell B genauso restriktiv und zeigt dennoch bessere und akzeptable Fit-Indizes als dieses. Die Zuordnung der Items zu den Faktoren kann inhaltlich und empirisch begründet werden. So repräsentieren die Items WAI1, WAI2, WAI6 und WAI7 einen Faktor, der die subjektive aktuelle und zukünftige Arbeitsfähigkeit sowie die individuellen Ressourcen abbildet. Der zweite Faktor mit den Items WAI3, WAI4 und WAI5 bildet dagegen die individuellen Gesundheitsbedingungen ab. Die zugelassene Korrelation der Faktoren ist dabei substantiell. Aufgrund von Multikollinearität wird dieser starke Zusammenhang für Analysen problematisch, in denen beide Faktoren gemeinsam als Prädiktoren zur Vorhersage eines Kriteriums, z. B. dem vorzeitigen Erwerbsausstieg, eingehen. Insgesamt wurden die Erwartungen der Hypothese 1b nicht vollständig erfüllt. Das Modell D zeigt im Vergleich mit den übrigen Modellen nicht die beste Modellpassung, doch weist es eine bessere Modellgüte als das ebenso sparsame Modell B auf. Durch das Vorliegen einer Einfachstruktur ist es zudem besser interpretierbar als das Modell C.

Vergleiche mit anderen Studien zur faktoriellen Struktur des WAI sind aufgrund der unterschiedlichen Stichproben und Analysemethoden nur unter Vorbehalt möglich. In der Studie von Martus et al. (2010) wird ebenfalls die Annahme der eindimensionalen Struktur des WAI anhand von konfirmatorischen Faktorenanalysen in AMOS unter Nichtberücksichtigung der ordinalen Metrik von fünf der sieben Items abgelehnt. Die Daten dieser Studie mit einer ad-hoc-Stichprobe bestehend aus 324 weiblichen Bürokräften, Lehrerinnen und Lehrern sowie Erzieherinnen stützen die Annahme eines zweifaktoriellen Modells mit Doppelladungen der Items WAI4 und WAI6 auf beiden Faktoren, welches dem Modell C der vorliegenden Arbeit entspricht. In der Studie von Bethge et al. (2012) an 1 036 Erwerbstätigen aus Deutschland wurde dagegen, ohne Überprüfung mehrfaktorieller Modelle, die Annahme einer eindimensionalen Struktur des WAI durch akzeptable Fit-Indizes gestützt. In der ländervergleichenden Studie von Radkiewicz und Widerszal-Bazyl (2005) an europäischen Krankenschwestern fanden sich in sieben von neun Ländern zweifaktorielle Strukturen des WAI gemäß dem Modell D der vorliegenden Arbeit. Lediglich für deutsche und finnische Krankenschwestern wurde die Annahme der Eindimensionalität des WAI gestützt. Die Ergebnisse basieren auf Hauptkomponentenanalysen ohne konfirmatorische Überprüfung der Modellgüte. Zu erwähnen ist, dass auch in dieser ländervergleichenden Studie für das Item WAI4 in den meisten Ländern substantielle Doppel-

ladungen gefunden wurden. Für den WAI6 fand sich lediglich für Italien eine Ladung auf beiden Faktoren. Radkiewicz und Widerszal-Bazyl (2005) leiten aus ihren Ergebnissen eine länderübergreifende Generalisierbarkeit (außer für Deutschland und Finnland) der faktoriellen Struktur des WAI ab. Diese ist jedoch lediglich auf die Stichprobe der Krankenschwestern anwendbar. Der Unterschied zu den vorliegenden Ergebnissen kann in der Spezifität der untersuchten Berufsgruppe der Krankenschwestern liegen. Künftig ist zu prüfen, ob die vorliegenden Ergebnisse, die auf einer Repräsentativstichprobe Erwerbstätiger aus Deutschland basieren, ebenfalls auf repräsentative Erwerbstätigenstichproben anderer Länder übertragbar sind.

Als Maß der internen Konsistenz der Skalen des WAI aus Modell D fanden sich Werte für das ordinale Alpha von  $\alpha_{\text{pol}_1} = .78$  und  $\alpha_{\text{pol}_2} = .69$  für beide Faktoren. Ein Vergleich ist für die Subskalen mit vorherigen Studien nicht möglich, da diese lediglich Werte für Cronbachs Alpha für den Gesamtindex des WAI berichten. Der berechnete Wert für das ordinale Alpha des Gesamtindex der vorliegenden Arbeit kann aufgrund der nicht vorhandenen Eindimensionalität nicht zur Interpretation der Reliabilität herangezogen werden. Nach dem niederländischen COTAN-System (Committee On Test Affairs Netherlands) zur Beurteilung der Qualität von Tests (Evers, 2001) muss zur Interpretation der Reliabilität eines Tests zuerst entschieden werden, auf welchem Niveau Entscheidungen als Folge der Testergebnisse vorgenommen werden können. Der WAI kann u. a. in Betrieben auf individueller Ebene für die Einzelfalldiagnostik, auf der betrieblichen Abteilungsebene oder im Kontext der Forschung auf Bevölkerungsebene angewandt werden. Je nach Niveau der Anwendung unterscheidet sich die Anforderung an die Höhe der Reliabilität eines Tests. So wäre die Reliabilität des WAI auf individueller Ebene für wichtige Entscheidungen mit  $\alpha_{\text{pol}} < .80$  als mangelhaft einzuschätzen. Dabei sollte jedoch beachtet werden, dass aufgrund der nicht vorliegenden tau-Äquivalenz des Modells eine Unterschätzung der Reliabilität auf Grundlage der internen Konsistenz möglich ist. Die Kürze der Skalen und die Heterogenität der Items sollte ebenfalls nicht außer Acht gelassen werden. Auch unter diesen Gesichtspunkten kann die Empfehlung von Amler et al. (2015) unterstützt werden, den WAI in Kombination mit anderen Instrumenten zur Erfassung der Arbeitsfähigkeit einzusetzen (siehe auch Kapitel 2.3). Welche Auswirkungen eine mangelhafte Reliabilität auf individuelle Testergebnisse haben kann, wird in Kapitel 5.1.3 an einem rechnerischen Beispiel erläutert.

Radkiewicz und Widerszal-Bazyl (2005) argumentieren auf Grundlage von Korrelationsanalysen der einzelnen WAI-Items mit dem eindimensionalen Gesamtindex, dass der WAI5 vom Fragebogen ausgeschlossen werden sollte. Dieser zeigte den geringsten Zusammenhang mit dem Gesamtindex ( $r = .21$ ) sowohl für die Gesamtstichprobe als auch für die einzelnen Länder. Der Ausschluss des Items würde eine marginale Verbesserung des Cronbachs Alphas hervorbringen. Dieser Argumentation kann anhand der Ergebnisse der vorliegenden Arbeit nicht gefolgt werden. Der WAI5 zeigt mit dem Gesamtindex (siehe Tabelle 4.1) unter der Annahme der Eindimensionalität zwar den geringsten, aber dennoch einen starken Zusammenhang. Auf Grundlage des zweidimensionalen Modells muss hier zusätzlich die Ladung des Items auf dem zweiten Faktor betrachtet werden. Die interne Konsistenz des zweiten Faktors würde sich bei Entfernung des Items auf  $\alpha_{\text{pol}_2} = .66$  verschlechtern, was zum einen der Verkürzung der Skala und zum anderen dem Informationsverlust des substantiell auf den Faktor ladenden Items geschuldet sein kann.

### 5.1.2 Zulässigkeit von Mittelwertvergleichen

In der Vergangenheit wurden bereits Vergleiche der Arbeitsfähigkeit zwischen verschiedenen Subgruppen vorgenommen, ohne zu überprüfen, ob der WAI in den jeweiligen Subgruppen das gleiche Konstrukt misst (Gould, Polvinen & Seitsamo, 2008; Ilmarinen, J. et al., 1997; Martus et al., 2011). So könnten z. B. unterschiedliche körperliche Voraussetzungen von Frauen gegenüber Männern oder Älteren gegenüber Jüngeren eine andere Wahrnehmung und Interpretation einzelner Items des WAI implizieren, was sich in unterschiedlichen Ausprägungen der latenten Mittelwerte niederschlägt. Voraussetzung für latente Mittelwertvergleiche ist das Vorliegen von skalärer Messinvarianz, d. h. die Faktorenstruktur, die Ladungen sowie die Intercepts und Thresholds sollten in den Subgruppen nicht verschieden sein. Für die vorliegende Arbeit war die Überprüfung der skalaren Messinvarianz der Subgruppen für das Geschlecht (männlich und weiblich) und für drei Altersgruppen (31-40, 41-50 und 51-60 Jahre) von Bedeutung. Aufgrund der Ergebnisse der Modellvergleiche wurden die multiplen Gruppenvergleiche auf Grundlage des Modells D nach dem Step-Up-Ansatz anhand konfirmatorischer Faktorenanalysen durchgeführt (Brown, 2006).

