

# 1218<sup>2025</sup>

**SOEP** papers  
on Multidisciplinary Panel Data Research

## Validität und Reliabilität der ein- und zweistufigen Version des Effort-Reward Imbalance Modells in der 33. Welle des Sozio-oekonomischen Panels

Mandy Müller

## **SOEPPapers on Multidisciplinary Panel Data Research** at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPPapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPPapers are available at <http://www.diw.de/soeppapers>

### **Editors:**

Jan **Goebel** (Spatial Economics)

Christian **Hunkler** (Migration)

Philipp **Lersch** (Sociology, Demography)

Levent **Neyse** (Behavioral and Experimental Economics)

Carsten **Schröder** (Public Economics)

Sabine **Zinn** (Statistics)

Conchita **D'Ambrosio** (Public Economics, DIW Research Fellow)

Denis **Gerstorff** (Psychology, DIW Research Fellow)

Martin **Kroh** (Political Science, Survey Methodology)

Stefan **Liebig** (Sociology)

David **Richter** (Psychology)

Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Fellow)

Thomas **Siedler** (Empirical Economics, DIW Research Fellow)

C. Katharina **Spieß** (Education and Family Economics)

Katharina **Wrohlich** (Gender Economics)

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)

DIW Berlin

Mohrenstrasse 58

10117 Berlin, Germany

Contact: [soeppapers@diw.de](mailto:soeppapers@diw.de)



# Validität und Reliabilität der ein- und zweistufigen Version des Effort-Reward Imbalance Modells in der 33. Welle des Sozio-oekonomischen Panels

**Müller, Mandy**

Institut für Personal und Arbeit

Helmut-Schmidt-Universität

E-Mail: [mueller@hsu-hh.de](mailto:mueller@hsu-hh.de)

<https://www.hsu-hh.de/ipa/mueller>

## Abstract

**Hintergrund:** Mit den Daten des Sozio-oekonomischen Panels (SOEP) 2016 ist erstmalig und einmalig eine Überprüfung der Vergleichbarkeit der ein- und zweistufigen Messmethode des Effort-Reward-Imbalance Modells im Sozio-oekonomischen Panel möglich.

**Methodik:** Die Reliabilität wird mit deskriptiven Statistiken, Inter-Item-Korrelation und Item-Skala Statistiken überprüft. Die Konstrukt- und Kriteriumsvalidität wird mit einer konfirmatorischen Faktoren- und einer multiplen Regressionsanalyse getestet.

**Ergebnisse:** Die deskriptiven Statistiken zeigen, dass die Mittelwerte der beiden Messverfahren nicht vergleichbar sind. Die Konstruktvalidität hingegen ist in beiden Messmethoden ähnlich, beide geben die theoretische Struktur des Effort-Reward-Imbalance Modells wieder. Für die Kriteriumsvalidität kann aufgrund der nur wenigen vorhandenen Variablen festgehalten werden, dass die Höhe und die Signifikanzen bei denselben Variablen weitestgehend übereinstimmen und daher unabhängig von der Messmethode sind.

**Schlussfolgerungen:** Die vorliegende Arbeit konnte die Forschungslücke bezüglich der Vergleichbarkeit der beiden existierenden Messmethoden in Bezug auf das Sozio-oekonomische Panel schließen. Die Ergebnisse sind ein erster Anhaltspunkt und müssen in weiteren umfangreichen Bevölkerungsumfragen oder Meta-Analysen bestätigt werden.

**Schlüsselwörter:** Stress, Modell beruflicher Gratifikationskrisen, Gesundheitliche Bedingungen, Berufliche Belastung

## I Einleitung

Die Forschung der letzten Jahrzehnte zeigt, dass eine dauerhaft höhere psychische Belastung im Beruf negative Auswirkungen auf die Gesundheit und die Arbeitsleistung bei Beschäftigten zur Folge hat (Rudow 2014). Eine besonders zentrale Rolle im Stressgeschehen und der damit verbundenen gesundheitlichen Auswirkungen spielen sogenannte berufliche Gratifikationskrisen, welche durch ein dauerhaftes Ungleichgewicht von übermäßigen Arbeitsbelastungen und unzureichenden Belohnungen charakterisiert werden (Siegrist & Dragano 2008). Das auf der theoretischen Grundlage dieses Effort-Reward Imbalance Modells entwickelte Messinstrument (Siegrist 1996a, 1996b) zeigt sich seit mehr als 20 Jahren als geeignetes Verfahren, um Arbeitsbelastungen und Belohnungen sowie damit verbundene gesundheitsbezogene Parameter messbar zu machen.

In der sozialwissenschaftlichen Umfrageforschung wird daran gearbeitet, psychologische Merkmale, wie die Arbeitsbelastung, möglichst kompakt zu messen, da insbesondere individualdiagnostische Messinstrumente für Bevölkerungsumfragen aus praktischen Kosten- und Zeitbeschränkungen bei der Datenerfassung zu umfangreich sind. Für diesen Einsatz entwickelte Siegrist (1996a, 1996b) im Rahmen seines Effort-Reward Imbalance Modells aus dem ursprünglichen Fragebogen mit 22 Items, eine Kurzversion mit 16 Items, die die drei Komponente effort, reward und overcommitment enthält. Siegrist (1996a, 1996b) zeigte in seinen Arbeiten, dass die Kurzskala für die Messung der Effort-Reward Imbalance für die gesamte Bevölkerungsbreite valide und universell anwendbar ist (Siegrist et al. 2009). Damit schaffte er ein für große Surveybefragungen einsetzbares Messmodell. Nicht nur die universelle Anwendbarkeit macht das Effort-Reward Imbalance Modell so beliebt, auch die Zusammenhänge mit gesundheitsbezogenen Parametern macht es wertvoll und geeignet für die Messung von Stress (van Vegchel 2005; Solis-Soto et al. 2019; Siegrist & Dragano 2008).

Die Kurzversion des Fragebogens existiert in zwei Versionen: Zweistufig, mit Filterfrage und einstufig ohne Filter. Die ursprüngliche Messmethode wurde unter der Prämisse entwickelt, dass für das Stressempfinden die kognitive Beurteilung eine Rolle spielt (Lazarus & Folkman 1984) und wurde dadurch zweistufig konzipiert (Montano 2016). Zur Weiterentwicklung der Skala kam es, weil sich die 5er-Likert-Skala auf psychometrischer Ebene und bei der Datenerhebung als problematisch erwies (ebd.). Durch den Wegfall der Filterfrage und der Übernahme von einem weit verbreiteten einstufigen Bewertungsverfahren mittels einer 4er-Likert-Skala wurde auf die Kritik und Nachteile der ursprünglichen Methode reagiert (ebd.). Wie aus Studien anderer Themen bekannt ist, kann die Art der Fragestellung einen Einfluss auf das Ergebnis haben (Tobsch et al. 2018). Bei einer so essenziellen Veränderung der Fragestellung, wie es bei dem Effort-Reward Imbalance Modell der Fall war, gilt dies zu überprüfen.

Die Effort-Reward Imbalance Skala ist ein fester Bestandteil im SOEP und wird alle fünf Jahre befragt (2006, 2011, 2016, 2021). In den ersten drei Zeitpunkten ist die ursprüngliche Messmethode implementiert, in den beiden neusten Jahren die einstufige, neuere Messvariante der Effort-Reward Imbalance Skala. Daher laufen erstmalig und einmalig ausschließlich in der 33. Welle Bevölkerungsdaten des Sozio-oekonomischen Panels aus dem Jahr 2016 beide Messmethoden des Effort-Reward Imbalance Modells, welches kein anderes der Autorin bekannter Survey weltweit beinhaltet.

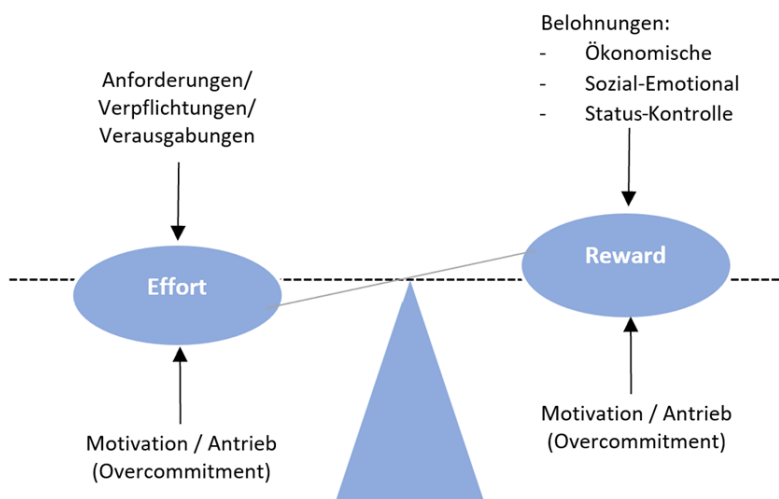
Diese besondere Datengrundlage wird dafür genutzt eine Validierung der weiterentwickelten Messmethode durch die Gegenüberstellung der Ergebnisse des zweistufigen Instrumentes vorzunehmen. Ziel des vorliegenden Beitrags ist daher die empirische Überprüfung der Vergleichbarkeit der beiden Messmethoden. Dazu werden die Reliabilität mit deskriptiven Statistiken, Inter-Item-Korrelationen und Item-Skala Statistiken sowie die Konstrukt- und Kriteriumsvalidität mit einer konfirmatorischen Faktoren- und einer multiplen Regressionsanalyse überprüft.

Im Anschluss an eine Darstellung des Forschungsstandes werden im empirischen Teil Methodik und Ergebnisse der Analysen vorgestellt. Der Beitrag schließt mit einer Diskussion der Ergebnisse und einem Ausblick.

