Organizational Practices for the Aging Workforce: Conceptualization, Operationalization, Validation, and Application of the Later Life Workplace Index

Von der Fakultät Wirtschaftswissenschaften

der Leuphana Universität Lüneburg zur Erlangung des Grades

Doktor der Wirtschaftswissenschaften

– Dr. rer. pol. –

genehmigte Dissertation von

Max R. Wilckens

geboren am 15. März 1988 in Itzehoe

Eingereicht am:	02. Februar 2021
Mündliche Verteidigung (Disputation) am:	20. Juli 2021
Erstbetreuer und Erstgutachter:	Prof. Dr. Jürgen Deller
Zweitgutachter:	Prof. Mo Wang, PhD
Drittgutachter:	Prof. Dr. Stephan Alexander Böhm

Die einzelnen Beiträge des kumulativen Dissertationsvorhabens sind oder werden wie folgt veröffentlicht:

Wilckens, M. R., Wöhrmann, A. M., Adams, C., Deller, J., & Finkelstein, R. (2020). Integrating the German and US perspective on organizational practices for later life work: The Later Life Work Index. In S. Czaja, J. Sharit, & J. James (Eds.), *Current and emerging trends in aging and work* (pp. 59–79). Springer. https://doi.org/10.1007/978-3-030-24135-3_4

Wilckens, M. R. (2021). Assessment of the reliability of multi-valued data: An analysis of Krippendorff's alpha coefficient. Manuscript in preparation.

Wilckens, M. R., Wöhrmann, A. M., Deller, J., & Wang, M. (2020). Organizational practices for the aging workforce: Development and validation of the Later Life Workplace Index. *Work, Aging and Retirement,* Advance online publication. https://doi.org/10.1093/workar/waaa012

Wilckens, M. R., Wöhrmann, A. M., Finsel, J. S., & Deller, J. (2021). Health and the intention to retire: Exploring the moderating effects of organizational practices. Manuscript submitted for publication.

Acknowledgements

I am very grateful to Prof. Dr. Jürgen Deller for his support over the past four years, for his confidence in me and my studies, for his extensive support for the *Later Life Workplace Index (LLWI)* project, and last but not least for providing his professional and personal contacts for the good of the LLWI project and this dissertation. Thank you. I am likewise very grateful to my mentor Dr. Anne Wöhrmann for her ideas and constructive feedback that supported me to develop and improve my studies, for her commitment for the LLWI project, and for providing me a second opinion in all major decisions throughout the years. Thank you. I thank Prof. Mo Wang, PhD, for his constructive criticism and very helpful contributions to Article 3, as well as valuable financial support for the studies.

My thanks go also to the members of the "Goinger Kreis", who allowed me to research their companies, provided valuable feedback from a practitioner's perspective, and supported the studies financially. I thank Ruth Finkelstein and Caitlin Adams for providing the U.S. data for Article 1; Julia Finsel for collecting parts of the data of Article 3; Martina Schmeink and Rudolf Kast from "Das Demographie Netzwerk" and Clemens Volkwein from "HessenChemie" for promoting my studies in their networks of companies and institutions in Germany; and several students from the Leuphana University of Lüneburg, who supported this dissertation by collecting data for my studies.

I also thank the "Boston Consulting Group" for granting me an educational leave of absence and the "Konrad Adenauer Stiftung" for a personal scholarship that allowed me to conduct my research independently of external funding.

I further thank my colleagues Julia Finsel, Jana Bucher, and Dina Gericke at Leuphana, and Cornelius Hafner and Nicolas Heinitz at the Boston Consulting Group for their support and many helpful discussions.

Last but not least, I am very grateful to my family and my friends who greatly supported me throughout all phases of this dissertation. I am especially grateful to my wife Annabelle for her patience and love. Thank you!

Abstract

This dissertation focused on the nature and role of organizational practices for the employment of older people and the extension of their working lives. The set of four articles is driven by the objective to further deepen our understanding of how organizations can facilitate ageing at work to the benefit of both, employees and employers. Findings are empirically based on qualitative expert interview data from Germany and the U.S. and several quantitative field studies among older employees in Germany. To bridge gaps in measurement of organizational practices related to aging at work, this dissertation proposes a new comprehensive, multifaceted, and thoroughly conceptualized measure of organizational practices related to aging at work, the Later Life Workplace Index (LLWI). Through the course of the four articles the LLWI is conceptually developed based on qualitative interview data, operationalized, validated based on multiple field studies among older workers, and applied in a multi-level study among older employees of 101 organizations. Results suggest that organizational practices are not uniform, but multifaceted in their presence within organizations and their effects for the employment of older workers. The LLWI distinguishes nine domains of practices including an age-friendly organizational climate, work design, individual development, and practices tailoring the retirement transition. Thus, it may lay the foundation for more granular organizational level research in the field. Further, this dissertation's fourth article applies the LLWI and argues based on person-environment fit and socio-emotional selectivity theory that organizational practices address different individual needs and, thus, affect employment depending on employees' individual characteristics. Results suggest that older employees' retirement intentions are effected by individual development, transition-to-retirement, and continued employment practices depending on their health resources. Application of the new measure in practice to improve organizations' response to the aging workforce and opportunities for future research based on the LLWI are discussed.