In den Hypothesen 2 und 3 wurde sowohl für die Subgruppen bezogen auf das Geschlecht als auch für die drei Altersgruppen skalare Messinvarianz angenommen. Diese Hypothesen konnten durch die Ergebnisse der vorliegenden Arbeit gestützt werden. So unterscheiden sich die Subgruppen weder in der Faktorenstruktur (konfigurale Messinvarianz), noch in der Höhe der Ladungen (metrische Messinvarianz) oder in der Höhe der Intercepts bzw. Thresholds (skalare Messinvarianz). Daraus folgt, dass der Vergleich von Mittelwerten der beiden latenten Faktoren für diese Subgruppen zulässig und eine Interpretation der Ergebnisse möglich ist. Der WAI erfasst somit in allen untersuchten Subgruppen das gleiche zweidimensionale Konstrukt der Arbeitsfähigkeit.

Im Rahmen der Überprüfung der skalaren Messinvarianz können in den Analysen direkt die Mittelwerte der Faktoren für die Gruppen untersucht werden. So zeigte sich im Vergleich der Mittelwerte zwischen Männern und Frauen zwar ein signifikanter Unterschied in beiden Faktoren, welcher jedoch durch die Betrachtung der Effektstärke relativiert und als nicht vorhandener Effekt interpretiert werden kann. Somit können bisherige Befunde einer mit dem WAI gemessenen niedrigeren Arbeitsfähigkeit von Frauen im Vergleich zu Männern, welche lediglich auf der Angabe von Mittelwerten mit dem Hinweis auf Signifikanz beruhen, nicht repliziert werden (Gould, Polvinen & Seitsamo, 2008). Diese Tendenz deutete sich bereits bei der Analyse der Zusammenhänge der Einzelitems mit dem Geschlecht an. So fanden sich für das Geschlecht lediglich substantiell negative Zusammenhänge mit dem WAI5 und positive Zusammenhänge mit dem WAI7, was sich in einer höheren Anzahl an Krankheitsdiagnosen (weniger Punkte im WAI5) und einer höheren Einschätzung der individuellen psychischen Leistungsreserven (mehr Punkte für WAI7) für Frauen ausdrückt. Auf die anderen Items des WAI hat das Geschlecht keinen Einfluss.

Anders sieht es bei dem Alter aus. So zeigt sich unter Berücksichtigung der Effektstärke besonders beim Vergleich der 51- bis 60-Jährigen mit der Referenzgruppe der 31- bis 40-Jährigen sowohl eine Verringerung des Faktormittelwertes für die subjektive Arbeitsfähigkeit und Ressourcen als auch des Faktormittelwertes der individuellen Gesundheitsbedingungen. Für die mittlere Altersgruppe der 41- bis 50-Jährigen ist der Mittelwertunterschied im Vergleich zur Referenzgruppe zwar signifikant, bezogen auf die Effektstärke allerdings nicht bedeutsam. Auch bei der Analyse der bivariaten Zusammenhänge der Items des WAI mit den Altersgruppen zeigte sich für alle Items

außer für den WAI5 und den WAI7 eine substantielle Verringerung der Werte bei Wechsel in eine höhere Altersgruppe. Diese Ergebnisse stützen die bisherigen Befunde zum Einfluss des Alters auf die Arbeitsfähigkeit (Gould, Polvinen & Seitsamo, 2008; Ilmarinen, J. et al., 1997; Martus et al., 2011), wobei in der vorliegenden Arbeit die Berücksichtigung der zweifaktoriellen Struktur des WAI Berücksichtigung fand.

### 5.1.3 Itemgewichtung

Ausgehend von einer zweifaktoriellen Struktur des WAI wurde im Rahmen der vorliegenden Arbeit die Berechnung eines einzelnen ungewichteten Summenwertes in Frage gestellt. Die gleichwertige Betrachtung der einzelnen Items ohne Berücksichtigung ihres Einflusses auf einen oder mehrere Faktoren zählt laut DiStefano et al. (2009) zu den non-refined Methoden, welche den refined Methoden unterlegen sind. So war ein weiteres Ziel der vorliegenden Arbeit die Ableitung einer für Anwender nutzbaren Formel zur Gewichtung der Items des WAI für die jeweiligen Faktoren unter Berücksichtigung der Ergebnisse der konfirmatorischen Faktorenanalysen. Die durch eine Annäherung berechneten Gewichtungskoeffizienten sollten dabei die exakten Faktorwerte der Faktoren des WAI adäquat abbilden (Hypothese 4). Anhand der neuen Berechnungsmethode sollten außerdem Normwerttabellen aus der Repräsentativstichprobe für Anwender des WAI bereitgestellt werden.

Die laut Grice (2001) valideste Methode zur gewichteten Berechnung von Faktorwerten erfolgt durch die Regressionsmethode von Thurstone (1935; zitiert durch DiStefano et al., 2009, S. 4). Dieses Vorgehen ist in Mplus zwar möglich, doch werden für kategoriale Items keine Gewichte in Form von Faktorwertkoeffizienten ausgegeben, da diese durch ein iteratives Vorgehen erzeugt werden müssten (B. Muthén, persönl. Mitteilung, 08.06.2016). Um die Gewichte für Anwender jedoch sichtbar zu machen, erfolgte die Berechnung dieser durch eine Annäherung. Dazu wurden die in Mplus ausgegebenen individuellen Faktorwerte für beide Faktoren in einer multiplen linearen Regressionsanalyse als Kriterium eingesetzt, während die Items des WAI für beide Faktoren als Prädiktoren in die Berechnung eingingen. Die daraus resultierenden standardisierten Regressionsgewichte stellten die Gewichte für die z-standardisierten Items des WAI dar, wodurch für jeden Faktor neue individuelle Faktorwerte unter Berücksichtigung aller Items berechnet werden konnten. Diese wurden zur Validierung mit den Faktorwerten resultierend aus einer Hauptkomponentenanalyse, Faktorwerten aus Mplus unter der Annahme metrischer Items sowie mit den exakten Faktorwerten unter Beachtung der ordinalen Items in Zusammenhang gesetzt.

Die Ergebnisse der Zusammenhangsanalysen zeigen, dass die neu berechneten Faktorwerte für beide Faktoren eine sehr gute Annäherung an die exakten Faktorwerte aus Mplus darstellen, wodurch die Hypothese 4 gestützt werden kann. Somit können die standardisierten Regressionsgewichte aus Tabelle 4.6 für die Gewichtung der individuellen z-standardisierten Items des WAI von Anwendern, welche keinen Zugriff auf eine repräsentative Stichprobe oder auf die notwendigen statistischen Programme haben, herangezogen werden. Allerdings zeigen die in Mplus erzeugten Faktorwerte unter der Annahme metrischer Items eine ebenso gute Annäherung an die exakten Faktorwerte wie jene durch die Regressionsgewichte berechneten. Somit könnten auch die in Mplus ausgegebenen Faktorwertkoeffizienten ( $FWK_{Mplus}$  in Tabelle 4.6) als Gewicht herangezogen werden, um den Umweg über die Regressionsanalysen, welche die Prädiktoren ebenfalls als metrische Variablen behandeln, zu vermeiden.



Um eine einfachere Interpretation der individuellen Messwerte des WAI in den Normwerttabellen zu ermöglichen, fand eine lineare Transformation und Rundung der Faktorwerte statt. Rein optisch betrachtet haben die Faktorwerte somit eine Metrik erhalten, welche der Darstellung des klassisch berechneten WAI ähnelt, wenn man die ungewichteten Indexwerte für die jeweiligen Items der zwei Faktoren getrennt summieren würde. Dennoch darf nicht davon ausgegangen werden, dass die dargestellten Normwerttabellen auch für die klassische Berechnung des WAI mit einer getrennten ungewichteten Berechnung beider Faktoren aus den ihnen zugeordneten Items anwendbar sind.

Anhand der berechneten Konfidenzintervalle ist erkennbar, wie sich die vergleichsweise geringe Reliabilität der Faktoren auf die Präzision bei der Schätzung individueller Testwerte auswirkt. So ergibt sich für den ersten Faktor WAI\_F1 eine Breite vom 1.8-fachen der Standardabweichung und für den zweiten Faktor WAI\_F2 eine Breite vom 2.2-fachen der Standardabweichung. Die berechneten kritischen Differenzen für beide Faktoren veranschaulichen ergänzend, in welchem Umfang sich Testwerte zweier Personen unterscheiden müssten, damit dieser Unterschied für eine Interpretation relevant ist. Die Verwendung sehr feiner Abstufungen bei der Darstellung der Testwerte in den Normwerttabellen impliziert den Eindruck von Präzision und einer hohen Differenziertheit des WAI. Dieser hohe Differenzierungsgrad ist bei Vorliegen einer mangelhaften Reliabilität im Rahmen der Einzelfalldiagnostik nicht angemessen (Lienert & Raatz, 1998). Hat beispielsweise ein 35-jähriger Mann einen Testwert von 15 auf dem zweiten Faktor erreicht, so errechnet sich für ihn ein KI von [12.5; 17.5], welches mit 95-prozentiger Wahrscheinlichkeit den wahren Testwert umschließt. Aufgrund seines ermittelten Testwerts erhält er einen Prozentrang von 40.7, d. h. 40.7% der Personen aus der Eichstichprobe haben einen schlechteren oder gleichen Wert erreicht. Neben dieser Punktschätzung sind jedoch auch die aus der gesamten Breite des KIs resultierenden Prozentränge zu berücksichtigen, welche im dargestellten Beispiel einen Bereich von 6.0 bis 88.6 umfassen können.