## II Theorie und Stand der Forschung

Nach Greif (1991) beschreibt Stress einen subjektiv intensiv unangenehmen Spannungszustand, der aus der Befürchtung entsteht, dass eine stark aversive, subjektiv zeitlich nahe (oder bereits eingetretene) und subjektiv lang andauernde Situation sehr wahrscheinlich nicht vollständig kontrollierbar ist, deren Vermeidung aber subjektiv wichtig erscheint (Greif 1991). Ressourcen oder Belohnungen können dabei in puffernder Funktion eingesetzt werden, um den Stress zu reduzieren, welcher durch arbeitsbezogene Belastungen ausgelöst wird (Bamberg & Staar 2019). Ein Modell, welches sich dem (fehlenden) Gleichgewicht zwischen Belastungen und Belohnungen widmet, ist das Modell der beruflichen Gratifikationskrisen von Johannes Siegrist (1996a, 1996b). Das „Effort-Reward Imbalance Modell“ ist vielseitig getestet und international eingesetztes medizinsoziologisch-theoretisches Modell, welches auf dem Prinzip der sozialen Reziprozität basiert. Es ist unterteilt in eine situations- und die personenbezogene Ebene (Siegrist 1996a, 1996b). Die situationsbezogene Ebene umfasst Arbeitsbelastungen (effort) des Beschäftigten in Form von Zeitdruck, Arbeitsunterbrechungen oder ein größer werdendes Arbeitsaufkommen und die Gratifikationen (reward) u.a. ein angemessenes Gehalt, Anerkennung vom Vorgesetzten und ein sicherer Arbeitsplatz. Die personenbezogene Ebene wird durch das overcommitment messbar gemacht und meint eine übersteigerte Verausgabungsneigung. Zeigt eine Person in dieser Skala hohe Werte, wirkt sich dies negativ auf die situationsbezogene Ebene aus, denn Arbeitsbelastungen werden weniger stark belastend und Belohnungen werden bedeutungsreicher wahrgenommen.

**Abbildung 1: Effort-Reward Imbalance Modell**



Quelle: Eigene Darstellung in Anlehnung an Siegrist 2012

Die Tauschbeziehung in dem Modell zeigt sich in Form von Arbeitsbelastungen (effort), die vom Arbeitnehmer wahrgenommen werden und Gratifikationen (reward), die vom Arbeitgeber ermöglicht werden. Die Gratifikationen decken drei Komponenten ab: Die ökonomischen und sozial-emotionalen Belohnungen und solche, die aus der so genannten Status-Kontrolle resultieren. Wird die Tauschgerechtigkeit durch den Arbeitgeber verletzt, kommt es zu einer Effort-Reward Imbalance. Wenn dieses Ungleichgewicht von übermäßigen Arbeitsbelastungen, die nicht durch ausreichende Belohnungen kompensiert werden, dauerhaft besteht, können die Beschäftigten Distress erleben (Siegrist & Dragano 2008). Das Besondere an dem Effort-Reward Imbalance Modell ist, dass Zusammenhänge mit psychischen und physischen Gesundheitsvariablen nachgewiesen werden

konnten, und die Effort-Reward Imbalance sich als ein guter Prädiktor für negativen gesundheitlichen Auswirkungen bewährt hat (Siegrist & Li 2016).

Systematische Überblicksarbeiten und Meta-Analysen zum Effort-Reward-Imbalance Model bestätigen, dass ein Zusammenhang zwischen der Effort-Reward-Imbalance und dem Risiko eines schlechten Gesundheitszustandes existiert (u.a. van Vegchel 2005; Solis-Soto et al. 2019). Stansfeld und Candy (2006) stellten bereits vor knapp 20 Jahren fest, dass die Effort-Reward Imbalance ein prospektiver Risikofaktor für häufige psychische Störungen sein kann (Stansfeld & Candy 2006). Weitere und auch neuere Studien untermauern diese Aussage und zeigen einen Zusammenhang mit psychischen oder depressiven Störungen (Siegrist & Dragano, 2008; Solis-Soto et al. 2019; Rugulies et al. 2017). de Araújo et al. (2019) bestätigen mit der ELSA-Brasil-Studie, dass ein konsistenter Zusammenhang zwischen Effort-Reward Imbalance und depressiven Episoden besteht und Mo et al. (2020) stellen effort als einen Risikofaktor und Reward als einen Schutzfaktor für Angst und Depression fest (de Araújo et al. 2019; Mo et al. 2020). Darüber hinaus zeigen sich negative Auswirkungen einer Effort-Reward-Imbalance auf die physische Gesundheit der Arbeitnehmer. Die Ergebnisse von Kivimaki et al. (2006) weisen bei Erwerbstätigen auf ein um durchschnittlich 50% erhöhtes koronares Herzerkrankungsrisiko mit Arbeitsstress hin, welches ebenso von Dragano et al. (2017) und Penz et al. (2019) bestätigt wurden (Kivimaki et al. 2006; Dragano et al. 2017; Penz et al. 2019). Siegrist & Li (2016) fanden Zusammenhänge zwischen Arbeitsstress und Herzfrequenzvariabilität, veränderten Blutfetten und dem Risiko des metabolischen Syndroms. Eddy et al. (2016) konnten aufzeigen, dass eine höhere Effort-Reward Imbalance mit einer geringeren Immunität verbunden ist (Siegrist & Li 2016; Eddy et al. 2016). Darüber hinaus zeigen verschiedene Studien statistisch signifikante Zusammenhänge zwischen der Effort-Reward Imbalance und Schmerzen im unteren Rücken (Solis-Soto et al. 2019), mit Muskel-Skelett-Erkrankungen (Taibi et al. 2021) sowie dem erhöhten Risiko für eine Reihe weit verbreiteter Krankheiten, wie der ischämischen Herzkrankheit und dem Schlaganfallrisiko (Penz et al. 2019). Gerade wegen der Zusammenhänge der Effort-Reward Imbalance mit einer Vielzahl physischer und psychischer Gesundheitsindikatoren hat sich das Modell der beruflichen Gratifikationskrisen als ein sinnvolles Messinstrument weltweit etabliert.

Wie oben beschrieben, existiert die Kurzversion des Messinstrumentes in zwei Versionen: Zweistufig, mit Filterfrage und einstufig ohne Filter. Während die ursprüngliche zweistufige Messmethode unter der Annahme der Bedeutsamkeit kognitiver Bewertungen für das Stressgeschehen als 5-stufige Skala entwickelt wurde (Lazarus & Folkman 1984), stellte die Weiterentwicklung der Skala zu einer 4-stufigen Erhebung mit dem Wegfall der Filterfrage eine Reaktion auf psychometrische Probleme und Kritik und Nachteile der ursprünglichen Methode dar. Insofern als bereits die Art der Fragestellung einen Einfluss auf das Ergebnis haben kann (Tobsch et al. 2018), werde bei entsprechend essenziellen Veränderungen wie beim Effort-Reward Imbalance Modell Fragen hinsichtlich der vorliegenden Gütekriterien relevant.

Ziel der vorliegenden Untersuchung ist es folglich, die Vergleichbarkeit der Ergebnisse der ein- und zweistufigen Messmethode des Modells der beruflichen Gratifikationskrisen mittels der Validität und Reliabilität zu prüfen.

### III Methodik

Im Folgenden wird das methodische Vorgehen beschrieben. Neben der Erläuterung der Daten, erfolgt die Datensatzbeschreibung und die Operationalisierung, die die verwendeten Variablen darstellt.

#### 3.1 Design

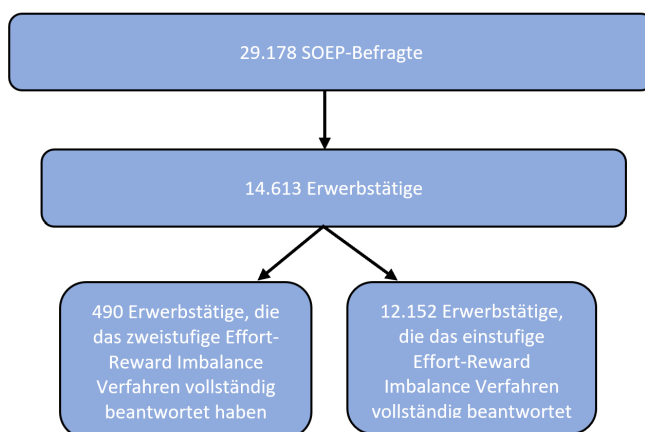
Die Datengrundlage für die nachfolgende Untersuchung ist die 33. Welle der SOEP-Vollerhebung aus dem Jahr 2016, welche eine repräsentative Bevölkerungsumfrage aller privaten Haushalte in Deutschland ist.<sup>1</sup> Mittels einer Querschnittsuntersuchung wird die Validität und Reliabilität der neueren Effort-Reward Imbalance Skala getestet. Die Besonderheit im Jahr 2016 ist, dass beide Fragebogenversionen des Modelles getestet wurden. Die neuere Version des Effort-Reward Imbalance Fragebogens ist erstmalig im SOEP und in der Hauptstichprobe und die Originalversion, die bereits 2006 und 2011 in der Hauptbefragung inkludiert war, in einer kleineren Innovationsstichprobe.

Die Reliabilität der beiden Skalen wird durch die Mittelwerte der Itemanalyse und Messung der internen Konsistenz getestet. Die faktorielle Struktur wird durch eine konfirmatorische Faktorenanalyse überprüft. Es wird getestet, ob sich das theoretische Modell, insbesondere das Konstrukt der Effort-Reward Imbalance, welches sich aus den zwei Skalen effort und reward errechnet, auch empirisch in beiden Messversionen nachweisen lässt. Anschließend wird die Kriteriumsvalidität mittels einer multiplen linearen Regressionsanalyse in Zusammenhang von demografischen, Arbeits- und gesundheitsbezogenen Variablen untersucht.