Keywords: aging at work, human resource practices, older workers, organizational practices, scale development

Table of Contents

List of Figures

List of Tables

Introduction

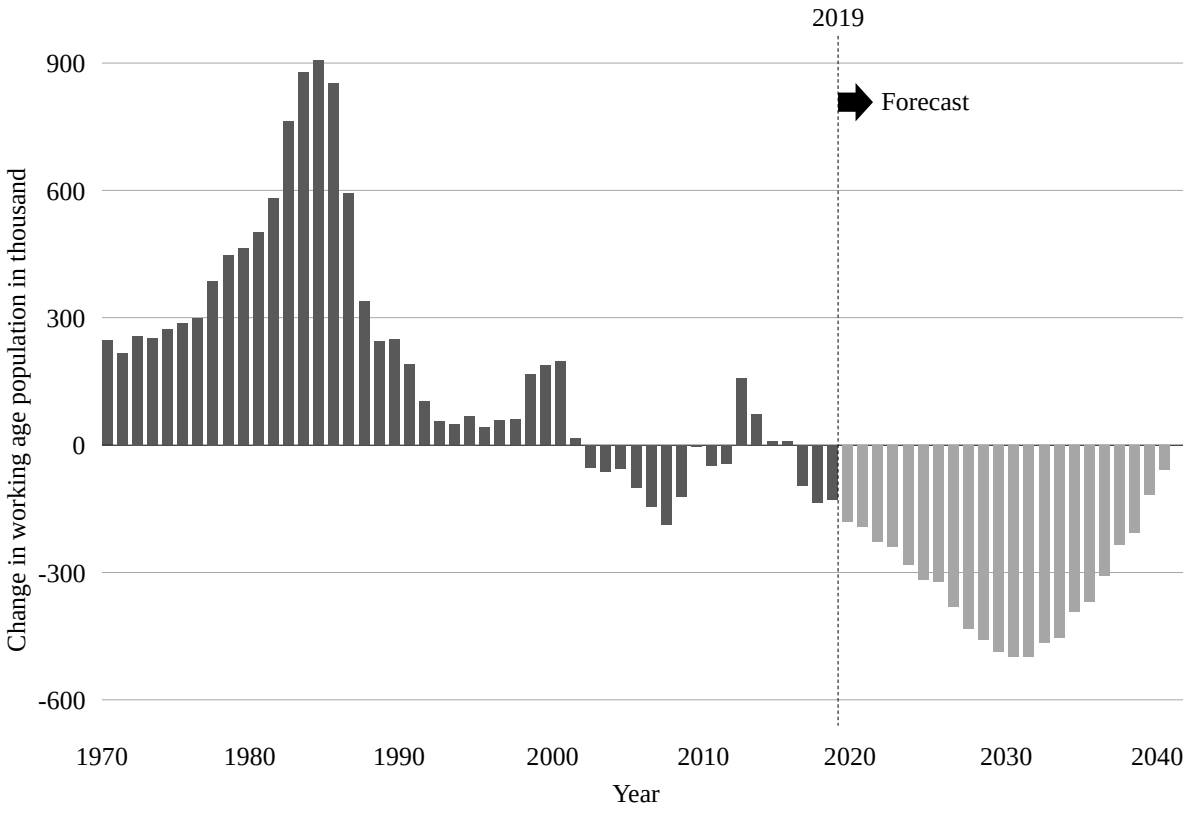
Life expectancies are raising almost everywhere in the world (United Nations, 2019). Living longer opens up great opportunities for every one of us to fulfill one's dreams and lifetime plans. However, in the wake of increased life expectancies and low birth rates, societies of most industrialized countries are aging as a whole. As a consequence, the share of people in working age gradually decreases, resulting in a raising dependency ratio between people dependent on insurances and family support (children and pensioners) and those financing society by means of working income. Consequently, pension systems suffer from a lack of funding, and social security premiums are raising (Mahon & Millar, 2014). In the upcoming decades, some countries with particularly low birth rates such as Germany will even face severe labor force shortages due to more employees retiring than entering the workforce (see Figure 1, Statistisches Bundesamt, 2015). As a consequence for organizations, recruiting skilled employees will become more and more challenging.

Working beyond normal retirement age has been argued to be an important part of the solution for several reasons: On the societal level, extended working lives relieve social security systems as employees pay into pension, health, and nursing care systems for a longer period of their lives, while receiving pensions for less years (Moen, 2016; Morrow-Howell et al., 2018). Besides public pensions, occupational pensions, and individual pension provision, partially or fully delaying retirement has thus been characterized as the *fourth pillar* of pension systems (Giarini, 2012).

Furthermore, on the individual level, research of recent years provided evidence that some individuals reaching normal retirement age actually want to work longer for personal reasons (e.g., Wöhrmann et al., 2014). A rapidly growing number of older employees takes the

Figure 1

Annual increases and decreases of the working age population aged 18 to 67 in Germany. Projection of the federal bureau of statistics, variant G2-L2-W2 (Statistisches Bundesamt, 2015, 2019)



opportunity to continue working, either in their former job, in a new job, or by becoming self-employed. These people value for example, that their work creates meaning for their lives, provides them with social contacts to colleagues and business partners, yields some extra money, and structures their day-to-day life (Sackreuther et al., 2017). Consequently, it is not only in the interest of society to promote prolonged working lives, but also in a growing number of peoples' own interest to work longer.