Dennoch ist die neue Berechnungsmethode unter Berücksichtigung der faktoriellen Struktur des WAI gegenüber einer ungewichteten Berechnung eines einzelnen Index für die Praxis vorzuziehen. Die gewichtete Berechnung der individuellen Werte getrennt für beide Faktoren beruht auf den Ergebnissen der konfirmatorische Faktorenanalyse, wodurch der individuelle Beitrag der einzelnen Items für die Faktorenbildung messfehlerbereinigt berücksichtigt wird. Die Bereitstellung der Normwerte liefert eine Interpretationshilfe für Anwendungen im Rahmen der Individualdiagnostik. Sie entstammen einer repräsentativen Stichprobe, deren Umfang auch in den Untergruppen gemäß dem COTAN-System (Evers, 2001) zufriedenstellend groß ist.

#### **5.1.4 Zusammenhang mit konstrukt-nahen Instrumenten**

Während sich die vorangestellten Analysen mit der Überprüfung der faktoriellen Validität als einen Teilaspekt der Konstruktvalidität beschäftigt haben, ging es im letzten Schritt um den Zusammenhang des mit dem WAI gemessenen Konstrukts der Arbeitsfähigkeit mit konstrukt-nahen Außenvariablen, welche in der Befragung zeitgleich zur Erfassung der Funktionsfähigkeit erhoben wurden. Für die Analyse der konvergenten Validität wurden die physischen und mentalen Skalen des SF-12 und der arbeitsbezogenen NFAS-w herangezogen, für welche gemäß den Hypothesen 5a-b bis 7a-b Zusammenhänge mit den Faktoren des WAI erwartet wurden.

Wie angenommen bestehen mittlere bis starke Zusammenhänge zwischen beiden Faktoren des WAI und den Skalen des SF-12 (Hypothese 5a). Je höher die mit dem

WAI gemessene subjektive Arbeitsfähigkeit und Ressourcen sowie die Gesundheitsbedingungen eingeschätzt werden, desto höher wird auch die gesundheitsbezogene Lebensqualität für die mentalen und physischen Skalen bewertet. Bei einer differenzierten Betrachtung der Skalen und Faktoren fällt der stärkere Zusammenhang bei der WAI-Faktoren mit der physischen Skala des SF-12 auf, wobei hier der gesundheitsbezogene Faktor des WAI dominiert (Hypothese 6a). Weiterhin ist der Zusammenhang der mentalen Skala des SF-12 höher mit dem Faktor der Arbeitsbedingungen und Ressourcen als mit dem gesundheitsbezogenen Faktor (Hypothese 6b).

Generell kann man aus den signifikanten Unterschieden in den Korrelationen schließen, dass die Faktoren des WAI zwischen den Skalen des SF-12 differenzieren. Jedoch fallen die Zusammenhänge insgesamt so hoch aus, dass diese differenzierte Betrachtung nur schwer zu interpretieren ist. Die bis zum jetzigen Zeitpunkt in der Literatur berichteten Analysen zur konvergenten Validität beziehen sich auf die Zusammenhänge des eindimensionalen WAI mit den Subskalen des SF-36. In den Analysen der vorliegenden Arbeit wurden daher zum Vergleich auch die Zusammenhänge der Skalen des SF-12 mit dem klassischen Gesamtindex der Arbeitsfähigkeit berichtet. Die Ergebnisse bestätigen die bisherigen Befunde, dass der WAI im stärkeren Zusammenhang mit der physischen Komponente der gesundheitsbezogenen Lebensqualität steht (Abdolalizadeh et al., 2012; Alexopoulos et al., 2013; Martinez et al., 2009; Radkiewicz & Widerszal-Bazyl, 2005). Dies steht im Einklang mit den stärkeren Zusammenhängen beider WAI-Faktoren und insbesondere des gesundheitsbezogenen Faktors mit der physischen Skala des SF-12. Somit hat gerade die körperliche Gesundheit einen besonderen Einfluss auf die individuelle Arbeitsfähigkeit.

Bezogen auf die arbeitsbezogene physische und mentale Funktionsfähigkeit gemessen mit dem NFAS-w zeigten sich ebenfalls die erwarteten Zusammenhänge mit beiden Faktoren des WAI, welche mittel bis stark ausfielen (Hypothese 5b). Je höher die mit dem WAI gemessene subjektive Arbeitsfähigkeit und Ressourcen sowie die Gesundheitsbedingungen eingeschätzt werden, desto geringer fallen die Einschränkungen in der arbeitsbezogenen Funktionsfähigkeit aus. Bei der differenzierten Betrachtung der einzelnen Skalen und Faktoren beider Instrumente fallen die Unterschiede in den Zusammenhängen lediglich gering aus. Tendenziell hängen beide Faktoren des WAI ähnlich wie beim SF-12 eher mit der physischen Domäne des NFAS-w zusammen. Wie erwartet ist der Zusammenhang des WAI-Faktors der Gesundheitsbedingungen höher mit der physischen als mit der mentalen Skala des NFAS-w (Hypothese 7a). Dagegen unterscheidet sich der Zusammenhang der mentalen Skala des NFAS-w nicht wie erwartet zwischen den Faktoren (Hypothese 7b). Zwar wurde der WAI in seiner klassischen Form bisher nicht mit der NFAS validiert, doch konnten die generellen Zusammenhänge aufgrund der Validierung der NFAS am SF-36 erwartet werden (Brage et al., 2004). Auch hier kann wie bereits im vorherigen Absatz abgeleitet werden, dass insbesondere die körperliche Gesundheit eine stärkere Rolle für die Arbeitsfähigkeit spielt.

Insgesamt betrachtet sind die Zusammenhänge zwischen den Faktoren des WAI und den Skalen des SF-12 und der NFAS-w bedeutend, so dass die Annahme einer gemeinsamen Erfassung von spezifischen Komponenten der Arbeitsfähigkeit gestützt werden kann. Dabei sind die Zusammenhänge jedoch nicht so stark, dass von einer Redundanz der Methoden ausgegangen werden kann, indem durch die Instrumente das gleiche Konstrukt gemessen wird. Der besondere Einfluss der Gesundheit auf die individuelle Arbeitsfähigkeit steht im Einklang mit den Ausführungen Tenglands (2011) zum Konzept der Arbeitsfähigkeit. Demnach ist die Gesundheit eine zentrale

Voraussetzung für das Vorhandensein von Arbeitsfähigkeit, jedoch nicht als allein-stehender Faktor.

## 5.2 Kritische Methodenwürdigung

Im Rahmen der Analysen der vorliegenden Arbeit sollten die besonderen Skalierungseigenschaften der Items des WAI Beachtung finden, wie es in der Literatur bis zum jetzigen Stand kaum der Fall war. Trotz der ordinalen Metrik der Items WAI3 bis WAI7 wurden häufig Methoden angewandt, welche diesen Items ein metrisches Messniveau zuschreiben. So wurden die Analysen als konfirmatorische Faktorenanalyse in SPSS AMOS oder als Hauptkomponentenanalysen in SPSS durchgeführt. SPSS AMOS bietet nicht die Möglichkeit der Berücksichtigung kategorialer Variablen, ebenso wenig wie die Anwendung der Hauptkomponentenanalyse, welche häufig fälschlicherweise als explorative Faktorenanalyse bezeichnet wird (Eid et al., 2013). Die in der Wissenschaft übliche Annahme eines metrischen Messniveaus von Ratingskalen kann für diese speziellen Items des WAI nicht angewandt werden. Aus diesen Gründen wurden die konfirmatorischen Faktorenanalysen sowie die Überprüfung der Messinvarianz in Mplus durchgeführt und die kategorialen Items auch als solche definiert. Den Ausgangspunkt für die Modellschätzung bildete statt der Kovarianzmatrix die polychorische Korrelationsmatrix der Items. Dabei wurde das robuste WLSMV-Schätzverfahren angewandt, welches weniger Anforderungen an die Verteilungseigenschaften der Items stellt.

In der empirischen Literatur wird zur Bestimmung der Reliabilität häufig das Maß der internen Konsistenz (Cronbachs Alpha) herangezogen. Dieses beruht wiederum auf der Kovarianzmatrix metrischer Variablen und kann lediglich dann als ein Maß für die wahre Reliabilität interpretiert werden, wenn das Modell essentiell tau-äquivalent ist. Diese Annahme wird oftmals nicht überprüft. In der vorliegenden Arbeit wurde daher das ordinale Alpha auf Grundlage der polychorischen Korrelationsmatrix berechnet und die tau-Äquivalenz des spezifizierten Modells geprüft (Brown, 2006; Gadermann et al., 2012).

Eine weitere Stärke der Arbeit ist die Verwendung einer umfangreichen Stichprobe, welche ein weitgehend unverzerrtes Abbild der Population der sozialversicherungspflichtig Beschäftigten in Deutschland zwischen 31 und 60 Jahren darstellt. In den bisherigen Validierungsstudien des WAI wurden dagegen häufig kleinere Stichproben oder nur ausgewählte Berufsgruppen herangezogen (Abdolalizadeh et al., 2012; Alexopoulos et al., 2013; Martinez et al., 2009; Martus et al., 2010; Radkiewicz & Widerszal-Bazyl, 2005). Mit fehlenden Werten in den untersuchten Items wurde durch die Full-Information-Maximum-Likelihood-Methode in Mplus in der vorliegenden Arbeit adäquat umgegangen.