#### 3.2 Samplebeschreibung

Das Sozio-oekonomische Panel ist sowohl für die sozialwissenschaftliche Forschung als auch in der Politikberatung eine wichtige Informationsquelle. Es handelt sich hierbei um eine repräsentative Bevölkerungsumfrage von etwa 30.000 Befragten aus über 15.000 Haushalten in Deutschland, die seit 1984 jährlich befragt werden. Neben einem etablierten Set an Items, das jedes Jahr befragt wird, werden wiederholend Schwerpunktthemen gesetzt in denen unterschiedliche Messinstrumente implementiert werden.

**Abbildung 2: Stichprobenverteilung**



Quelle: Eigene Darstellung

<sup>1</sup> Version 37, SOEP, 2022, doi: 10.5684/soep.core.v37eu.



Im Auftrag des DIW Berlin wird die Befragung von Kantar Public Deutschland durchgeführt, die Teilnahme ist freiwillig. Die SOEP-Befragung findet über persönliche Interviews statt, der Personenfragebogen, der für die vorliegende Analyse verwendet wird, umfasst eine Befragungsdauer von 35-45 Minuten und wird von allen Haushaltsmitgliedern, die über 16 Jahre alt sind, beantwortet (Goebel et al. 2019; Schupp 2012). Neben Fragen über Arbeitsmarkt- und Beschäftigungsbedingungen, Bildungsmerkmale und Gesundheit wurden in den Jahren 2006, 2011, 2016 und 2021 die Effort-Reward-Imbalance Kurzsкала von Johannes Siegrist (1996a, 1996b) eingesetzt.

Insgesamt wurden 29,187 Personen im SOEP 2016 befragt, wovon 14,613 Erwerbstätige waren. Von diesen Erwerbstätigen haben 12,152 den einstufigen Effort-Reward Imbalance Fragebogen erhalten und beantwortet sowie 490 Erwerbstätige aus der Auffrischungstichprobe F (F1 & F2 ISOEP (Innovation)), die seit 2000 in die Stichprobe aufgenommen wurde, den zweistufigen Effort-Reward Imbalance Fragebogen vollständig ausgefüllt haben. Die Stichprobe beim SOEP wird über die Jahre laufend erweitert, um eine Stabilisierung der Fallzahlen zu gewährleisten und um mit dem Innovationssample innovative Forschungsprojekte Längsschnittdaten zur Verfügung zu stellen (Goebel et al. 2019; Schupp 2012).

### 3.3 Operationalisierung

In der folgenden Untersuchung steht das Effort-Reward Imbalance Model im Fokus. Die drei Siegrist-Skalen effort (berufliche Verausgabung), reward (Belohnungen) und overcommitment (übersteigerte berufliche Verausgabebereitschaft) wurden in der SOEP-Vollerhebung 2016 in den Fragebogen integriert. Sie werden neben den deskriptiven Analysen auch für die multivariaten Verfahren verwendet.

#### **Zweistufiges Verfahren der Effort-Reward Imbalance Skala**

Die ursprüngliche Kurzversion des Effort-Reward Imbalance Skala startet in der 33. Welle des SOEPs mit der Frage: *Ich lese Ihnen Aussagen über mögliche berufliche Belastungen an Ihrem derzeitigen Arbeitsplatz vor. Bitte geben Sie an, ob der jeweilige Punkt bei Ihnen zutrifft und, falls ja, wie stark Sie das belastet.*

Die Skala effort wird mit den drei Items „Aufgrund des hohen Arbeitsaufkommens besteht häufig großer Zeitdruck“, „Bei meiner Arbeit werde ich häufig unterbrochen und gestört“ und „Im Laufe der letzten beiden Jahre ist meine Arbeit immer mehr geworden“ gemessen. Auf das jeweilige Item wird die Frage gestellt, ob diese zutreffend sei und sofern dieses bejaht wird, wird um eine Einschätzung bezüglich der Belastung auf einer Skala von „belastet mich gar nicht“ bis „belastet mich sehr stark“ gebeten. Beide Fragebogenversionen befinden sich im Anhang.

Auch bei den sieben Items der Reward Skala wird mit derselben Filterfrage gearbeitet, zumindest bei den ersten drei Items, da diese negativ formuliert sind. Die weiteren vier Reward Items sind positiv formuliert, und es wird somit nur um eine Einschätzung der Belastung gebeten, wenn diese verneint werden. Die Reward Skala enthält die drei Subskalen esteem, die die zwei Items „Ich erhalte von meinen Vorgesetzten die Anerkennung, die ich verdiene“ und „Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich die erfahrene Anerkennung für angemessen“ enthält. Die Subskala job security wird mit den zwei Items „Ich erfahre – oder erwarte – eine Verschlechterung meiner Arbeitssituation“ und „Mein eigener Arbeitsplatz ist gefährdet“ gemessen und die dritte Subskala ist job protection, welche mit den Items „Die Aufstiegschancen in meinem Betrieb sind

**Tabelle 1 Zweistufiger Effort-Reward Imbalance Fragebogen im SOEP 2016**

		Zutreffend		Und wie stark belastet Sie das?			
		Nein	Ja	Gar nicht	mäßig	stark	Sehr stark
		(1)		(2)	(3)	(4)	(5)
E1	Aufgrund des hohen Arbeitsaufkommens besteht häufig Zeitdruck	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2	Bei meiner Arbeit werde ich häufig unterbrochen und gestört	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E3	Im Laufe der letzten beiden Jahre ist meine Arbeit immer mehr geworden	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
R1	Meine Aufstiegschancen in meinem Betrieb sind schlecht <i>Rekodierung</i>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
R2	Ich erfahre – oder erwarte – eine Verschlechterung meiner Arbeitssituation <i>Rekodierung</i>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
R3	Mein eigener Arbeitsplatz ist gefährdet <i>Rekodierung</i>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			(5)	(4)	(3)	(2)	(1)
R4	Ich erhalte von meinen Vorgesetzten die Anerkennung, die ich verdiene		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
R5	Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich die erfahrende Anerkennung für angemessen		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
R6	Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich meine persönlichen Chancen des beruflichen Fortkommens für angemessen		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
R7	Wenn ich an all die erbrachten Leistungen denke, halte ich mein Gehalt/meinen Lohn für angemessen		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quelle: Eigene Darstellung

schlecht“, „Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich meine persönlichen Chancen des beruflichen Fortkommens für angemessen“ und „Wenn ich an all die erbrachten Leistungen denke, halte ich mein Gehalt / meinen Lohn für angemessen“ erfragt wird.

### **Einstufiges Verfahren (I)**

Die neue Kurzversion der Effort-Reward Imbalance Skala beinhaltet dieselben zehn Items, wie die ursprüngliche Version. Dementsprechend müssen auch hier die ersten drei Reward Items rekodiert werden. Was wegfällt ist die Filterfrage bei der neuen Version. Es wird ebenfalls auf einer 4er-Skala gemessen allerdings wird nicht mehr nach der Einschätzung der vorhandenen Belastung von „gar nicht“ bis „sehr stark“ gefragt, sondern auf die Frage:

*Die folgenden Aussagen beschreiben mögliche Situationen im Zusammenhang mit Ihrer Arbeit, Ihrem Arbeitsplatz und dem Betrieb, in dem Sie arbeiten. Bitte geben Sie an, in welchem Maße Sie den folgenden Aussagen zustimmen.*

werden die Befragten gebeten der Aussage zuzustimmen auf einer Skala von „stimme gar nicht zu“ bis „stimme voll zu“. Damit stimmen die Skalenausprägungen mit der des overcommitments überein,

welches von Beginn der Konzeption an auf einer 4er-Likert-Skala gemessen wurde. Hier gewinnt das Messinstrument an Vereinheitlichung. Der große Unterschied zu der Version zuvor ist, dass die Befragten nicht mit dem Wort „Belastung“ in Berührung kommen, wodurch zumindest die Möglichkeit einer verzerrten, weil geprimten, Einschätzung besteht. Es ist in der neuen Version eine neutrale Formulierung gewählt wurden, welche keinen negativen Bias in der Fragestellung verankert. Durch den Wegfall der Filterfrage, könnte darüber hinaus die Zahl der missings reduziert werden (Masouel et al. 2012).

### Overcommitment

Neben den beiden Hauptskalen effort und reward gehört zu dem Modell der Gratifikationskrisen auch die Skala overcommitment (übersteigerte Verausgabungsneigung), welche in einem positiven Zusammenhang mit dem Effort-Reward Imbalance Wert steht. Es wirkt bei hoher Ausprägung verstärkend auf die Effort-Reward Imbalance, da ein Erwerbstätiger mit übersteigelter Verausgabungsneigung die Belastungen subjektiv geringer wahrnimmt und den Belohnungen ein höheres Gewicht beimisst. Bei der Messung des overcommitments gab es keine Änderungen.

**Tabelle 2 Overcommitment Fragebogen im SOEP 2016**

		Stimme gar nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme voll zu
		(1)	(2)	(3)	(4)
OC1	Beim Arbeiten komme ich leicht in Zeitdruck	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OC2	Es passiert mir oft, dass ich schon beim Aufwachen an Arbeitsprobleme denke	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OC3	Wenn ich nach Hause komme, fällt mir das Abschalten von der Arbeit sehr leicht <i>Rekodierung</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OC4	Diejenigen, die mir am nächsten stehen sagen, dass ich mich für meinen Beruf zu sehr aufopfere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OC5	Die Arbeit lässt mich selten los, das geht mir abends im Kopf rum	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OC6	Wenn ich etwas verschiebe, was ich eigentlich heute tun müsste, kann ich nachts nicht schlafen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quelle: Eigene Darstellung

Die Skala wird durch die sechs Items (siehe Tabelle 2) gebildet. Die Befragten antworten auf die Aussagen mit einer Einschätzung auf einer vierer Likert-Skala von „Stimme gar nicht zu“ bis „Stimme voll zu“. Ein hoher Wert steht für eine hohe übersteigerte Verausgabungsneigung. Das dritte Item „Wenn ich nach Hause komme, fällt mir das Abschalten von der Arbeit sehr leicht“ ist negativ für die Skala overcommitment formuliert ist und wird deshalb für die Analysen rekodiert.