On the organizational level, older employees working longer expand the workforce and support solving skill and labor shortages. Organizations gain new opportunities to find the right people for their jobs, may retain knowledge within the organization by retaining older workers (e.g., Li et al., 2021), and may benefit from age diversity, for example in terms of different perspectives of younger and older colleagues (Kunze et al., 2013). However, to benefit from

extended working lives, organizations need to successfully include and employ older employees in an economically meaningful manner.

Whether working lives can be extended not only depends on the political regulatory framework, but also on how well older employees and organizations satisfy each other's demands (e.g., Lahlouh et al., 2019). Both, the organizations and older employees have to mutually benefit from working in older age. Therefore, researchers have focused on the needs and motives of older workers to continue working on the one hand side (e.g., Fisher et al., 2016). Yet, on the other hand side, research on how organizations successfully facilitate employment of older workers remains limited (e.g., Pak et al., 2019). In their well-received commentary Henkens et al. (2018, p. 809) ask the question, "How can employers make an aging work staff 'work'?" The authors lament a lack of evidence-based human resource management in the context of aging at work and little knowledge about employers' decision making in different contexts and under different regulatory frameworks. Certain organizational practices addressing older employees' needs, such as idiosyncratic deals, training and development, and an age-friendly organizational climate, have been found to increase older employees' motivation to continue working (Bal et al., 2012; Pak et al., 2021; Vignoli et al., 2019; Zaniboni, 2015) and to benefit older employees, for example, by means of or their job satisfaction (Visser et al., 2020). Yet, we do not know much about the effects of organizational practices for older workers on whether they are actually longer employed, on their individual performance at work, and on the organization's overall performance (see also von Bonsdorff et al., 2018). Neither, do we sufficiently understand why and under what conditions organizations implement certain practices for an aging workforce (Henkens et al., 2018).

An important restriction for current research results from a lack of validated and sufficiently detailed measures of organizational practices (Boehm et al., 2014; Pak et al., 2021). Given that researchers have named very different organizational practices that facilitate successful employment of older employees (e.g., Armstrong-Stassen, 2008; Kooij et al., 2014),

multifaceted measures with thorough conceptual coverage are required to further deepen our understanding of how organizational practices support aging at work.

This dissertation adds to our understanding of how organizations can support extended working lives by developing a new taxonomy and measure of organizational practices for the aging workforce, the Later Life Workplace Index (LLWI[1]). The LLWI surpasses existing measures in the literature by a multifaceted conceptualization distinguishing nine domains of practices and a thoroughly developed and validated operationalization. The nine domains cover practices addressing an organizational climate[2], leadership, work design, health management, individual development, knowledge management, the transition to retirement, continued employment, and health and retirement coverage. The new measure is intended to not only facilitate more profound and detailed research on organizational practices for the aging workforce closing outlined research gaps, but also to allow organizations to assess themselves in order to identify areas for improvement regarding the employment of an aging workforce in practice. In the course of this dissertation, the new measure is developed, operationalized, validated, and applied in a multi-level organizational study investigating the effects of organizational practices in the relation between older employees' health and their retirement intentions. Thereby, the articles of this dissertation provide answers to three main research questions: (1) Results show which organizational practices are relevant in the context of aging at work and provide a comprehensive conceptualization. (2) Results propose how

[1] The label 'Later Life Workplace Index' was adapted throughout the research project. It was first named Silver Work Index (SWI) in the initial qualitative study conducted in Germany and renamed to Later Life Work Index (LLWI) after the incorporation of the U.S. perspective. Finally, to be more precise we changed "work" to "workplace" during the validation phase.

[2] The organizational climate domain was initially named 'organizational culture' (see Article 1). We changed the name during operationalization, because "climate" better captures, what the LLWI assesses in this domain.

organizational practices for the aging workforce can be measured validly. And (3), results add to our understanding, how individual organizational practices effect retirement intentions and, consequently, how organizations can facilitate aging at work.

Organizational Practices for the Aging Workforce

Organizational practices tailored to aging at work are an important part of organizations' response to the aging workforce. While Human Resource (HR) practices such as high involvement work practices (Lawler III, 1992) are often researched irrespectively of workers' age, researchers have concurrently argued that organizational practices beneficial for certain age groups may differ from general practices. Lifespan theory suggests that needs and circumstances change throughout the lifespan because of diverting life courses and aging processes that do not affect individuals uniformly (Rudolph, 2016; Zacher, 2015b). Consequently, researchers have named a variety of practices particularly relevant for older employees, such as work design practices that account for potential decreases of physical resources among older employees, mentoring programs in which mature employees can pass on their knowledge to younger employees, an age-inclusive organizational climate, phased retirement options, and opportunities to continue working post normal retirement age (see Article 1 and 3 for more detail). Thus, organizational practices for the aging workforce include not only HR practices, but also attributes of an age-friendly organizational climate and leadership (Silver et al., 2019; Taneva & Arnold, 2018). Moreover, the organizational practices not only refer to formal practices as for example, offering flexible work arrangements but also informal practices that exist as habits without active management by the organization's leadership (cf. Oostrom et al., 2016; Taneva & Arnold, 2018). For example, knowledge sharing between older and younger employees may result from informal exchange without being actively encouraged or promoted.