Die Berechnung der gewichteten Summenindizes stellt grundsätzlich eine besondere Stärke der Arbeit dar. Da das spezifizierte Modell nicht tau-parallel ist (die Ladungen der Items unterscheiden sich innerhalb der Faktoren) und die Items eine unterschiedliche Metrik aufweisen, ist ein ungewichteter Summenindex nicht interpretierbar (DiStefano et al., 2009; Eid et al., 2013). Die Gewichte resultierten dabei aus dem konfirmatorischen Modell unter Beachtung des Messfehlers, wobei aufgrund der nicht ausgegebenen Faktorwertkoeffizienten auf eine regressionsanalytische Berechnung als Annäherung an die wahren Faktorwerte zurückgegriffen werden musste. Die durchgeführten Analysen zeigen jedoch, dass dieser Umweg über die Regressionsanalyse nicht unbedingt notwendig ist, da die Faktorwertkoeffizienten aus der konfirmatorischen Faktorenanalyse unter der Annahme metrischer Variablen zu einem na-

hezu identischen Ergebnis bzgl. der Validität gegenüber den exakten Faktorwerten führen. Bei den Regressionsanalysen wird ebenfalls ein metrisches Messniveau der Prädiktoren angenommen. Alternativ wäre die Anwendung von Dummy-Variablen für die ordinalen Items methodisch betrachtet angemessener, jedoch aufgrund des Umfangs der Items nicht praktikabel gewesen.

Einschränkend muss zusätzlich bemerkt werden, dass zur Berechnung der Punktwerte der einzelnen Items des WAI weiterhin die festgelegte Berechnungsvorschrift der Entwickler (Tuomi et al., 1998) herangezogen wurde (siehe Tabelle 3.2 in Kapitel 3.2.1). Eine Überprüfung alternativer und empirisch begründeter Berechnungen wird an dieser Stelle für künftige Forschungen angeraten. Des Weiteren hätten die berechneten Faktorwerte für die weiteren Analysen aus den Ergebnissen der Überprüfung der skalaren Messinvarianz resultieren müssen (B. Muthén, persönl. Mitteilung, 04.08.2016), obwohl für beide Gruppen das Vorliegen skalarer Messinvarianz angenommen werden kann. Aufgrund der untersuchten Merkmale Geschlecht (zwei Subgruppen) und Altersklasse (drei Subgruppen) wären für die Berechnung der gewichteten Summenindizes für beide Faktoren jeweils sechs Formeln entstanden, was aus Sicht der Anwender des WAI weniger praktikabel ist.

Insgesamt betrachtet ergibt sich aus der neuen Berechnungsmethode als gewichteter Summenindex zweier Faktoren für die Anwender in Betrieben ein höherer Berechnungsaufwand im Vergleich zur klassischen Berechnung des WAI. Bisher war die Berechnung des WAI ohne Zuhilfenahme eines statistischen Programms auch handschriftlich auf dem Papier möglich. Für die empfohlene gewichtete Berechnung wird dies durch die sieben Gewichte je Faktor, welche mit den z-standardisierten Werten der erfassten Itemrohwerte multipliziert werden müssen, deutlich erschwert. Für die anschließenden Vergleiche mit der Normwerttabelle muss zusätzlich eine lineare Transformation stattfinden.

Die Überprüfung der konvergenten Validität anhand korrelativer Zusammenhangsanalysen mit verwandten Instrumenten stellt eine einfache Variante von verschiedenen Möglichkeiten dar (Eid et al., 2013). Positiv zu bemerken ist hier die Verwendung der ausschließlich arbeitsbezogenen Items der NFAS. Weiterhin wäre es möglich, ein gemeinsames Strukturgleichungsmodell mit den verschiedenen Instrumenten zu untersuchen. Dabei könnte z. B. unter Berücksichtigung des Messfehlers und der Methodenspezifität der Zusammenhang der manifesten Items des SF-12 mit den latenten Faktoren des WAI betrachtet werden. Weitere Analysemöglichkeiten bieten die umfangreichen Multi-Trait-Multi-Method-Ansätze, mit denen die konvergenten und diskriminanten Zusammenhänge unterschiedlicher Erhebungsmethoden für verschiedene Konstrukte untersucht werden können.

### **5.3 Fazit und Ausblick**

Die Analysen der vorliegenden Arbeit haben gezeigt, dass die von den Entwicklern des WAI postulierte Eindimensionalität des Konstrukts nicht haltbar ist. Vielmehr wird mit diesem Fragebogen die Arbeitsfähigkeit durch zwei zusammenhängende Faktoren abgebildet. Der erste Faktor stellt dabei die subjektiv eingeschätzte Arbeitsfähigkeit unter Berücksichtigung der Arbeitsbedingungen und die individuellen Ressourcen dar. Der zweite Faktor wird durch die individuellen Gesundheitsbedingungen abgebildet. Die Untersuchung der Äquivalenz der Messungen für die Subgruppen Geschlecht und Altersgruppen wurde für den WAI erstmalig berichtet. Es konnte die Annahme gestützt werden, dass der WAI in den Subgruppen das gleiche Konstrukt

der Arbeitsfähigkeit erfasst. Dadurch sind Mittelwertvergleiche für die untersuchten Subgruppen zulässig.

Die neu entwickelte Berechnungsmethode anhand von gewichteten Summenwerten für beide Faktoren stellt zwar eine Herausforderung für die Anwender in der betrieblichen Praxis dar, wird jedoch begründet durch die Ergebnisse der konfirmatorischen Faktorenanalysen. Eine Nichtberücksichtigung würde in einer inkorrekten Interpretation der erfassten Arbeitsfähigkeit resultieren, sowohl bezogen auf die Einzelfalldiagnostik als auch auf Bevölkerungsanalysen. Zur Vereinfachung der Anwendung in Praxis und Forschung sollte in der Zukunft die Entwicklung eines Berechnungstools, welches Anwendern im Internet frei zur Verfügung gestellt werden kann, sowie die Bereitstellung der Syntax unter Berücksichtigung der Ergebnisse der vorliegenden Arbeit angestrebt werden. Dabei wäre es praktikabel, wenn unter Berücksichtigung des Geschlechts und der Altersgruppe, ein Vergleich mit im Tool hinterlegten Normwerten ermöglicht wird und zudem eine Kategorisierung der erfassten Arbeitsfähigkeit erfolgt. In diesem Fall wäre zukünftig auch die unterschiedliche Gewichtung unter Berücksichtigung der Gruppenzugehörigkeit umsetzbar.

Die bereitgestellten Normwerttabellen ermöglichen eine Einschätzung individueller Testwerte unter Berücksichtigung des Alters und des Geschlechts. Aufgrund der aus der mangelnden Messgenauigkeit des WAI resultierenden Spannweite der wahren Testwerte im Vergleich zum gemessenen Testwert sollte die Interpretation von Einzelergebnissen nur mit Vorsicht vorgenommen werden. In der Vergangenheit wurde bereits angeregt, dass Intervention und Prävention nicht erst bei älteren Beschäftigten ansetzen sollten (Freude & Pech, 2005; Ilmarinen, J., 2009). Da jedoch gerade die jüngeren Personen bessere Werte für die Faktoren der Arbeitsfähigkeit liefern, erhöht sich aufgrund der eingeschränkten Messgenauigkeit des WAI die Gefahr der Überschätzung der individuellen Arbeitsfähigkeit.

Insgesamt ist der Einsatz des WAI in Betrieben aufgrund seiner freien Verfügbarkeit, Verständlichkeit und Kürze angezeigt. Im Einklang mit den Empfehlungen von Bie-neck et al. (2005) sollte er einzig als Screening-Instrument eingesetzt werden, um einen ersten Überblick über die Arbeitsfähigkeit zu erhalten. Wie auch von Amler et al. (2015) angeregt, sollte der WAI dabei in Kombination mit anderen Instrumenten der Arbeits- und Funktionsfähigkeit Verwendung finden. Die Items des WAI bilden lediglich einen kleinen Teil des komplexen Konstrukts der Arbeitsfähigkeit ab und stützen sich dabei hauptsächlich auf die individuelle Gesundheit und die direkte Nachfrage der selbsteingeschätzten Arbeitsfähigkeit. Eine umfangreichere Erfassung der individuellen Arbeitsbedingungen sollte parallel oder spätestens im darauffolgenden Schritt durch den Einsatz etablierter Instrumente erfolgen. Der WAI allein ist eher für einen Überblick des Standes und der Entwicklung der Arbeitsfähigkeit in Bevölkerungs- und Beschäftigtenanalysen geeignet.

Die Analysen zur Validierung des WAI haben gezeigt, dass gerade die körperliche Gesundheit eine besondere Rolle für die individuelle Arbeitsfähigkeit spielt. Dies steht nicht im Einklang mit den Ergebnissen von J. Ilmarinen und Tuomi (2004), bei denen insbesondere die nicht-gesundheitsbezogenen Dimensionen des WAI die höchste Prädiktivität für die Berufsunfähigkeit und Mortalität aufzeigten. Zukünftige Analysen mit den Längsschnittdaten der S-MGA sollten daher die Kriteriumsvalidität der einzelnen Faktoren des WAI auf den vorzeitigen Erwerbsausstieg untersuchen.