## Berechnung des Effort-Reward Imbalance Scores

Die Effort-Reward Imbalance zeigt die Stärke des Ungleichgewichts auf individueller Ebene zwischen effort und reward. Es kann mittels der Formel die ER-Ratio:

$$ER = k \frac{E}{R} \quad (1.1)$$

ein Indikator für jeden einzelnen Befragten errechnet werden, der zeigt, wie es um das individuelle Ungleichgewicht steht. Dabei sind E die aufsummierten Effort- und R die aufsummierten Rewardscores. Sofern eine unterschiedliche Anzahl von Items vorhanden ist, muss ein Korrekturfaktor k eingefügt werden, dieses ist hier nicht der Fall, wird allerdings der Vollständigkeit halber erläutert. Der Schwellen-Wert, also der Wert, bei dem es zu einer Imbalance kommt, indem die Arbeitsbelastungen die Belohnungen übersteigen, liegt bei einer ER-Ratio von  $> 1$ , übersteigen die Belohnungen die Belastungen fällt die ER-Ratio unter 1 und eine ER-Ratio = 1 bedeutet, dass die Befragte Person von gleich viel vorhandenen Belohnungen und Belastungen spricht (Siegrist 2014).

Diese Formel ist für beide Messverfahren, also sowohl für das neuere einstufige als auch für das zweistufige Verfahren gültig, allerdings ist die Streuung der neueren Version durch den niedrigeren Maximalwert geringer, wie in der Tabelle 3 dargestellt ist. Die zweistufige Version wird auf einer 5er-Likert-Skala gemessen und hat bei effort einen Minimalwert von 3 und Maximalwert von 15, für reward beträgt die Spanne 7-35 und für die Effort-Reward Imbalance gilt ein Wertebereich von 0.2 – 5. Die neuere Messmethode zeigt eine geringere Spannweite durch die Messung auf einer 4er-Likert-Skala auf, deshalb hat effort eine Spanne von 3-12, reward befindet sich im Wertebereich von 7-28 und die Effort-Reward Imbalance hat den Minimalwert 0.25 und den Maximalwert von 4.

**Tabelle 3: Berechnung der Summenscores des einstufigen (I) und des zweistufigen (II) Effort-Reward Imbalance Messinstrumentes**

Skala	Items	Einstufiges Verfahren (I)	Zweistufiges Verfahren (II)
		Range	Range
effort	E1 – E3	3 – 12	3 – 15
reward	R1 – R7	7 – 28	7 – 35
overcommitment	OC1 – OC6	6 – 24	6 – 24
Effort-Reward Imbalance	ERI	0.25 – 4	0.2 – 5

Quelle: Eigene Darstellung

## Transformation der Skalen

Die Effort-Reward Imbalance Skala existiert in zwei allgemeinen Formaten: Lang- (22 Items) und Kurzformat (16 Items) und kann in einem ein- oder zweistufigen Verfahren gemessen werden. Durch die unterschiedlichen Messungen der Effort-Reward Imbalance Skala durch ihre unterschiedliche Spannweite (range), sind die Ergebnisse nicht direkt miteinander vergleichbar. Deshalb werden für die folgende Analyse die Originalwerte in eine gemeinsame Skala mittels Anwendung einer linearen Transformation umgewandelt, die die angepassten Originalwerte in einen Bereich von 1-100 abbilden (Montano 2016; Siegrist 2014). Der transformierte Wert ( $y_i$ ) für eine Person (i) wird wie folgt definiert:

$$y_i = (x_i * 100 - 1) / (m_l - m_u) + 1 \quad (1.2)$$

Dabei steht  $x_i$  für einen beliebigen Originalwert für die Person  $i$  (effort, reward, overcommitment, Effort-Reward Imbalance) und  $m_l > 0$  für das Maximum und  $m_u$  für das Minimum des Originalrange. Mittels dieser Transformation lassen sich die Ergebnisse der Effort-Reward Imbalance aus der nachfolgenden Analyse des Sozio-oekonomischen Panels (2016), die ein- und zweistufig gemessen wurden, vergleichen, und es können Aussagen darüber getroffen werden, ob es Unterschiede in den Ergebnissen je nach Messung gibt. Um vorher zu prüfen, ob die beiden Samples im SOEP repräsentativ sind, folgt im nächsten Abschnitt eine Repräsentativitätsprüfung und im Weiteren die deskriptiven und multivariaten Analysen.

## IV Ergebnisse

Die Datenanalyse wurde mit SPSS 29 durchgeführt mit Ausnahme der konfirmatorischen Faktorenanalyse, diese wurde mit dem Statistikprogramm R durchgeführt.

### 4.1 Repräsentativitätsprüfung

Das Sozio-oekonomische Panel ist eine repräsentative Umfrage für Deutschland. Es wird geprüft, ob beide Datensätze sich hinsichtlich verschiedener Merkmale ähnlich verteilen, die zu einer Repräsentativitätsprüfung herangezogen werden können. Die Hauptbefragung (I) beinhaltet insgesamt 12,152 Fälle, welche gültige Antworten abgegeben haben und den einstufigen Effort-Reward Imbalance Fragebogen ohne Filterfrage erhalten haben. Die Innovationsstichprobe (II) hat den zweistufigen Fragebogen erhalten und wurde von 490 Personen ausgefüllt.

**Tabelle 4: Repräsentativitätsprüfung (Angaben in Prozent), Sample I (N=12,152) und Sample II (N=490)**

<b>Merkmal</b>	<b>SOEP 2016 I</b>		<b>SOEP 2016 II</b>	
<b>Geschlecht</b>				
weiblich	6224	51.2	253	51.6
männlich	5928	48.8	237	48.4
<b>Alter</b>				
18 bis unter 25	548	4.5	13	2.7
25 bis unter 35	1671	13.8	66	13.5
35 bis unter 45	2953	24.3	79	16.1
45 bis unter 55	3726	30.7	132	26.9
55 bis unter 65	2613	21.5	143	29.2
65 und älter	641	5.3	57	11.6
<b>Familienstand</b>				
Verheiratet	7565	62.2	265	54.1
Ledig	3260	26	149	30.4
Geschieden	1219	10	64	13.1
Verwitwet	175	1.4	11	2.2
Eing. gleichg. Partnerschaft	29	0.2	1	0.2
keine Angabe	4	0	0	0
<b>Erwerbsstatus</b>				
Voll erwerbstätig	7549	62.1	310	63.3
Teilzeitbeschäftigung	3091	25.4	111	22.7
Ausbildung, Lehre	494	4.1	15	3.1
Geringfügig beschäftigt	998	8.2	52	10.6
Werkstatt für behinderte Menschen	20	0.2	2	0.4
<b>Berufliche Stellung</b>				
Selbstständige	614	5.1	19	3.9
Arbeiter	2694	22.2	98	20
Beamte	806	6.6	27	5.5
Auszubildende	531	4.4	16	3.3
Angestellte	7484	61.6	329	67.1
keine Angabe	23	0.2	1	0.2
<b>Öffentlicher Dienst</b>				
Öffentlicher Dienst	2918	24	110	22.4

Privatwirtschaft	8516	70.1	361	73.7
keine Angabe/trifft nicht zu	718	6	19	3.9

Quelle: Eigene Darstellung

Als Vergleichsmerkmale wurden demographische Variablen, wie das Geschlecht, Alter und Familienstand geprüft sowie die Arbeitsvariablen: Erwerbsstatus, Berufliche Stellung und Öffentlicher Dienst/Privatwirtschaft (siehe Tabelle 4). Die prozentuale Verteilung beim Geschlecht, Erwerbsstatus und öffentlicher Dienst ist nahezu identisch und bei der beruflichen Stellung gibt es nur minimale Abweichungen der beiden Stichproben. Kleinere, aber zu vernachlässigende Abweichungen, sind bei den Variablen Alter und Familienstand zu verzeichnen. Etwas mehr Erwerbstätige sind in der größeren Stichprobe (I) mit der einstufigen Messmethode verheiratet und mehr Erwerbstätige sind in den Alterskategorien 35 bis unter 45 und 45 bis unter 55 zu finden. Die Gegenüberstellung zeigt, dass die Datensätze sich hinsichtlich der Verteilung gleichen und weitere Analysen durchgeführt werden können.

#### 4.2 Item- und Reliabilitätsanalyse

Um eine Itemanalyse der Effort-Reward Imbalance Skala durchzuführen, werden der Mittelwert und die Standardabweichung berechnet sowie zur Prüfung der internen Konsistenz Cronbachs Alpha Koeffizient und Trennschärfe geprüft.