Mechanisms researchers have proposed to reason the benefits of certain practices are as broad as the practices themselves. For example, organizational practices have been argued to increase the fit between older employees and their jobs by accommodating the employee in case of an age-related decrease in personal or contextual resources (e.g., Taylor & Walker, 1998), to preventively maintain older employees resources, to adjust jobs to better utilize employees' resources (e.g., Kooij et al., 2014), to support their individual ability to optimally utilize their resources and compensate for losses (e.g., Taneva & Arnold, 2018), to signal older employees that they are valued and recognized by the organization (e.g., Boehm et al., 2014), to strengthen the relation between older employees and their organization and to intensify social exchange (e.g., Bal et al., 2012), or to provide them with perspectives for personal growth at work (e.g., Oostrom et al., 2016; Zacher & Yang, 2016). Different organizational practices serve the management of an aging workforce and support employment of older people via various theoretical mechanisms. However, an evidence-based picture, which practice serves which purpose under what conditions is still missing. To support an evidence-based response to the aging workforce within organizations, research need to examine both, how each practices works, under which conditions, for whom, and which practices do not pay off for older employees or their organizations (Henkens et al., 2018).

A prerequisite to disentangle the various practices and their effects is a clear conceptualization and measurement of the practices. Researchers have used several measures to assess organizational practices for the aging workforce, such as Kooji's (2014) bundles of HR practices, Zacher's and Yang's (2016) organizational climate for successful aging, and Taneva's and Arnold's (2018) 8-item scale on organizational practices (see Article 3 for an analysis of existing measures). However, all these measures fall short either in psychometric validation or in assessing organizational practices for the aging workforce sufficiently detailed to capture the multifacetedness of practices within organizations. Yet, a multifaceted assessment of practices is important to further deepen our understanding of how organizations

can improve aging at work and deal with an aging workforce. The present dissertation aims to close this gap by proposing and applying the LLWI as a new measure.

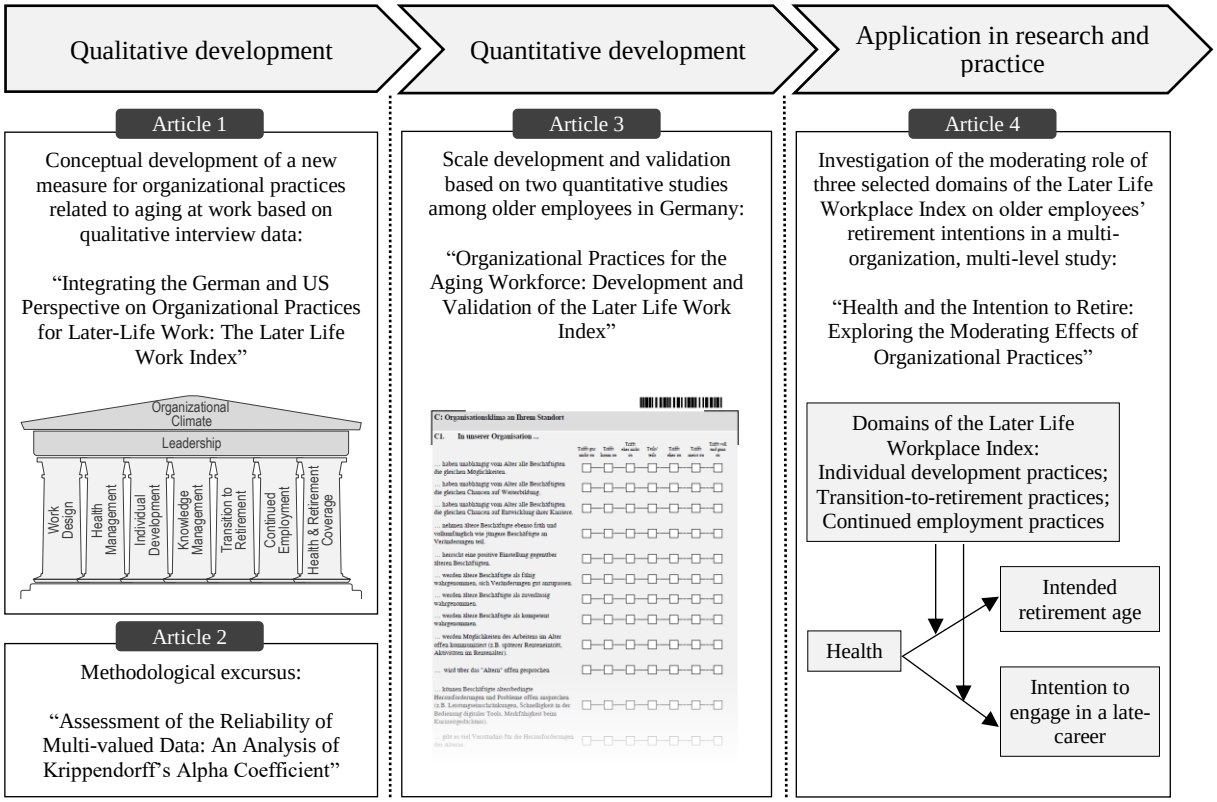
Outline of This Dissertation

This dissertation comprises four articles describing the conceptualization, operationalization, validation and application of the LLWI. By developing the LLWI this dissertation targets a better understanding of how organizations can support the aging workforce. The four articles are sequential as depicted in Figure 2.