## Literaturverzeichnis

**Abdolalizadeh, M., Arastoo, A. A., Ghsemzadeh, R., Montazeri, A., Ahmadi, K. & Azizi, A.** (2012). The psychometric properties of an Iranian translation of the Work Ability Index (WAI) questionnaire. *Journal of Occupational Rehabilitation*, 22, 401-408. doi:10.1007/s10926-012-9355-3

**Airila, A., Hakanen, J. J., Schaufeli, W. B., Luukkonen, R., Punakallio, A. & Lusa, S.** (2014). Are job and personal resources associated with work ability 10 years later? The mediating role of work engagement. *Work & Stress*, 28, 87-105. doi:10.1080/02678373.2013.872208

**Alavinia, S. M., de Boer, A. G., van Duivenbouden, J. C., Frings-Dresen, M. H. & Burdorf, A.** (2009). Determinants of work ability and its predictive value for disability. *Occupational medicine*, 59, 32-37. doi:10.1093/occmed/kqn148

**Alexopoulos, E., Merikoulias, G., Gnardellis, C. & Jelastopulu, E.** (2013). Work Ability Index: Validation of the Greek version and descriptive data in heavy industry employees. *British Journal of Medicine and Medical Research*, 3, 608-621. doi:10.9734/bjmmr/2013/2552

**Amler, N., Felder, S., Mau, W., Merkesdal, S., Schöffski, O. & und Mitglieder des Arbeitsgruppe.** (2015). Instrumente zur Messung von Effekten einer Frühintervention auf den Erhalt bzw. die Wiederherstellung der Arbeitsfähigkeit in Deutschland - Stellungnahme einer interdisziplinären Arbeitsgruppe (Online-Publikation). *Gesundheitswesen*. Zugriff am 25.08.2016. Verfügbar unter: doi: 10.1055/s-0041-110678

**Arbeit.** (2004). In J. Mittelstraß (Hrsg.), *Enzyklopädie Philosophie und Wissenschaftstheorie* (Unveränderte Sonderausgabe, Bd. 1, S. 151-152). Stuttgart: J.B. Metzler.

**Badura, B., Walter, U. & Hehlmann, T.** (2010). *Betriebliche Gesundheitspolitik: Der Weg zur gesunden Organisation*. Berlin: Springer.

**Bethge, M., Radoschewski, F. M. & Gutenbrunner, C.** (2012). The Work Ability Index as a screening tool to identify the need for rehabilitation: longitudinal findings from the Second German Sociomedical Panel of Employees. *Journal of rehabilitation medicine*, 44, 980-987. doi:10.2340/16501977-1063

**Bieneck, H. J., Sedlatschek, C., Kuhn, K., Freude, G. & Pech, E.** (2005). Position der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin in der Debatte um den „Work Ability Index“. *Gute Arbeit. Zeitschrift für Gesundheitsschutz und Arbeitsgestaltung*, 17 (7), 36-39.

**Brage, S., Fleten, N., Knudsrød, O., Reiso, H. & Ryen, A.** (2004). Norsk Funksjonsskjema – et nytt instrument ved sykmelding og uførhetsvurdering [Norwegian Functional Scale - a new instrument in sickness certification and disability assessments]. *Tidsskr Nor Laegeforen*, 124, 2472-2474.

**Brown, T. A.** (2006). *Confirmatory Factor Analysis for applied research*. New York: Guilford Press.

**Brussig, M. & Knuth, M.** (2009). Individuelle Beschäftigungsfähigkeit: Konzept, Operationalisierung und erste Ergebnisse. *WSI Mitteilungen*, 6, 287-294.

**Bühner, M.** (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. akt. Aufl.). München: Pearson Studium.

**Bullinger, M., Kirchberger, I. & Ware, J.** (1995). Der deutsche SF-36 Health Survey Übersetzung und psychometrische Testung eines krankheitsübergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualität. *Journal of Public Health*, 3, 21-36. doi:10.1007/bf02959944

**Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.** (2002). *Europäische Erfahrungen mit dem Arbeitsbewältigungsindex (Work Ability Index)*. Dortmund: Autor.

**Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.** (2013). *Why WAI? - Der Work Ability Index im Einsatz für Arbeitsfähigkeit und Prävention. Erfahrungsberichte aus der Praxis* (5. akt. Aufl.). Dortmund: Autor.

**Bundesministerium für Arbeit und Soziales.** (2015). *Grünbuch Arbeiten 4.0 - Arbeit weiter denken*. Berlin: Autor.

**Byrne, B. M., Shavelson, R. J. & Muthén, B.** (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456-466. doi:10.1037/0033-2909.105.3.456

**Cheak-Zamora, N. C., Wyrwich, K. W. & McBride, T. D.** (2009). Reliability and validity of the SF-12v2 in the medical expenditure panel survey. *Quality of life research*, 18, 727-735. doi:10.1007/s11136-009-9483-1

**Chen, J. J.** (2007). Functional Capacity Evaluation & Disability. *The Iowa Orthopaedic Journal*, 27, 121-127.

**Cheung, G. W. & Rensvold, R. B.** (2002). Evaluating Goodness-of-Fit indexes for testing measurement invariance. *Structural equation modeling: A Multidisciplinary Journal*, 9, 233-255. doi:10.1207/S15328007SEM0902\_5

**Christ, O. & Schlüter, E.** (2012). *Strukturgleichungsmodelle mit Mplus – Eine praktische Einführung*. München: Oldenbourg.

**Cohen, J.** (1988). *Statistical power analysis for the behavioral sciences* (2. Auflage). Hillsdale, NJ: Erlbaum.

**DiStefano, C., Zhu, M. & Mîndrilă, D.** (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, 14 (20), 1-11.

- Eichhorst, W.** (2015). Der Wandel der Erwerbsformen in Deutschland. *Gesundheits- und Sozialpolitik*, 69 (1), 15-22. doi:10.5771/1611-5821-2015-1-15
- Eichhorst, W., Tobsch, V. & Wehner, C.** (2016). Neue Qualität der Arbeit? Zur Entwicklung von Arbeitskulturen und Fehlzeiten. In B. Badura, A. Ducki, H. Schröder, J. Klose & M. Mayer (Hrsg.), *Fehlzeiten-Report 2016*. Berlin: Springer.
- Eid, M., Gollwitzer, M. & Schmitt, M.** (2013). *Statistik und Forschungsmethoden* (3., korr. Auflage). Weinheim: Beltz.
- El Fassi, M., Bocquet, V., Majery, N., Lair, M. L., Couffignal, S. & Mairiaux, P.** (2013). Work ability assessment in a worker population: comparison and determinants of Work Ability Index and Work Ability score. *BMC Public Health*, 13, 305-314. doi:10.1186/1471-2458-13-305
- Elsner, G.** (2005). Der Arbeitsbewältigungsindex: Eine Bewertung aus arbeitsmedizinischer Sicht. *Gute Arbeit. Zeitschrift für Gesundheitsschutz und Arbeitsgestaltung*, 17 (2), 18-21.
- Eskelinen, L., Kohvakka, A., Merisalo, T., Hurri, H. & Wägar, G.** (1991). Relationship between the self-assessment and clinical assessment of health status and work ability. *Scandinavian journal of work, environment & health*, 17 (suppl 1), 40-47.
- Evers, A.** (2001). The revised Dutch Rating System for Test Quality. *International Journal of Testing*, 1, 155-182. doi:10.1207/S15327574IJT0102\_4
- Field, A.** (2013). *Discovering statistics using IBM SPSS Statistics* (4. Auflage). London: SAGE Publications.
- Flora, D. B. & Curran, P. J.** (2004). An empirical evaluation of alternative methods of estimation for Confirmatory Factor Analysis with ordinal data. *Psychological Methods*, 9, 466-491. doi:10.1037/1082-989x.9.4.466
- Freude, G. & Pech, E.** (2005). Demographischer Wandel, Gesundheit und Arbeitsfähigkeit. In J. Kerschbaumer & W. Schroeder (Hrsg.), *Sozialstaat und demographischer Wandel* (S. 185-222). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gadermann, A. M., Guhn, M. & Zumbo, B. D.** (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, 17 (3), 1-13.
- Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E. et al.** (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *Journal of clinical epidemiology*, 51, 1171-1178.
- Geiser, C.** (2011). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung* (2. Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.



**Gemeinsamer Bundesausschuss.** (2015). Richtlinie des Gemeinsamen Bundesausschusses über die Beurteilung der Arbeitsunfähigkeit und die Maßnahmen zur stufenweisen Wiedereingliederung nach § 92 Abs. 1 Satz 2 Nr. 7 SGB V. In Kraft getreten am 28. Januar 2014. Zuletzt geändert am 17. Dezember 2015.