**Table 5: Transformierte Mittelwerte (M), Standardabweichung (SD), Cronbach's Alpha und Trennschärfe (Item-Total Korrelationskoeffizienten) der Effort-Reward Imbalance Skala, Sample I (N=12,152) und Sample II (N=490)**

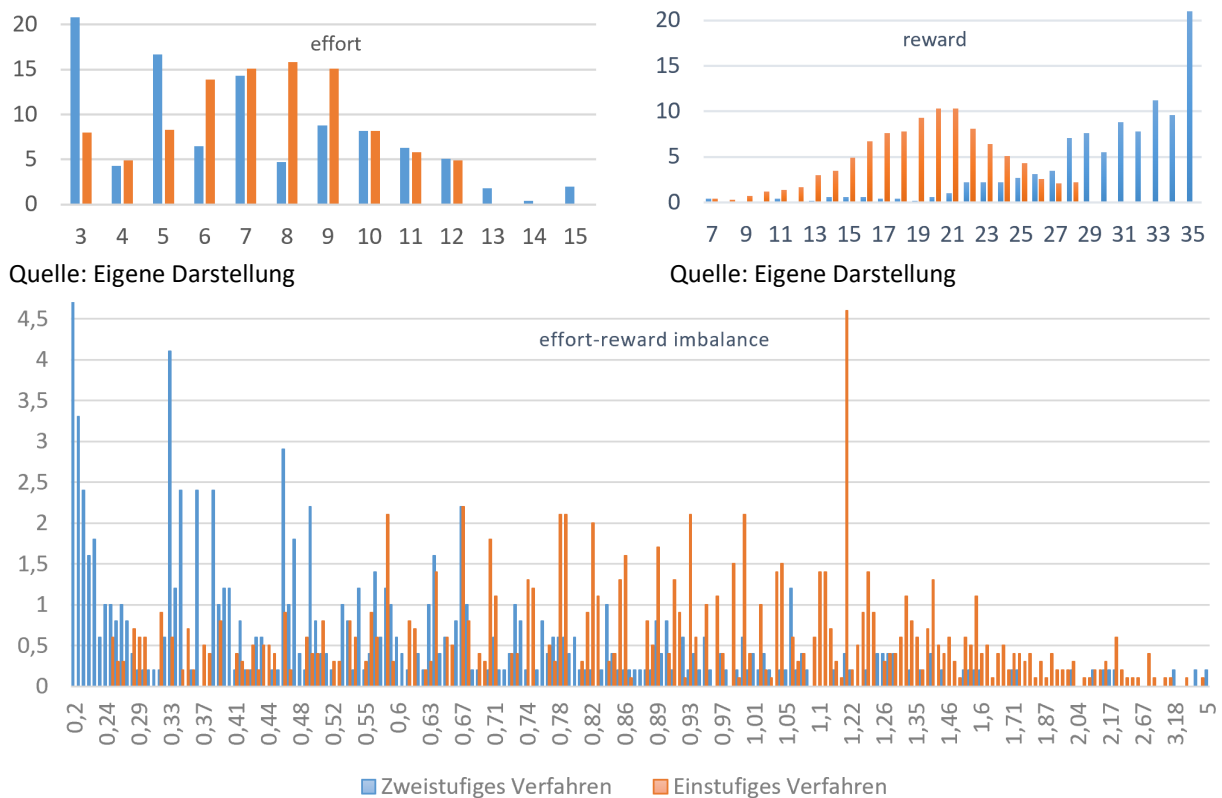
	Alpha			M		SD		Effort		esteem		Job security		Job promotion		Overcommitment		
	I	II		I	II	I	II	I	II	I	II	I	II	I	II	I	II	
Effort	0.7	0.76	E1	47.27	32.62	31.91	31.96	0.81	0.83									
			E2	46.64	33.66	31.96	31.93	0.78	0.80									
			E3	54.65	33.32	36.05	31.68	0.78	0.83									
Reward	0.77	0.8	R4	58.14	30.91	83.58	26.38			0.94	0.94							
			R5	57.12	30.78	83.99	25.67			0.94	0.94							
			R3	83.04	24.32	93.23	21.63					0.90	0.91					
Job security			R2	72.24	29.26	89.39	25.47					0.85	0.88					
			R3	83.04	24.32	93.23	21.63											
			R1	43.62	34.17	74.80	26.08							0.74	0.77			
Job promotion			R6	52.74	29.80	82.17	26.26							0.78	0.82			
			R7	49.56	30.68	74.19	30.53							0.72	0.76			
			OC1	2.39	0.92	2.38	0.92										0.60	0.58
Overcommitment	0.77	0.77	OC2	2.01	0.95	2.00	0.96									0.78	0.77	
			OC3	2.17	0.98	2.10	1.01									0.61	0.61	
			OC4	2.12	0.97	2.08	1.00									0.67	0.70	
			OC5	2.00	0.92	1.98	0.94									0.81	0.81	
			OC6	1.81	0.87	1.73	0.86									0.62	0.61	

Quelle: Eigene Darstellung

Tabelle 5 zeigt die transformierten Mittelwerte, die zwischen 1 und 100 liegen können und zwischen einer maximalen Punktzahl von 83.04 (I) und 93.23 (II) beim dritten reward Item und einem Minimum von 27.75 (I) und 25.24 (II) beim sechsten overcommitment Item variieren. An der Tabelle 5 ist zu sehen, dass die Mittelwerte in den zwei Samples augenscheinlich in der Höhe deutlich voneinander abweichen. Dieses verstärkt sich in den aussummierten Skalen effort und reward und auch bei der ER Ratio, die aus den beiden Skalen errechnet wird. Die Rangfolge der Items, wie an der Höhe der Mittelwerte stimmen in beiden Samples weitestgehend überein (I – R3, R2, R4, R5, R6, R7, R1 / E3, E1, E2 & II – R3, R2, R5, R4, R6, R1, R7 / E3, E2, E1). Lediglich die Positionen des vierten und fünften reward Items und dem zweiten und ersten effort Item sind vertauscht. Die interne Konsistenz der Effort-Reward Imbalance Skala wurde mittels Cronbachs Alpha und Item-Total Korrelationen geprüft, welche ebenfalls in Tabelle 5 zu finden sind. Die Analyse zeigt, dass für beide Samples die Cronbachs Alpha Werte mit: 0.70/0.76 (I/II) für effort, 0.77/0.80 (I/II) für reward und 0.77/0.77 (I/II) für

overcommitment für beide Samples ausreichend (Bühner & Ziegler 2017) sind. Die Koeffizienten für die korrigierten Item-Total Korrelationen sind für alle Items zwischen 0.60 (I) bzw. 0.58 (II) und 0.95 (I/II) und zeigen damit ebenfalls akzeptable Werte für beide Samples und weisen auf eine gute Konsistenz der Items hin, die zu den jeweiligen Skalen zusammengefasst werden.

**Abbildung 3: Prozentuale Häufigkeitsverteilungen der Mittelwerte von effort, reward und Effort-Reward Imbalance des zweistufigen und einstufigen Verfahrens des Effort-Reward Imbalance Model, Sample I (N= 12,152) und Sample II (N=490)**



Quelle: Eigene Darstellung

Um eine Erklärung zu finden, wie es zu den Unterschieden in der Höhe der Mittelwerte kommen kann, wird das Antwortverhalten genauer betrachtet. Die Abbildung 3 zeigt die prozentualen Häufigkeitsverteilungen der Originalwerte beider Messmethoden, die jeweils in den Diagrammen zur besseren Vergleichbarkeit übereinandergelegt werden. Auf der horizontalen Achse sind die aufaddierten Werte der effort- und reward-Skalen sowie der Effort-Reward Imbalance Skala abgebildet und auf der vertikalen Achse die prozentualen Häufigkeiten.

Beim Blick auf die Häufigkeitsverteilung der Skalen effort und reward ist zu erkennen, dass bei der ursprünglichen Messmethode der überwiegende Anteil der Antworten auf die erste Antwortkategorie (bei effort: 1, bei reward: 5) fällt, zu sehen an den drei aufsummierten effort Items die einen Wert von 3 annehmen und bei den aufsummierten sieben reward Items ein Wert von 35 herauskommt. Dieses führt bei effort zu einer rechtsschiefen und reward einer linksschiefen Verteilung. Die neuere Messmethode zeigt hingegen die häufigsten Antworten in den beiden mittleren Antwortkategorien (Kategorie 2 und 3), weshalb sich hier die Skalen effort und reward tendenziell normalverteilt zeigen. Dadurch dass die Effort-Reward Imbalance aus den beiden Skalen errechnet wird, wirken sich die



schiefen Verteilungen auch auf die Effort-Reward Imbalance aus und zeigen bei der ursprünglichen Messmethode wie bei effort eine rechtsschiefe Verteilung und bei der neueren Messmethode eine weitestgehende Normalverteilung.

### 4.3 Konstruktvalidität

Die Konstruktvalidität bzw. Faktorielle Validität wurde mit einer konfirmatorischen Faktorenanalyse für beide Samples gerechnet und mit einem Vergleichswert aus dem SOEP 2006 abgeglichen unter Anwendung des Statistikprogramms R.