In the first article (*Integrating the German and US Perspective on Organizational Practices for Later-Life Work: The Later Life Workplace Index*) my co-authors and I laid out the qualitative basis of the LLWI. By integrating qualitative data from 27 expert interviews in Germany and findings from the *Age Smart Employer Award* in the U.S. we advanced previous

Figure 2

This dissertation's process model



endeavors to categorize organizational practices for the aging workforce (see Wöhrmann et al., 2018). The article describes the nine LLWI domains, differences between practices in Germany and the U.S., and our qualitative approach to develop the LLWI's taxonomy.

The second article (*Assessment of the Reliability of Multi-Valued Data: An Analysis of Krippendorff's Alpha Coefficient*) addresses a methodological challenge during reliability assessment of the LLWI's taxonomy. Reliability assessment for qualitative research requires independent coders to confirm the findings extracted from data such as qualitative interviews (e.g., Mayring, 2010). Agreement coefficients allow researchers to assess the reliability of their findings based on multiple coders' evaluations. Recent advancements by Krippendorff (2019) complemented existing coefficients to allow reliability assessment for multi-valued data. Multi-valued data occurs if a single unit of analysis (e.g., an interview section, paragraph, sentence, or thought) applies to more than one code. The qualitative LLWI data was multi-valued, because interviewed experts discussed challenges, programs, and interventions, which often covered more than one single LLWI domain. To assess the LLWI's reliability, I undertook Krippendorff's alpha for multi-valued data (Krippendorff, 2019) a critical review based on Monte Carlo simulation and proposed an amendment to his formula.

The third article (*Organizational Practices for the Aging Workforce: Development and Validation of the Later Life Workplace Index*) outlines the scale development and validation of the LLWI along three sequential studies. My co-authors and I operationalized the LLWI's nine domains (Study 1), developed a psychometrically sound scale based on a sample of 609 workers in Germany (Study 2) and validated the resulting 80-item measure with several criterion measures from the literature on a second sample of 349 older workers in Germany (Study 3). The resulting measure allows researchers and practitioners to assess organizational practices in the context of aging at work in a psychometrically sound and multifaceted manner. Thereby, the measure enables researchers to study organizational practices for the aging workforce more

precise and practitioners to derive more fine-grained conclusions to improve practices in their organizations.

Finally, the fourth article (*Health and the Intention to Retire: Exploring the Moderating Effects of Organizational Practices*) adds to our understanding of how organizational practices facilitate employment of older employees. Based on a multi-level study among 101 organizations in Germany, we explored the moderating role of three career-development and retirement related domains of the LLWI on the relation between older employees' health and their intentions to retire early or to continue working even beyond normal retirement age. Drawing upon work adjustment and socioemotional selectivity theory we argue that selected organizational practices improve older employees' retirement timing by supporting healthy older employees to extend their working lives, while allowing those with poorer health to retire early. The study provides additional evidence for the effect of selected LLWI practices in managing an aging workforce. Moreover, the study provides evidence for differences between practices and that various practices need to be disentangled to improve our understanding and to derive meaningful recommendations for practice.

In total, the four articles cover the journey from developing a psychometrically sound measure for organizational practices in the context of aging at work to its application in research and practice. The dissertation tackles methodological issues during concept development and shows exemplarily how the new measure can advance research in the field of aging at work.

1

Integrating the German and US Perspective on Organizational Practices for Later-Life Work: The Later Life Work Index

Max R. Wilckens, Anne M. Wöhrmann, Caitlin Adams, Jürgen Deller, and Ruth Finkelstein

Published as: Wilckens, M. R., Wöhrmann, A. M., Adams, C., Deller, J., & Finkelstein, R. (2020). Integrating the German and US perspective on organizational practices for later life work: The Later Life Work Index. In S. Czaja, J. Sharit, & J. James (Eds.), *Current and emerging trends in aging and work* (pp. 59–79). Springer. https://doi.org/10.1007/978-3-030-24135-3_4

Abstract

Later life work is on the rise in most developed countries and organizational practices are important for its successful implementation. However, an integrated holistic perspective on successful management practices is still missing. Drawing on two qualitative frameworks of organizational practices for later life work, this chapter examines similarities and differences between management of older employees in the U.S. and Germany. Based on evidence from the Age Smart Employer Award in New York City and the Silver Work Index (SWI) in Germany an integrated, inter-cultural framework for organizational later life work practices is proposed. The comparison reflects the differences in the countries' social systems and legislation. While Germany's rigid social security system requires emphasis on more individualized and flexible transition solutions into retirement, U.S. practices emphasize the importance of retirement savings and health coverage. Findings suggest a revised integrated set of practices regarding organizational culture, leadership, and several specific human resource (HR) management domains including work design, health management, individual development, and transition to retirement. This integration thereby contributes to the development of a well-founded index for good organizational management of later life work. The revised index is called Later Life Work Index (LLWI) and aims to enable organizations to self-assess their capabilities regarding successful employment of older employees, to identify areas for improvement, and to serve as a source for best practices.