**Gesellschaft für Arbeitswissenschaft e.V.** (2016). Selbstverständnis der Gesellschaft für Arbeitswissenschaft e.V. (GfA). Zugriff am 25.08.2016. Verfügbar unter: [http://www.gesellschaft-fuer-arbeitswissenschaft.de/wir-ueber-uns\\_selbstverstaendnis-gesellschaft-fuer-arbeitswissenschaft-gfa.htm](http://www.gesellschaft-fuer-arbeitswissenschaft.de/wir-ueber-uns_selbstverstaendnis-gesellschaft-fuer-arbeitswissenschaft-gfa.htm)

**Gould, R., Ilmarinen, J., Järvisalo, J. & Koskinen, S.** (2008). Dimensions of work ability - Results of the Health 2000 Survey. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

**Gould, R., Koskinen, S., Seitsamo, J., Tuomi, K., Polvinen, A. & Sainio, P.** (2008). Data and methods. In R. Gould, J. Ilmarinen, J. Järvisalo & S. Koskinen (Hrsg.), Dimensions of work ability - Results of the Health 2000 Survey. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

**Gould, R. & Polvinen, A.** (2008). Work ability in different population groups - occupation. In R. Gould, J. Ilmarinen, J. Järvisalo & S. Koskinen (Hrsg.), Dimensions of work ability - Results of the Health 2000 Survey. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

**Gould, R., Polvinen, A. & Seitsamo, J.** (2008). Work ability in different population groups - age and gender. In R. Gould, J. Ilmarinen, J. Järvisalo & S. Koskinen (Hrsg.), Dimensions of work ability - Results of the Health 2000 Survey. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

**Graham, J. W.** (2009). Missing data analysis: making it work in the real world. Annual review of psychology, 60, 549-576. doi:10.1146/annurev.psych.58.110405.085530

**Grice, J. W.** (2001). A comparison of factor scores under conditions of factor obliquity. Psychological Methods, 6, 67-83. doi:10.1037/1082-989x.6.1.67

**Hoffmann, E. & Walwei, U.** (1998). Normalarbeitsverhältnis: ein Auslaufmodell? Überlegungen zu einem Erklärungsmodell für den Wandel der Beschäftigungsformen. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 31, 409-425.

**Hu, L. & Bentler, P. M.** (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3, 424. doi:10.1037/1082-989X.3.4.424

**Hu, L. & Bentler, P. M.** (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6, 1-55. doi:10.1080/10705519909540118

**Ilmarinen, J.** (2004). Preface. In J. Ilmarinen & S. Lehtinen (Hrsg.), *Past, Present and Future of Work Ability*. Helsinki: Finnish Institute of Occupational Health.

**Ilmarinen, J.** (2009). Work ability—a comprehensive concept for occupational health research and prevention. *Scandinavian journal of work, environment & health*, 35, 1-5. doi:10.5271/sjweh.1304

**Ilmarinen, J.** (2011). Arbeitsfähig in die Zukunft. In M. Giesert (Hrsg.), *Arbeitsfähig in die Zukunft. Willkommen im Haus der Arbeitsfähigkeit*. Hamburg: VSA.

**Ilmarinen, J.** (2016, März). WAI 2.0, Vortrag auf der Jahrestagung "Arbeitsleben 2025 - Arbeitsfähig in die Zukunft" des Instituts für Arbeitsfähigkeit, Berlin.

**Ilmarinen, J., Gould, R., Järvikoski, A. & Järvisalo, J.** (2008). Diversity of work ability. In R. Gould, J. Ilmarinen, J. Järvisalo & S. Koskinen (Hrsg.), *Dimensions of work ability - Results of the Health 2000 Survey*. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

**Ilmarinen, J. & Tempel, J.** (2002). *Arbeitsfähigkeit 2010. Was können wir tun, damit Sie gesund bleiben?* Hamburg: VSA.

**Ilmarinen, J. & Tuomi, K.** (2004). Past, present and future of work ability. In J. Ilmarinen & S. Lehtinen (Hrsg.), *Past, present and future of work ability*. Helsinki: Finnish Institute of Occupational Health.

**Ilmarinen, J., Tuomi, K. & Klockars, M.** (1997). Changes in the work ability of active employees over an 11-year period. *Scandinavian Journal of Work Environment & Health*, 23 (suppl 1), 49-57.

**Ilmarinen, V., Ilmarinen, J., Huuhtanen, P., Louhevaara, V. & Näsman, O.** (2015). Examining the factorial structure, measurement invariance and convergent and discriminant validity of a novel self-report measure of work ability: work ability - personal radar. *Ergonomics*, 58, 1445-1460. doi:10.1080/00140139.2015.1005167

**Jankowiak, S. & Kersten, N.** (2016, September). Norwegian Function Assessment Scale: Reduktion und Validierung für die Arbeitswelt. Paper presented at the Gesundheit – bio.psycho.sozial 2.0. Gemeinsame Jahrestagung der DGMS und der DGMP, Berlin.

**Kerschbaumer, J. & Schroeder, W.** (2005). *Sozialstaat und demographischer Wandel*. Wiesbaden: VS Verlag für Sozialwissenschaften.

**Kessler, R. C., Barber, C., Beck, A., Berglund, P., Cleary, P. D., McKenas, D. et al.** (2003). The world health organization health and work performance questionnaire (HPQ). *Journal of Occupational and Environmental Medicine*, 45, 156-174.

**Kline, R. B.** (2011). *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press.

**Knieps, F. & Pfaff, H.** (2015). Langzeiterkrankungen. BKK Gesundheitsreport 2015. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft mbH & Co. KG.

**Lederer, V., Loisel, P., Rivard, M. & Champagne, F.** (2014). Exploring the diversity of conceptualizations of work (dis)ability: A scoping review of published definitions. *Journal of Occupational Rehabilitation*, 24, 242-267. doi:10.1007/s10926-013-9459-4

**Lerner, D., Amick, B. C., Rogers, W. H., Malspeis, S., Bungay, K. & Cynn, D.** (2001). The work limitations questionnaire. *Medical Care*, 39, 72-85.

**Lienert, G. A. & Raatz, U.** (1998). Testaufbau und Testanalyse (6. Auflage). Weinheim: Beltz.

**Little, T. D., Card, N. A., Slegers, D. W. & Ledford, E. C.** (2007). Representing contextual effects in multiple-group MACS models. In T. D. Little, J. A. Bovaird & N. A. Card (Hrsg.), *Modeling contextual effects in longitudinal studies*. (S. 121–147). Mahwah, NJ: Erlbaum.

**Marsh, H. W., Hau, K.-T. & Wen, Z.** (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling: A Multidisciplinary Journal*, 11, 320-341.

**Martinez, M. C., Latorre, M. d. R. D. d. O. & Fischer, F. M.** (2009). Validity and reliability of the Brazilian version of the Work Ability Index questionnaire. *Revista de Saúde Pública*, 43, 525-532.

**Martus, P., Freude, G., Rose, U., Seibt, R. & Jakob, O.** (2011). Arbeits- und gesundheitsbezogene Determinanten von Vitalität und Arbeitsfähigkeit. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.

**Martus, P., Jakob, O., Rose, U., Seibt, R. & Freude, G.** (2010). A comparative analysis of the Work Ability Index. *Occupational medicine*, 60, 517-524. doi:10.1093/occmed/kqq093

**Meredith, W.** (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543. doi:10.1007/bf02294825

**Mosier, C. I.** (1943). On the reliability of a weighted composite. *Psychometrika*, 8, 161-168. doi:10.1007/bf02288700

**Muthén, L. K. & Muthén, B. O.** (1998-2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Autor.

**Nerdinger, F. W., Blickle, G. & Schaper, N.** (2014). *Arbeits-und Organisationspsychologie* (3. Auflage). Heidelberg: Springer. doi: 10.1007/978-3-642-41130-4

**Nöllenheidt, C. & Brenscheidt, S.** (2015). *Arbeitswelt im Wandel - Zahlen - Daten - Fakten*. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.

**Nordenfelt, L.** (2008). The concept of work ability. Brüssel: P.I.E. Peter Lang.

**Nübling, M., Andersen, H. H. & Mühlbacher, A.** (2006). Entwicklung eines Verfahrens zur Berechnung der körperlichen und psychischen Summenskalen auf Basis der SOEP-Version des SF 12 (Algorithmus). Data Documentation 16. Berlin: Deutsches Institut für Wirtschaftsforschung.

**Nübling, M., Hasselhorn, H. M., Seitsamo, J. & Ilmarinen, J.** (2005). Work Ability Index Fragebogen: Vergleich kurzer und langer Krankheitsliste. Arbeits-med.Sozialmed.Umweltmed., 40 (3), 180-181.

**Østerås, N., Brage, S., Garratt, A., Benth, J. S., Natvig, B. & Gulbrandsen, P.** (2007). Functional ability in a population: normative survey data and reliability for the ICF based Norwegian Function Assessment Scale. BMC Public Health, 7, 278-286. doi:10.1186/1471-2458-7-278

**Østerås, N., Gulbrandsen, P., Garratt, A., Benth, J. S., Dahl, F. A., Natvig, B. et al.** (2008). A randomised comparison of a four- and a five-point scale version of the Norwegian Function Assessment Scale. Health and quality of life outcomes, 6, 14-22. doi:10.1186/1477-7525-6-14

**Peralta, N., Godoi Vasconcelos, A. G., Härter Griep, R. & Miller, L.** (2012). Validity and reliability of the Work Ability Index in primary care workers in Argentina. Salud Colectiva, 8, 163-173. doi:10.1590/S1851-82652012000200005

**Prümper, J. & Riechenhagen, G.** (2011). Von der Arbeitsunfähigkeit zum Haus der Arbeitsfähigkeit: Der Work Ability Index und seine Anwendung. In B. Seyfried (Hrsg.), Ältere Beschäftigte: Zu jung, um alt zu sein. Konzepte - Forschungsergebnisse - Instrumente (S. 135-146). Bielefeld: Bertelsmann.