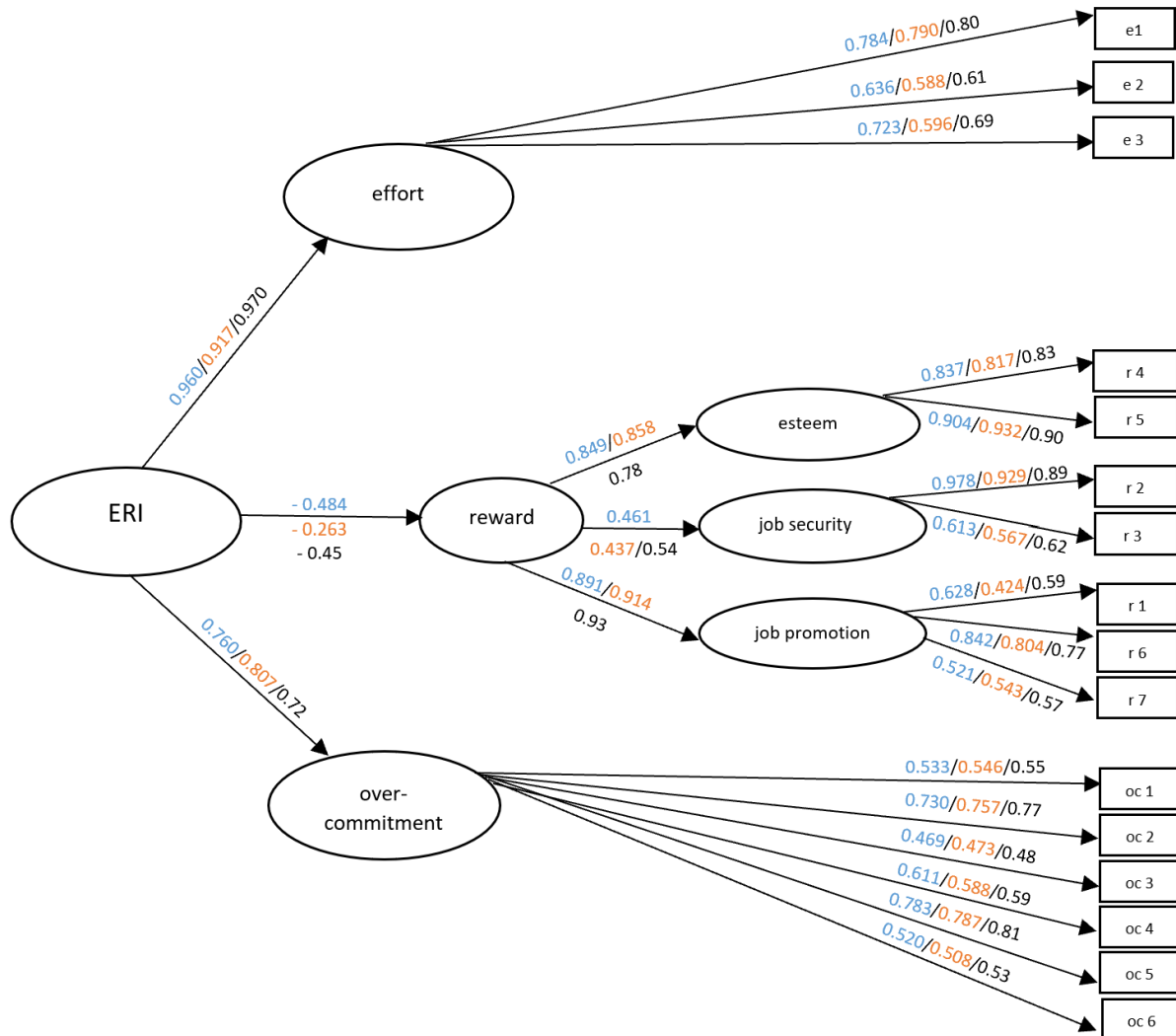
**Tabelle 6 Zusammenfassung der Fit-Indizes für die 3 Modelle der konfirmatorischen Faktorenanalyse, Sample I (N=12,152) und Sample II (N=490)**

Fit-indices	Model 0		Model 1		Model 2	
	I	II	I	II	I	II
$\chi^2$	1262.274	31232.965	580.663	11243.345	321.813	6954.172
$\chi^2/df$	104	104	101	101	98	98
RMSEA	0.151	0.157	0.099	0.095	0.068	0.076
CI: 90% low – 90% up)	0.144 – 0.0159	0.156 – 0.159	0.091 – 0.107	0.094 – 0.097	0.060 – 0.077	0.075 – 0.078
CFI	0.558	0.492	0.817	0.818	0.915	0.888

Quelle: Eigene Darstellung

Es wurden insgesamt drei Modelle geschätzt. Das Model 0 geht davon aus, dass alle 16 Items auf eine zugrundeliegende Dimension laden. Die Kennzahlen bei diesem Modell weisen, eine schlechte globale Anpassung auf (Tabelle 6). Die Modelle 1 und 2 prüfen die theoretische Struktur des Effort-Reward Imbalance Modell, das erstere Modell schätzt die drei Faktoren effort, reward und overcommiment als Faktoren erster Ordnung und die Effort-Reward Imbalance als Faktor zweiter Ordnung. Die Statistiken zeigen hier eine bessere Anpassung als im Model 0 allerdings immer noch keine akzeptablen Werte bei dem Model 1 (Tabelle 6). Das Model 2 prüft darüber hinaus zusätzlich die Subskalen der Reward Skalen: esteem, job security und job promotion. Das Modell 2, welches die theoretische Struktur der Effort-Reward Imbalance Modell mit den Haupt- und Subskalen darstellt, zeigt einen guten Fit, wie die Tabelle 6 zeigt (Hu & Bentler 1999).

**Abbildung 4 Konfirmatorische Faktorenanalysen zur Überprüfung des theoretischen Konstrukts des ein- und zweistufigen Verfahrens (SOEP Daten 2016) sowie Vergleichswerte von Siegrist et al. 2009 (SOEP Daten 2006)**



■ Zweistufiges Verfahren (SOEP 2016) ■ Einstufiges Verfahren (SOEP 2016) ■ Zweistufiges Verfahren (SOEP 2006)

Quelle: Eigene Darstellung in Anlehnung an Siegrist et al. 2009

Zu erwarten war, je detailreicher und näher an dem theoretischen Modell von Siegrist (1996) gerechnet wird, desto besser die Passung. Wie in Tabelle 6 gezeigt, verbesserte sich die Passform dieses Modells im Vergleich zu den früheren Modellen erheblich. Auch in dem Strukturbaum in der Abbildung 4 laden die Items in einer ähnlichen Höhe auf die übergeordneten Konstrukte, dies gilt sowohl die beiden Samples als auch für die Vergleichswerte von Siegrist aus dem Sozio-oekonomischen Panel aus dem Jahr 2006. Zusammenfassend lässt sich feststellen, dass die theoretische Kernstruktur des Effort-Reward Imbalance Modell angemessen in beiden Samples widerspiegelt wird. Die Kennzahlen sind mit der einstufigen etwas besser als mit der zweistufigen Messmethode.

#### 4.4 Kriteriumsvalidität

Zur Überprüfung der Kriteriumsvalidität der Effort-Reward Imbalance Skala wurden in beiden Samples als unabhängige Variablen Merkmale aus dem Bereich der Demographie, Arbeit, Gesundheit und Zufriedenheit herangezogen und eine multiple Regression gerechnet. Die Variablen wurde so

ausgewählt, dass sie in vergangenen Studien signifikante Ergebnisse im Zusammenhang mit der Effort-Reward Imbalance (abhängige Variable) geliefert haben, wenn gleich kein hohes  $R^2$  zu erwarten ist, da die Auswahl an Variablen, die einen starken Zusammenhang in der Regel mit der abhängigen Variable zeigen, nicht im SOEP-Repertoire enthalten sind. Für die Validierung der Skala fällt dieses jedoch nicht ins Gewicht, weil es ausreicht zu prüfen, wie stark und signifikant die unabhängigen Variablen wirken und ob die Höhe und Signifikanz unabhängig von der Messmethode ist. Wir sehen in der Tabelle 7 die standardisierten Beta Koeffizienten, die uns zeigen, dass hohe signifikante Werte in beiden Samples bei den relevanten und erwarteten Variablen zu finden sind. Die Wochenstunden, das overcommitment und die Arbeitszufriedenheit haben in beiden Samples hoch signifikante standardisierte Beta Koeffizienten zu verzeichnen. Die Wochenstunden und die Arbeitszufriedenheit wirken negativ auf die Effort-Reward Imbalance, das overcommitment verstärkt hingegen, wie die Theorie es auch sagt, die Effort-Reward Imbalance.

**Tabelle 7 Lineare Regressionsanalyse der Effort-Reward-Imbalance, Sample I (N=12,152) und Sample II (N=490)**

	I		II	
	st. Beta	Sig.	st. Beta	Sig.
<b>Demographische Variablen</b>				
Geschlecht (männlich)	-0.022	0.162	-0.066	0.158
Verheiratet (nicht verheiratet)	-0.027	0.071	0.074	0.085
Alter	0.011	0.574	-0.059	0.229
<b>Arbeitsvariablen</b>				
Betriebszugehörigkeit	-0.084	0.000	-0.001	0.990
Wochenstunden	0.184	0.000	0.104	0.036
Teilzeit	0.061	0.000	-0.053	0.255
Unbefristet	0.078	0.000	0.071	0.098
overcommitment	0.323	0.000	0.386	0.000
<b>Gesundheitsvariablen</b>				
Gesundheitszustand	-0.036	0.059	0.014	0.787
Sorge um die Gesundheit	-0.034	0.049	-0.035	0.446
<b>Zufriedenheitsvariablen</b>				
AZ	-0.365	0.000	-0.196	0.000
LZ	-0.023	0.225	-0.150	0.001
Gesundheitszufriedenheit	0.006	0.805	-0.061	0.269
Schlafzufriedenheit	0.062	0.001	0.056	0.235
<b>R<sup>2</sup> adj.</b>	<b>0.372</b>		<b>0.333</b>	

Quelle: Eigene Darstellung

Diese Ergebnisse stimmen mit den Befunden anderer Studien überein (z.B. Fietze 2011; Calnan et al. 2008; Li et al. 2005; Kinman & Jones 2008). In dem zweistufigen Messverfahren ist darüber hinaus noch die Lebenszufriedenheit mit einem signifikanten negativen standardisierten Beta-Koeffizienten von -0.150 zu sehen. Das korrigierte  $R^2$  ist in beiden Samples in einer vergleichbaren Höhe, einstufig:  $R^2$ : 0.372 und zweistufig:  $R^2$ : 0.333. Diese Analyse lässt den Schluss zu, dass die Kriteriumsvalidität gut ist und sich die beiden Messverfahren in den Zusammenhängen mit den relevanten Variablen nicht in Höhe oder Richtung der standardisierten Beta-Koeffizienten groß unterscheiden.