Keywords: active ageing, bridge employment, index, later life work, older employees, organizational practices, retirement, work design, working conditions

Fed by the demographic change in many countries and increasing life expectancies later life work is on top of the agenda for not only older adults and politicians, but also for business. By hiring and employing older employees under appropriate working conditions organizations are the enablers of later life work. This implies new perspectives to actively and successfully participate in work life through meaningful and age-friendly employment for employees aged 55 and above, as well as individuals in retirement-age. Differences in labor market regulations and social systems between countries, however, require organizations to prioritize different organizational practices in order to allow for successful later life work within their respective regulatory frame. Literature on the impact of countries' legislation, regulation, and work culture on the organizational practices required is so far, however, very limited. In order to understand how to cope with and successfully leverage later life work from an organizational perspective, it is hence important to extend the debate on organizational practices and conditions by an understanding of the country-specificity of the practices required.

The debate on how to manage later life work within organizations is still ongoing (Henkens et al., 2018), but the potential of employment for older employees prior and beyond their retirement age has generally been acknowledged on several levels: On a societal level, later life work relieves social systems by generating an own, independent income and decreases poverty among the elderly. Moreover, older employees solve labor and skill shortages, which are existent particularly in rural areas today already (European Commission & Economic Policy Committee, 2017). On the organizational level, older employees contribute their expert knowledge, not only regarding the growing customer group of older people, and their long-term experience which is beneficial for specific tasks (Göbel & Zwick, 2013). Finally, on the individual level, flexible models of retirement work ('bridge employment') ease the life-changing event of being retired and increase well-being (Dingemans & Henkens, 2015).

Thereby, adequate work for older employees positively impacts physical, mental and cognitive health (Hershey & Henkens, 2014; Staudinger et al., 2016).

On the other hand, successful later life work requires dedicated organizational practices and work conditions to leverage older employees potential. While many organizational conditions for successful employment obviously apply to employees of all age groups, such as ergonomic work design, individual development, knowledge management, and a supportive leadership style, some are specifically relevant for older employees or need to be specified for older employees, such as an age-friendly organizational culture and management of the transition into retirement. Research has come up with country-specific perspectives on the organizational antecedents of successful later life work. Within Germany, Wöhrmann, Deller, and Pundt (2018) developed the Silver Work Index (SWI), which is intended to serve as a diagnostic tool for organizations to assess their capabilities regarding successful employment of older employees. The index includes elements of organizational culture, leadership and specific human resource (HR) practices, which are defined conceptually, but are yet to be operationalized for use as an assessment tool. In the U.S., organizational conditions for successful later life work have been identified by the Age Smart Employer Award (see Finkelstein et al., 2013). The award honors employers in New York City that implement practices to engage, successfully employ, and retain older workers.

By comparing organizational practices for later life work identified by the SWI in Germany and the Age Smart Employer Award in the U.S. this chapter firstly provides a revised version of the index representing both, the German, as well as the U.S. perspective on later life work, and secondly identifies country-specific practices and conditions for Germany and the U.S. Intercoder reliability results confirm the index' revised category system. Thereby, this chapter contributes towards a final, intercultural index for good organizational management of later life work. Future project phases will include the operationalization of the indicators as well as the validation of the index.

Differences in Demographics, Work Culture and Legislation

Steadily increasing life expectancies within our societies pose the question of how to manage extended working lives. However, the severity of the demographic problem and the prerequisites to solve the problem differ significantly between countries. Work practice in Germany and the U.S. is based on different legal labor market systems and work customs. Compared to Germany, the U.S. labor market is substantially more flexible, which is best characterized by the in- and outflows of unemployment. Across periods of economic prosperity and recession in- and outflow rates of unemployment in Germany are by a factor 5-10 lower than in the U.S. (Hertweck & Sigrist, 2015; Jung & Kuhn, 2014). Consequently, the average employee in Germany stays twice as long with the same employers as employees in the U.S. (see also Bureau of Labor Statistics, 2016; Eurofound, 2015). Differences are reasoned by higher employment protection and union bargaining power, higher unemployment benefits, as well as a lower matching efficiency (longer search periods) in Germany (Jung & Kuhn, 2014). Unionization has a historic tradition in Germany leading to strong regional collective employment agreements between employer associations and the unions for many industries, which limit freedom for individual negotiations and focus on job stability. Stability also results from statutory codetermination, which – contrarily to the U.S. – provides the works councils of companies larger than 2,000 employees with just under the half of the company's supervisory board's seats. Work practices moreover differ, as social security systems in the U.S. are not as comprehensive as in Germany. Besides higher unemployment benefits, most employees are mandatorily enrolled in public health insurance and pensions system, whereas the latter's future pension level becomes increasingly uncertain for the next generation due to the demographic change (Börsch-Supan & Wilke, 2004).

The demographic development towards longer lives in Germany is paired with constantly low birth rates leading to a substantially increase in the share of older employees over the next decades. This increase is driven by aging 'baby-boomer' age groups, but also

influenced by a higher labor market participation rate of older employees. Studies prove a significant increase in retirees working beyond the official retirement age (Eurofound, 2012). And this effect is further escalated by the politically raised standard retirement ages in Germany and many other European countries. Consequently, the European Commission (2017) projects a 35% increase in employees aged 55-74 between 2016 and 2030 raising the share of this age group among all employees from 18% to 24%. Not only politics, but also leading industries have declared extension of working lives and retention of older employees to a key priority in human resource management. Especially for industries and rural areas with skilled labor force shortages, working beyond retirement age becomes more and more common. However, in Germany for example retirement age legislation and regulation is still very inflexible for many industries. The standard retirement age is defined by the beginning of public pension payment, which has been at the age of 65 historically and will now be gradually increased to 67 by 2029 (European Commission & Economic Policy Committee, 2017). Despite the fact that retirement is not required by law at that age, many collective agreements foresee retirement at the beginning of pension payment (European Commission & Economic Policy Committee, 2017).