**Radkiewicz, P. & Widerszal-Bazyl, M.** (2005). Psychometric properties of Work Ability Index in the light of comparative survey study. International Congress Series, 1280, 304-309. doi:10.1016/j.ics.2005.02.089

**Rebscher, H.** (2016). DAK Gesundheitsreport 2016. Analyse der Arbeitsunfähigkeitsdaten. Schwerpunkt: Gender und Gesundheit. Hamburg: medhochzwei.

**Reilly, M. C., Zbrozek, A. S. & Dukes, E. M.** (1993). The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics, 4, 353-365. doi:10.2165/00019053-199304050-00006

**Riechenhagen, G.** (2011). Arbeitsfähigkeit - Arbeitsunfähigkeit - Arbeitsschutz. In M. Giesert (Hrsg.), Arbeitsfähigkeit in die Zukunft. Willkommen im Haus der Arbeitsfähigkeit. Hamburg: VSA.

**Rongen, A., Robroek, S. J., Schaufeli, W. & Burdorf, A.** (2014). The contribution of work engagement to self-perceived health, work ability, and sickness absence beyond health behaviors and work-related factors. Journal of occupational and environmental medicine, 56, 892-897. doi:10.1097/JOM.000000000000196

**Rose, U., Schiel, S., Schroder, H., Kleudgen, M., Tophoven, S., Rauch, A. et al.** (2017). The Study on Mental Health at Work: Design and sampling. *Scandinavian Journal of Public Health*, 45(6), 584-594. doi: 10.1177/1403494817707123

**Rosenthal, R. & DiMatteo, M. R.** (2001). Meta-analysis: recent developments in quantitative methods for literature reviews. *Annual review of psychology*, 52, 59-82. doi:10.1146/annurev.psych.52.1.59

**Salonen, P., Arola, H., Nygard, C.-H., Huhtala, H. & Koivisto, A.-M.** (2003). Factors associated with premature departure from working life among ageing food industry employees. *Occupational medicine*, 53, 65-68. doi:10.1093/occmed/kqg012

**Schermelleh-Engel, K., Moosbrugger, H. & Müller, H.** (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8 (2), 23-74. Zugriff am 25.08.2016. Verfügbar unter <http://www.dgps.de/fachgruppen/methoden/mpr-online/>

**Schröder, H., Schiel, S., Schulz, S. & Kleudgen, M.** (2015). Mentale Gesundheit bei der Arbeit (S-MGA) Methodenbericht zur Repräsentativerhebung an Erwerbstätigen in Deutschland. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.

**Schuntermann, M. F.** (2005). Einführung in die ICF. Grundkurs - Übungen - offene Fragen (1. Auflage). Landsberg am Lech: ecomed Medizin.

**Seitsamo, J., Tuomi, K., Ilmarinen, J. & Gould, R.** (2008). Work and the work environment. In R. Gould, J. Ilmarinen, J. Järvisalo & S. Koskinen (Hrsg.), *Dimensions of work ability - Results of the Health 2000 Survey*. Helsinki: Finnish Centre for Pensions, The Social Insurance Institution, National Public Health Institute, Finnish Institute of Occupational Health.

**Semmer, N. K. & Meier, L. L.** (2014). Bedeutung und Wirkung von Arbeit. In H. Schuler & K. Moser (Hrsg.), *Lehrbuch Organisationspsychologie*. Bern: Hogrefe.

**Soer, R., van der Schans, C. P., Groothoff, J. W., Geertzen, J. H. & Reneman, M. F.** (2008). Towards consensus in operational definitions in functional capacity evaluation: A Delphi Survey. *Journal of Occupational Rehabilitation*, 18, 389-400. doi: 10.1007/s10926-008-9155-y

**Sozialgesetzbuch (SGB III).** Drittes Buch Arbeitsförderung. Inkraftgetreten am 01. Januar 1998. Letzte Änderung vom 31. Juli 2016 (Bundesgesetzblatt Teil I Nr. 39 S. 1939).

**Statistisches Bundesamt.** (2015). *Statistisches Jahrbuch 2015*. Wiesbaden: Autor.

**Steyer, R. & Eid, M.** (2001). *Messen und Testen*. Berlin: Springer. doi:10.1007/978-3-642-56924-1

**Tempel, J. & Ilmarinen, J.** (2013). *Arbeitsleben 2025. Das Haus der Arbeitsfähigkeit im Unternehmen bauen*. Hamburg: VSA.

**Tengland, P. A.** (2011). The concept of work ability. *Journal of Occupational Rehabilitation*, 21, 275-285. doi:10.1007/s10926-010-9269-x

**Tuomi, K., Ilmarinen, J., Jahkola, A., Katajarinne, L. & Tulkki, A.** (1998). *Work Ability Index* (2. Auflage). Helsinki: Finnish Institute of Occupational Health.

**Tuomi, K., Ilmarinen, J., Seitsamo, J., Huuhtanen, P., Martikainen, R., Nygard, C.-H. et al.** (1997). Summary of the Finnish research project (1981-1992) to promote the health and work ability of ageing workers. *Scandinavian Journal of Work Environment & Health*, 23, 66-71.

**Tuomi, K., Wägar, G., Eskelinen, L., Järvinen, E., Huuhtanen, P., Suurnäkki, T. et al.** (1985). Terveys, työkyky ja työolot kunnallisissa ammattiryhmissä [Health, work capacity and work conditions in municipal occupations]. *Työterveyslaitoksen tutkimuksia*, 3, 95-132.

**van den Berg, T. I., Elders, L. A., de Zwart, B. C. & Burdorf, A.** (2009). The effects of work-related and individual factors on the Work Ability Index: A systematic review. *Occupational and environmental medicine*, 66, 211-220. doi:10.1136/oem.2008.039883

**Voß, G. G.** (2010). Was ist Arbeit? Zum Problem eines allgemeinen Arbeitsbegriffs. In F. Böhle, G. G. Voß & G. Wachtler (Hrsg.), *Handbuch Arbeitssoziologie* (S. 23-80). Wiesbaden: VS Verlag für Sozialwissenschaften.

**WAI-Netzwerk.** (2015). *WAI-Manual. Anwendung des Work-Ability Index*. Zugriff am 25.08.2016. Verfügbar unter: <http://www.arbeitsfaehig.com/de/wai-manual-347.html>

**Ware Jr., J. E.** (2000). SF-36 Health Survey Update. *Spine*, 25, 3130-3139. doi:10.1097/00007632-200012150-00008

**Ware Jr., J. E., Kosinski, M., Dewey, J. E. & Gandek, B.** (2001). How to score and interpret single-item health status measures: A manual for users of the SF-8 Health Survey. Lincoln, RI: QualityMetric Incorporated.

**Ware Jr., J. E., Kosinski, M. & Keller, S. D.** (1996). A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care*, 34, 220-233.

**Ware Jr., J. E., Kosinski, M., Turner-Bowker, D. & Gandek, B.** (2002). SF12v2: How to score version 2 of the SF-12 health survey. Lincoln, RI: QualityMetric Incorporated.

**Ware Jr., J. E., Snow, K. K., Kosinski, M. & Gandek, B.** (1993). *SF-36 Health Survey: Manual and interpretation guide*. Boston: The Health Institute, New England Medical Center.

**World Health Organization.** (2001). *International Classification of Functioning, Disability and Health: ICF*. Genf: Autor.

## Abkürzungsverzeichnis

(w_)MD	(arbeitsbezogene) psychische Domäne des NFAS
(w_)PD	(arbeitsbezogene) körperliche Domäne des NFAS
BAuA	Bundesanstalt für Arbeitsschutz und Arbeitsmedizin
BGM	Betriebliches Gesundheitsmanagement
CFI	Comparative-Fit-Index
COTAN	Committee On Test Affairs Netherlands
FCE	Functional Capacity Evaluation
FIOH	Finnisches Institut für Arbeitsmedizin (Finnish Institute of Occupational Health)
HPQ	Health and Work Performance Questionnaire
ICF	Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit
infas	Institut für angewandte Sozialwissenschaft GmbH
KI	Konfidenzintervall
MCS	Mental Component Summary (psychische Komponente des SF-12)
NFAS(-w)	(arbeitsbezogene) Norwegian Function Assessment Scale
PCS	Physical Component Summary (körperliche Komponente des SF-12)
RMSEA	Root-Mean-Square-Error of Approximation
SF-8	Short-Form-8 Health Survey
SF-12	Short-Form-12 Health Survey
SF-36	Short-Form-36 Health Survey
S-MGA	Studie zur Mentalen Gesundheit bei der Arbeit
TLI	Tucker-Lewis-Index
VIF	Varianz-Inflations-Faktor
WAI	Work Ability Index
WAS	Work Ability Score
WHO	World Health Organization
WLSMV	Mean- and variance-adjusted weighted least squares
WLQ	Work Limitations Questionnaire
WPAI	Work Produktivity and Activity Impairment Questionnaire

## Anhang 1

### Work Ability Index (nach Tuomi et al., 1998)<sup>1</sup>

#### (0) Art der Tätigkeit: geistig/körperlich

Sind Sie bei Ihrer Arbeit...