## V Diskussion

Ziel der vorliegenden Untersuchung war es, die Vergleichbarkeit der Ergebnisse der ein- und zweistufigen Messmethode des Modells der beruflichen Gratifikationskrisen mittels der Reliabilität und Validität zu prüfen. Zusammenfassend zeigen die Analysen der deskriptiven Statistiken, dass die Mittelwerte der einzelnen Items und dadurch auch für die aufaddierten Skalen effort, reward und der Effort-Reward Imbalance der beiden Messverfahren trotz Transformation zumindest in der vorliegenden SOEP-Erhebung der 33. Welle nicht vergleichbar sind, und diese sogar stark voneinander abweichen. Ursache hierfür ist ein unterschiedliches Antwortverhalten bei den beiden Messinstrumenten. In der ursprünglichen Messmethode, vermutlich bedingt durch die vorangegangene Filterfrage, gibt es eine Antworttendenz zu den unteren äußeren Ausprägungen. Es kommt so zu einer rechtsschiefen Verteilung bei effort und einer linksschiefen Verteilung bei reward. Tsusumi (2008) stellte eine ähnliche Verteilung der Antworten bei der zweistufigen Messmethode fest. Auch bei ihnen fiel bis auf wenige Ausnahmen fast die Hälfte oder mehr als die Hälfte aller Antworten auf die erste Antwortkategorie (Tsusumi 2008). Die einstufige Messmethode zeigt normalverteilte Antworten, weil die Befragten am häufigsten die mittleren Kategorien ankreuzen. Dadurch kommt es zu deutlich höheren Werten bei den Belastungen und niedrigeren Werten bei der Bewertung der Gratifikationen. Da die Effort-Reward Imbalance aus diesen beiden Konstrukten errechnet wird, kommt es auch zu insgesamt höheren Effort-Reward Imbalance Werten, die den neuen Fragebogen erhalten haben und die Befragten zeigen höhere Belastungswerte auf.

Die Art der Messung und die Formulierung der Items, hat einen Einfluss auf das Antwortverhalten der Teilnehmer und führt so zu unterschiedlichen Ergebnissen in der Verteilung und Mittelwerten. Die Rangfolge der Items stimmt allerdings weitgehend bei den beiden Messmethoden überein.

Die interne Konsistenz mit Cronbachs Alpha und den Item-Total-Korrelationen ist bei beiden Messmethoden gut, ebenso wie sich die theoretische Struktur mit beiden Messverfahren in den Ergebnissen der konfirmatorischen Faktorenanalyse als gut erwiesen hat. Somit können wir für die Konstruktvalidität ohne Einschränkungen die Aussage tätigen, dass beide Messmethoden das Modell der beruflichen Gratifikationskrisen in der theoretischen Struktur hervorragend mit den Daten des SOEPs widerspiegeln, die einstufige Messmethode zeigt sogar einen noch besseren Fit der Daten. Auch die Anzahl der Missings ist tatsächlich geringer in der neueren Messmethode mit Ausnahme der dritten effort und reward Items (siehe Tabelle 8 im Anhang). Daher war die Weiterentwicklung des Effort-Reward Imbalance Messinstrumentes nicht nur hinsichtlich der Konstruktvalidität richtig und wichtig, sondern ging auch mit einer Simplifizierung der Codierung und Berechnung einher und die Befragungszeit hat sich verringert, was insbesondere für die Surveys relevant ist.

Die vorliegende Studie weist einige zu diskutierende Limitationen auf: Zum einen wurden lediglich Daten aus dem SOEP-Survey als Datengrundlage herangezogen. Damit beschränkt sich die Aussagekraft auf die deutschsprachige Version und eine auf Deutschland begrenzte Teilnehmerschaft, welche zudem im Querschnitt lediglich zu einem Zeitpunkt befragt wurde. Zum anderen zeigt die Regressionsanalyse zwar, dass die gleichen unabhängigen Variablen recht hohe und signifikante Effekte in beiden Samples haben. Allerdings wären hier mehr Variablen für aussagekräftigere Ergebnisse wünschenswert gewesen, daher kann hier nur mit etwas Vorsicht die Kriteriumsvalidität beider Messmethoden bestätigt werden. Die Hinzunahme relevanter Gesundheitsvariablen, die im Zusammenhang mit einer Effort-Reward Imbalance stehen, würden verlässlichere Aussagen in Bezug auf die Kriteriumsvalidität zulassen.

Weitere Ansätze für die zukünftige Beforschung beziehen sich auf die Erhebung des komplexen Konstrukts Gesundheit mittels subjektiver und objektiver Messungen: Gesundheit ist generell aber auch im SOEP ein großes Thema, weshalb es fest im SOEP Standard verankert wurde. Auch hat das SOEP in den letzten Jahren immer mehr verhaltenswissenschaftliche Skalen zusätzlich in ihr Variablenset aufgenommen. Was noch fehlt, sind physische Stressmessungen im SOEP. Eine gut handhabbare physische Stressmessung für einen Survey ist die Überprüfung des Cortisolspiegels bei den Befragten. Bisherige Übersichtsarbeiten, wie Siegrist und Li (2016), zeigen widersprüchliche Ergebnisse bei der Untersuchung vom Effort-Reward Imbalance Modell und dem Stresshormon Cortisol. Dieses lässt sich teilweise auf die Messgrößen und unterschiedliche Designs zurückführen. Daher wäre es umso wichtiger, dieses mittels Survey zu standardisieren und wiederholt in einem repräsentativen Datensatz laufen zu lassen. Nicht nur die Standardisierung kann durch ein erfahrenes Umfrageinstitut gut umgesetzt werden, sondern überhaupt die Kombination von subjektiver Messung und der Messung physiologischer Parameter von Stress möglich gemacht werden.

Stressmessung in Surveys sind wichtig und nicht nur die Organisationsforschung benötigt solche Messungen wegen der Vergleichbarkeit und Einordnung der eigenen Daten und Ergebnisse der Forschung, weshalb die Eichung der Skala so wichtig ist. Das vorliegende Paper zeigt, dass die Mittelwerte sich, in dem vorliegenden SOEP Survey der 33. Welle nicht vergleichbar sind und bestätigen wie andere Studien zuvor die unterschiedliche Verteilung der Antworten je nachdem ob mit der ursprünglichen (Tsusumi 2008) oder mit der weiterentwickelten Messmethode (Kurioka 2013) gerechnet wird. Da die Datengrundlage bei den bisherigen Studien recht unterschiedlich und teils nicht repräsentativ (Kurioka 2013; Montano 2016) sind, wäre eine Meta-Analyse, die die Messmethode berücksichtigt und die Ergebnisse vergleicht ein zweckmäßiger Weg, um hier Klarheit zu erlangen. Der direkte Vergleich von Verteilungsparametern und Effektgrößen zwischen Studien unter Berücksichtigung des Ratingverfahren/Messmethoden gäbe Klarheit zur Vergleichbarkeit der unterschiedlichen Messmethode, da hier auch Selektions- oder länderspezifische Effekte berücksichtigt werden.

Mittlerweile hat sich die neuere Messmethode als verwendete Methode weitestgehend durchgesetzt, auch im SOEP läuft die neue Welle 2021 (Daten seit 2024 verfügbar) ausschließlich das neue Messinstrument, so dass ein Vergleich der beiden Messmethoden daher im SOEP ausschließlich mit der 33. Welle aus dem Jahr 2016 möglich ist. Erstmals wurde in Siegrist et al. (2014) das einstufige Verfahren mit nur vier Kategorien vorgestellt und empfohlen nur noch diese zu verwenden. Da der Einsatz dieser weiterentwickelten Messmethode schleichend passiert, ist nicht ganz klar, ab welchem Zeitpunkt die komplette Wende sich vollzogen hat. Darüber hinaus wird in der Forschung die Effort-Reward Imbalance Skala nicht immer einheitlich angewendet, weshalb es nicht immer klar ist, welches Messinstrument in den Untersuchungen Anwendung findet. In Zukunft sollte ausschließlich, wie von Siegrist (2014) empfohlen, einheitlich die einstufige Messmethode des Effort-Reward-Imbalance Modell angewendet werden.

## Anhang

**Tab. 8 Missing Values nach Messmethode des Effort-Reward-Imbalance Model (Angaben in Prozent), Sample I (N=14,591) und Sample II (N=598)**

		zweistufig	einstufig
E1	Aufgrund des hohen Arbeitsaufkommens besteht häufig Zeitdruck	1.17	0.65
E2	Bei meiner Arbeit werde ich häufig unterbrochen und gestört	1.17	0.64
E3	Im Laufe der letzten beiden Jahre ist meine Arbeit immer mehr geworden	2.17	3.47
R1	Meine Aufstiegschancen in meinem Betrieb sind schlecht	7.36	6.87
R2	Ich erfahre – oder erwarte – eine Verschlechterung meiner Arbeitssituation	2.34	2.28
R3	Mein eigener Arbeitsplatz ist gefährdet	1.67	1.93
R4	Ich erhalte von meinen Vorgesetzten die Anerkennung, die ich verdiene	9.36	6.44
R5	Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich die erfahrende Anerkennung für angemessen	8.19	5.17
R6	Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich meine persönlichen Chancen des beruflichen Fortkommend für angemessen	11.04	8.55
R7	Wenn ich an all die erbrachten Leistungen denke, halte ich mein Gehalt/meinen Lohn für angemessen	5.35	2.52