In the U.S. on the other hand employees aged 55 and above already account for 22% of the civilian workforce, increasing moderately to 24% in 2026 (Bureau of Labor Statistics, 2017). A more balanced population pyramid and historically higher flexibility in retirement age leads to much smaller changes to the workforce age distribution compared to the EU. However, the older age groups nevertheless account for highest growth rates also in the U.S. with an absolute increase between 2016 and 2026 of 18% for the 55+ age group and 58% for the 65+ age group. Similar to the trend in the EU, sectors of agriculture and skilled trades (i.e., tailoring) have some of the highest median ages as less young people are choosing or being trained to work in these industries (Bureau of Labor Statistics, 2017). As life expectancy has increased, the retirement age in the U.S. has also gradually been increasing from 65 to 67 for people born after 1959. This is the age where a person can begin collecting the full amount of their social

security benefit (Social Security Administration, 2018). Unlike most countries within the EU, the U.S. have abolished mandatory retirement ages (OECD, 2017) and are not affected by far-reaching collective agreements with mandatory retirement age, so that the transition to retirement is more flexible as in Germany. Aside from social security benefits, access to retirement savings plans, employer contributions, and pensions are hence largely dependent on the employer and the individual employment contract. In the U.S., as well as in the EU organizations hence face an aging workforce, so that it is in the interest of both, industry and society to retain and develop the potential of older employees.

Later-Life Work From an Organizational Perspective

While much research has been carried out on the individual antecedents and preferences for older employees and post-retirement work in recent years (Davis, 2003; Fasbender et al., 2014; Wang et al., 2008; Wang & Shultz, 2010), it is not yet understood, how successful employment of older employees can be enabled by the organization (Henkens et al., 2018). In contrast to the common opinion that a higher share of older employees lowers organizational productivity, recent studies have shown that this is not necessarily the case. Cross-sectional, as well as interventional longitudinal studies found individual organizational measures such as e.g., redesigning the assembly line and ergonomic adoption of the work place to sufficiently counteract the productivity disadvantages caused by older employees (Göbel & Zwick, 2013; Loch et al., 2010). Organizational practices have shown to improve performance, work ability, and motivation of older employees, even leading to increased willingness to continue work beyond retirement. Areas of action range from leadership and organizational culture to health promotion, knowledge management, and work design (Armstrong-Stassen & Templer, 2006; Klaffke, 2014; Kunze et al., 2013; Naegele & Walker, 2006; Schuett, 2014; Zacher & Yang, 2016). Moreover, older employees show lower error rates leading to quality advantages, if leveraged by appropriate work design practices, as e.g., in mixed teams of young and old aged

employees, so that competences and capabilities of younger and older employees complement each other (Börsch-Supan & Weiss, 2016; Göbel & Zwick, 2013).

The effects of appropriate organizational practices and work conditions on organizational outcomes have theoretically been reasoned by both, a resource-based perspective on the organization leveraging human capital, as well as a behavioral perspective leveraging the practices' effect to encourage productive behaviors from the employees (Jiang et al., 2012). Accordingly, studies found skill-, motivation-, and opportunity-enhancing practices to positively influence financial outcomes of the firm mediated by human capital and employee motivation (Jiang et al., 2012). Thus, those practices have also been found to reduce negative effects of age on work ability and subsequently organizational performance (von Bonsdorff et al., 2018). Kooij et al. (2013) however showed that practices' impact on performance differs with age. Moreover, organizational commitment as a key antecedent of organizational performance is influenced by different sets of organizational practices dependent on the employee's age (Conway, 2004). While many organizational practices, as for example a supportive leadership style, are beneficial for employment of all age groups, it is hence argued that older employees require dedicated practices in certain domains.