- 1 vorwiegend geistig tätig
- 2 vorwiegend körperlich tätig
- 3 etwa gleichermaßen geistig und körperlich tätig
- 97 verweigert
- 98 weiß nicht

#### (1) Derzeitige Arbeitsfähigkeit im Vergleich zu der besten je erreichten Arbeitsfähigkeit

Wenn Sie Ihre beste, je erreichte Arbeitsfähigkeit mit 10 Punkten bewerten: Wie viele Punkte würden Sie dann für Ihre derzeitige Arbeitsfähigkeit geben? Null bedeutet, dass Sie derzeit arbeitsunfähig sind.

- 0 völlig arbeitsunfähig
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 die beste jemals erreichte Arbeitsfähigkeit
- 97 verweigert
- 98 weiß nicht

#### (2) Arbeitsfähigkeit im Vergleich zu den Anforderungen der Arbeitstätigkeit\*

(a) Wie schätzen Sie Ihre derzeitige Arbeitsfähigkeit in Bezug auf die körperlichen Arbeitsanforderungen ein? Ist sie...

- 1 sehr gut
- 2 eher gut
- 3 mittelmäßig
- 4 eher schlecht
- 5 sehr schlecht
- 97 verweigert
- 98 weiß nicht

(b) Und wie schätzen Sie Ihre derzeitige Arbeitsfähigkeit in Bezug auf die geistigen Arbeitsanforderungen ein?

- 1 sehr gut
- 2 eher gut

<sup>1</sup> Mit \* versehene Items müssen rekodiert werden. Berechnung gemäß WAI-Handbuch (WAI-Netzwerk, 2015).



- 3 mittelmäßig
- 4 eher schlecht
- 5 sehr schlecht
- 97 verweigert
- 98 weiß nicht

### (3) Anzahl der aktuellen vom Arzt diagnostizierten Krankheiten

Nun möchte ich Ihnen einige Fragen zu Ihrer Krankheitsgeschichte stellen. Auf dieser Liste sind mehrere Krankheiten aufgeführt. Hat ein Arzt bei Ihnen jemals eine der hier aufgeführten Krankheiten oder Gesundheitsstörungen festgestellt?

- (a) Unfallverletzungen (z. B. des Rückens, der Glieder, Verbrennungen)
- (b) Erkrankungen des Muskel-Skelett-Systems von Rücken, Gliedern oder anderen Körperteilen (z. B. wiederholte Schmerzen in Gelenken oder Muskeln, Ischias, Rheuma, Wirbelsäulenerkrankungen)
- (c) Herz-Kreislauf-Erkrankungen (z. B. Bluthochdruck, Herzkrankheit, Herzinfarkt)
- (d) Atemwegserkrankungen (z. B. wiederholte Atemwegsinfektionen, chronische Bronchitis, Bronchialasthma)
- (e) Psychische Beeinträchtigungen (z. B. Depressionen, Angstzustände, chronische Schlaflosigkeit, psychovegetatives Erschöpfungssyndrom)
- (f) Neurologische und sensorische Erkrankungen (z. B. Tinnitus, Hörschäden, Augenerkrankungen, Migräne, Epilepsie)
- (g) Erkrankungen des Verdauungssystems (z. B. der Gallenblase, Leber, Bauchspeicheldrüse, Darm)
- (h) Erkrankungen im Urogenitaltrakt (z. B. Harnwegsinfektionen, gynäkologische Erkrankungen)
- (i) Hautkrankheiten (z. B. allergischer Hautausschlag, Ekzem)
- (j) Tumore / Krebs
- (k) Hormon- / Stoffwechselerkrankungen (z. B. Diabetes, Fettleibigkeit, Schilddrüsenprobleme)
- (l) Krankheiten des Blutes (z. B. Anämie)
- (m) Angeborene Leiden/Erkrankungen
- (n) Andere Leiden oder Krankheiten, und zwar \_\_\_\_\_ (nachkodiert)
  - 1 ja
  - 2 nein
  - 7 verweigert
  - 8 weiß nicht

Jeweils: Wann ist dies zuletzt aufgetreten?

- 1 innerhalb der letzten vier Wochen
- 2 innerhalb der letzten 12 Monate
- 3 länger her
- 7 verweigert
- 8 weiß nicht

### (4) Geschätzte Beeinträchtigung der Arbeitsleistung durch die Krankheiten\*

Nun geht es um eine Erkrankung oder Verletzung, die Sie möglicherweise derzeit bei der Arbeit behindern. Bitte sagen Sie mir, welche der folgenden Aussagen am ehesten auf Sie zutrifft.

- 1 Ich habe keine Beeinträchtigung, ich habe keine Erkrankung
- 2 Ich kann meine Arbeit ausführen, habe aber Beschwerden
- 3 Manchmal gezwungen, langsamer zu arbeiten/Methoden zu ändern
- 4 Oft gezwungen, langsamer zu arbeiten/Methoden zu ändern
- 5 Wegen Krankheit bin ich nur zu Teilzeitarbeit in der Lage
- 6 Meiner Meinung nach bin ich völlig arbeitsunfähig
- 7 verweigert
- 8 weiß nicht

**(5) Krankenstandstage in den vergangenen 12 Monaten\***

Wie viele ganze Tage blieben Sie auf Grund eines gesundheitlichen Problems, wie einer Krankheit, Gesundheitsvorsorge oder Untersuchung, in den letzten 12 Monaten der Arbeit fern?

- 1 überhaupt keinen Tag
- 2 höchstens 9 Tage
- 3 10 - 24 Tage
- 4 25 - 99 Tage
- 5 100 Tage und mehr
- 7 verweigert
- 8 weiß nicht

**(6) Einschätzung der eigenen Arbeitsfähigkeit in zwei Jahren\***

Glauben Sie, dass Sie, ausgehend von Ihrem jetzigen Gesundheitszustand, Ihre derzeitige Arbeit auch in den nächsten Jahren ausüben können?

- 1 nein, halte ich für unwahrscheinlich
- 2 ich bin mir nicht sicher
- 3 ja, ziemlich sicher
- 7 verweigert
- 8 weiß nicht

**(7) Psychische Leistungsreserven\***

(a) Haben Sie in der letzten Zeit Ihre täglichen Aufgaben mit Freude erledigt?

- 1 häufig
- 2 eher häufig
- 3 manchmal
- 4 eher selten
- 5 niemals
- 7 verweigert
- 8 weiß nicht

(b) Waren Sie in letzter Zeit aktiv und rege?

- 1 immer
- 2 eher häufig
- 3 manchmal
- 4 eher selten
- 5 niemals
- 7 verweigert
- 8 weiß nicht

(c) Waren Sie in der letzten Zeit zuversichtlich, was die Zukunft betrifft?

- 1 ständig
- 2 eher häufig
- 3 manchmal
- 4 eher selten
- 5 niemals
- 7 verweigert
- 8 weiß nicht

## Anhang 2

**Anh. 2, Tab. 1** Soziodemografische Daten der Stichprobe (N = 3 968)

	n <sup>a</sup>	Prozent
Altersgruppen		
5-Jahresschritte		
31 - 35 Jahre	399	10.1
36 - 40 Jahre	566	14.3
41 - 45 Jahre	792	20.0
46 - 50 Jahre	866	21.8
51 - 55 Jahre	746	18.8
56 - 60 Jahre	599	15.1
10-Jahresschritte		
31 - 40	965	24.3
41 - 50	1 658	41.8
51 - 60	1 345	33.9
Geschlecht		
männlich	2 029	51.1
weiblich	1 939	48.9
höchster Schulabschluss		
Hauptschulabschluss	1 013	25.5
Realschulabschluss	1 608	40.5
(Fach-)Hochschulreife	1 278	32.2
anderer/ohne Abschluss beendet	68	1.7
höchster Ausbildungsabschluss		
Abschluss einer beruflich-betrieblichen Berufsausbildung	2 013	50.7
Abschluss einer beruflich-schulischen Ausbildung	315	7.9
Abschluss an einer Fach-, Meister-/Technikerschule, Berufsakademie	599	15.1
Fachhochschulabschluss	346	8.7
Hochschulabschluss	483	12.2
Keinen/einen anderen Abschluss	211	5.3
derzeit hauptberufliche Tätigkeit		
Vollzeit erwerbstätig mit mindestens 35 Stunden pro Woche	2 879	72.6
Teilzeit erwerbstätig zwischen 14 und 34 Stunden pro Woche	945	23.8
geringfügig, gelegentlich oder unregelmäßig beschäftigt	144	3.6
Art des Arbeitsvertrages		
zeitlich unbefristet	3 690	93.0
zeitlich befristet	222	5.6
sonstiges/fehlend	56	1.4
in Deutschland geboren		
ja	3 548	89.4
nein	419	10.6

Anmerkung. <sup>a</sup>Abweichungen vom Gesamt-N durch fehlende Werte/Verweigerungen.