Quelle: Eigene Darstellung

## Quellenverzeichnis

- de Araújo, T. M., Siegrist, J., Moreno, A. B., de Jesus Mendes da Fonseca, M., Barreto, S. M., Chor, D., & Griep, R. H. (2019). Effort-reward imbalance, over-commitment and depressive episodes at work: Evidence from the ELSA-Brasil cohort study. *International Journal of Environmental Research and Public Health*, 16(17), 3025. <https://doi.org/10.3390/ijerph16173025>
- Bamberg, E., & Staar, H. (2019). Gesundheit und Sicherheit. In H. Schuler & K. Moser (Hrsg.), *Lehrbuch Organisationspsychologie* (6. überarb. Aufl., S. 427–472). Hogrefe.
- Bühner, M., & Ziegler, M. (Hrsg.). (2017). *Statistik für Psychologen und Sozialwissenschaftler* (2. Aufl.). Pearson Deutschland.
- Calnan, M., Wainwright, D., & Almond, S. (2000). Job strain, effort-reward imbalance and mental distress: A study of occupations in general medical practice. *Work & Stress*, 14(4), 297–311. <https://doi.org/10.1080/02678370110040920>
- Dragano, N., Siegrist, J., Nyberg, S. T., Lunau, T., Fransson, E. I., Alfredsson, L., Bjorner, J. B., Borritz, M., Burr, H., Erbel, R., Fahlén, G., Goldberg, M., Hamer, M., Heikkilä, K., Jöckel, K.-H., Knutsson, A., Madsen, I. E. H., Nielsen, M. L., Nordin, M., Oksanen, T., Pejtersen, J. H., Pentti, J., Rugulies, R., Salo, P., Schupp, J., Singh-Manoux, A., Steptoe, A., Theorell, T., Vahtera, J., Westerholm, P. J. M., Westerlund, H., Virtanen, M., Zins, M., Batty, G. D., & Kivimäki, M. (2017). Effort-reward imbalance at work and incident coronary heart disease: A multicohort study of 90,164 individuals. *Epidemiology*, 28(4), 619–626. <https://doi.org/10.1097/EDE.0000000000000666>
- Eddy, P., Heckenberg, R., Wertheim, E. H., Kent, S., & Wright, B. J. (2016). A systematic review and meta-analysis of the effort-reward imbalance model of workplace stress with indicators of immune function. *Journal of Psychosomatic Research*, 91, 1–8. <https://doi.org/10.1016/j.ipsychores.2016.10.003>
- Fietze, S. (2011). Arbeitszufriedenheit und Persönlichkeit: „Wer schaffen will, muss fröhlich sein!“. German Socio-Economic Panel Study (SOEP), 2011(388), 1–47. <https://doi.org/10.6102/soep.v388>
- Greif, S., Bamberg, E., & Semmer, N. (Hrsg.). (1991). Psychischer Streß am Arbeitsplatz. Hogrefe.
- Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., & Schupp, J. (2019). The German Socio-Economic Panel Study (SOEP). *Journal of Economics and Statistics*, 239(2), 345–360. <https://doi.org/10.1515/jbnst-2018-0022>
- Häfner, H. (Hrsg.). (1999). Gesundheit - unser höchstes Gut?. Springer.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kinman, G., & Jones, F. (2008). A life beyond work? Job demands, work-life balance, and wellbeing in UK academics. *Journal of Human Behavior in the Social Environment*, 17(1–2), 41–60. <https://doi.org/10.1080/10911350802165478>
- Kivimäki, M., Vahtera, J., Elovainio, M., Virtanen, M., & Siegrist, J. (2007). Effort-reward imbalance, procedural injustice and relational injustice as psychosocial predictors of health: Complementary or redundant models? *Occupational and Environmental Medicine*, 64(10), 659–665. <https://doi.org/10.1136/oem.2006.031310>
- Kurioka, S., Inoue, A., & Tsutsumi, A. (2013). Optimum cut-off point of the Japanese short version of the effort-reward imbalance questionnaire. *Journal of Occupational Health*, 55(3), 340–348. <https://doi.org/10.1539/joh.12-0179-OA>
- Lazarus, R. S., & Folkman, S. (1984). Stress, appraisal, and coping. Springer.
- Li, J., Yang, W., Cheng, Y., Siegrist, J., & Cho, S.-I. (2005). Effort-reward imbalance at work and job dissatisfaction in Chinese healthcare workers: A validation study. *International Archives of*

- Occupational and Environmental Health, 78(3), 198–204. <https://doi.org/10.1007/s00420-004-0580-y>
- 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*, 11(3), 320–341. [https://doi.org/10.1207/s15328007sem1103\\_2](https://doi.org/10.1207/s15328007sem1103_2)
- Mo, P. K., Cheng, Y., & Lau, J. T. (2020). Work-related factors on mental health among migrant factory workers in China: Application of the demand-control and effort-reward imbalance model. *Health & Social Care in the Community*. <https://doi.org/10.1111/hsc.13176>
- Montano, D., Li, J., & Siegrist, J. (2016). The measurement of effort-reward imbalance (ERI) at work. In D. Montano, J. Li, & J. Siegrist (Hrsg.), *Work stress and health in a globalized economy: The model of effort-reward imbalance* (S. 21–42). Springer. [https://doi.org/10.1007/978-3-319-32937-6\\_2](https://doi.org/10.1007/978-3-319-32937-6_2)
- Msaouel, P., Keramaris, N. C., Apostolopoulos, A. P., Syrmos, N., Kappos, T., Tasoulis, A., Tripodaki, E.-S., Kagiampaki, E., Lekkas, I., & Siegrist, J. (2012). The effort-reward imbalance questionnaire in Greek: Translation, validation and psychometric properties in health professionals. *Journal of Occupational Health*, 54(2), 119–130. <https://doi.org/10.1539/joh.11-0197-oa>
- Penz, M., Siegrist, J., Wekenborg, M. K., Rothe, N., Walther, A., & Kirschbaum, C. (2019). Effort-reward imbalance at work is associated with hair cortisol concentrations: Prospective evidence from the Dresden Burnout Study. *Psychoneuroendocrinology*, 109, 104399. <https://doi.org/10.1016/j.psyneuen.2019.104399>
- Rugulies, R., Aust, B., & Madsen, I. E. (2017). Effort-reward imbalance at work and risk of depressive disorders: A systematic review and meta-analysis of prospective cohort studies. *Scandinavian Journal of Work, Environment & Health*, 43(4), 294–306. <https://doi.org/10.5271/sjweh.3632>
- Schupp, J. (2012). Das Sozio-oekonomische Panel (SOEP). *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*. <https://doi.org/10.1515/ibnst-2018-0022>
- Siegrist, J. (1996a). *Soziale Krisen und Gesundheit: Eine Theorie der Gesundheitsförderung am Beispiel von Herz-Kreislauf-Risiken im Erwerbsleben*. Hogrefe.
- Siegrist, J. (1996b). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology*, 1(1), 27–41. <https://doi.org/10.1037/1076-8998.1.1.27>
- Siegrist, J., & Dragano, N. (2008). Psychosoziale Belastungen und Erkrankungsrisiken im Erwerbsleben: Befunde aus internationalen Studien zum Anforderungs-Kontroll-Modell und zum Modell beruflicher Gratifikationskrisen. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 51(3), 305–312. <https://doi.org/10.1007/s00103-008-0461-5>
- Siegrist, J., Wege, N., Pühlhofer, F., & Wahrendorf, M. (2008). A short generic measure of work stress in the era of globalization: Effort-reward imbalance. *International Archives of Occupational and Environmental Health*, 82(8), 1005–1014. <https://doi.org/10.1007/s00420-008-0384-3>
- Siegrist, J. (2012). *Effort-Reward Imbalance at Work: Theory, Measurement and Evidence*. Dordrecht: Springer.
- Siegrist, J., Dragano, N., Nyberg, S. T., Lunau, T., Alfredsson, L., Erbel, R., Fahlén, G., Goldberg, M., Jöckel, K.-H., Knutsson, A., Leineweber, C., Magnusson Hanson, L. L., Nordin, M., Rugulies, R., Schupp, J., Singh-Manoux, A., Theorell, T., Wagner, G. G., Westerlund, H., Zins, M., Heikkilä, K., Fransson, E. I., & Kivimäki, M. (2014). Validating abbreviated measures of effort-reward imbalance at work in European cohort studies: The IPD-Work Consortium. *International Archives of Occupational and Environmental Health*, 87(3), 249–256. <https://doi.org/10.1007/s00420-013-0855-z>
- Siegrist, J., Li, J., & Montano, D. (2014). Psychometric properties of the effort-reward imbalance questionnaire. *Duesseldorf, Germany: Department of Medical Sociology*.



- Siegrist, J., & Wahrendorf, M. (Hrsg.). (2016). *Work Stress and Health in a Globalized Economy: The Model of Effort-Reward Imbalance*. Springer International Publishing.  
<https://doi.org/10.1007/978-3-319-32937-6>
- Siegrist, J., & Li, J. (2016). Associations of extrinsic and intrinsic components of work stress with health: A systematic review of evidence on the effort-reward imbalance model. *International Journal of Environmental Research and Public Health*, 13(4), 432.  
<https://doi.org/10.3390/ijerph13040432>
- Solis-Soto, M. T., Schön, A., Parra, M., & Radon, K. (2019). Associations between effort-reward imbalance and health indicators among school teachers in Chuquisaca, Bolivia: A cross-sectional study. *BMJ Open*, 9(3), e025121. <https://doi.org/10.1136/bmjopen-2018-025121>
- Stansfeld, S., & Candy, B. (2006). Psychosocial work environment and mental health—A meta-analytic review. *Scandinavian Journal of Work, Environment & Health*, 32(6), 443–462.  
<https://doi.org/10.5271/sjweh.1050>
- Taibi, Y., Metzler, Y. A., Bellingrath, S., & Müller, A. (2021). A systematic overview on the risk effects of psychosocial work characteristics on musculoskeletal disorders, absenteeism, and workplace accidents. *Applied Ergonomics*, 95, 103434. <https://doi.org/10.1016/j.apergo.2021.103434>
- Tobsch, V., Matiaske, W., Holst, E., Schmidt, T., Seifert, H., & Hartmut, T. (2018). Mehr oder weniger arbeiten? Es kommt darauf an, wie man fragt. *SOEPpapers on Multidisciplinary Panel Data Research*, No. 960. <https://doi.org/10.13140/RG.2.2.25815.78244>
- Tsutsumi, A., Iwata, N., Wakita, T., Kumagai, R., Noguchi, H., & Kawakami, N. (2008). Improving the measurement accuracy of the effort-reward imbalance scales. *International Journal of Behavioral Medicine*, 15(2), 109–119. <https://doi.org/10.1080/10705500801929718>
- van Vegchel, N., de Jonge, J., Bosma, H., & Schaufeli, W. (2005). Reviewing the effort-reward imbalance model: Drawing up the balance of 45 empirical studies. *Social Science & Medicine*, 60(5), 1117–1131. <https://doi.org/10.1016/j.socscimed.2004.06.043>