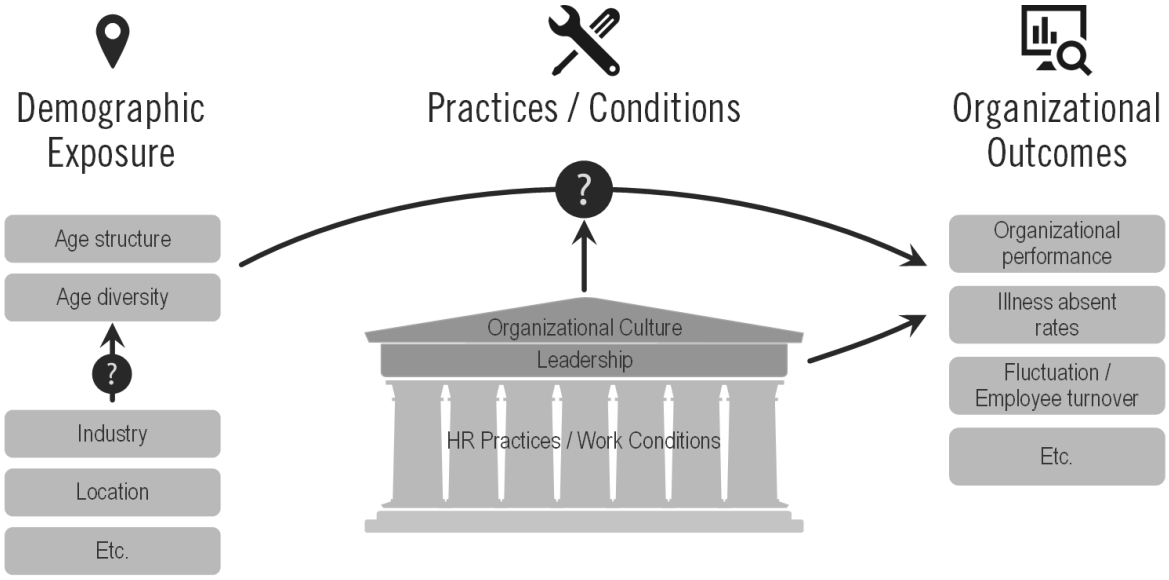
In practice, however, later life work is substantially retarded by age stereotypes and norms at the workplace (Posthuma & Campion, 2009). Henkens (2005) found that managers stereotype older employees regarding their productivity, reliability and adaptively. Stereotypes result from differing age norms and related work ability and productivity assumptions about older employees (Conen et al., 2012; Karpinska et al., 2013). Employers recognize a link between an aging workforce and an increased gap of labor cost and productivity (Conen et al., 2012). And this is also supported by research findings. Within their review Boehm and Dwertmann (2015) analyzed 22 studies on the relationship between age diversity and productivity and found eight studies reporting a negative effect of increased age diversity on productivity, as well as 10 studies with a null effect. Only 3 out of 22 studies reported a positive

impact of age diversity. However, these studies do not test specific organizational practices as potential moderators, which have shown to be effective in many cases.

In order to face stereotypes in practice, a more thorough understanding of the moderating levers for successful employment of older employees is required. Qualitatively, Wöhrmann et al. (2018) began to collect and define organizational practices for later life work within the SWI, however, based on a German perspective so far. The SWI's structure consisting of organizational culture, leadership and specific HR practice elements is also supported by Boehm and Dwertmann's (2015) review, in which they hypothesized organizational conditions moderating the relationship between age diversity among the workforce and organizational productivity. First, they identified leadership including elements of transformational leadership, health-focused leadership, and top management leadership as an important moderator. Second, they proposed a moderating effect for an organizational climate that values age-diversity and inclusion. And third, they found a positive influence of "age-specific and age-inclusive human resource practices" that tailor human resource management towards older employees, while allowing for individuality within the employment conditions. Taken together, these three categories of organizational measures are theoretically reasoned and partially empirically proven to positively influence the effect of age diversity on organizational performance and hence provide a valuable framework for further research. Going beyond Boehm's and Dwertmann's (2015) focus on age diversity and performance, the SWI is intended to identify organizational practices positively impacting not only performance, but also on illness absence rates and employee fluctuation given an ageing workforce. Figure 1.1 outlines the organizational level model, in which the three categories of organizational measures are hypothesized to firstly impact organizational outcomes directly, and secondly to moderate the impact of increased age diversity and average age of the workforce on organizational outcomes.

Figure 1.1

Proposed moderating effect of organizational later life work practices on the effect of demographic exposure on organizational outcomes



Note. Based on Boehm & Dwertmann (2015).

Thereby, those measures would support organizations to achieve high levels of organizational results despite an older and more age diverse workforce.

The German Perspective: Silver Work Index (SWI)

So far, organizations lack profound tools to assess their readiness and capabilities to leverage older employees' potential. In order to cope with an ageing workforce and enable later life work successfully moderating organizational practices need to be firstly validated and secondly accessible in practice to be effective and reduce current stereotyping. Wöhrmann, Deller and Pundt (2018) hence proposed the Silver Work Index (SWI) integrating the most relevant organizational practices regarding later life work. As a diagnostic tool for organizations the SWI is intended to allow for assessment, comparison, and evaluation of organizational conditions identified as good practices regarding later life work. Thereby, organizations shall be enabled to firstly assess their individual areas for improvement among the index dimensions internally and secondly to benchmark results against peers on industry or regional level.

The specific organizational measures needed vary for example, between different industries, organizational size, and different age structures within the organization. In order to fully cover relevant perspectives the index dimensions and indicators were developed in an iterative process based on 27 expert interviews. Interviewees were researchers from various disciplines (demographics, economics, gerontology, human resources (HR) management, and psychology) as well as employees of retirement age, HR executives, HR managers, management consultants, executives of placement agencies for paid and voluntarily later life work, and representatives of strategic and operational management in various industries. All experts were either able to share personal experience regarding later life employment, or had dealt with later life employment as part of their job responsibilities or in research.

The interviewees were aged 35 to 83 years (M = 52.7; SD = 10.6). Most were male (74%). Experts had, on average, 28 years of professional experience (M = 28.4; SD = 11.4; range 7 to 59 years). The experts worked in various industries: 29% professional, scientific, and technical; 26% finance and insurance; 15% manufacturing; 15% administrative, support and other services; 11% human health and social work activities; and 4% information and communication. The industry heterogeneity of the sample was intended to cover a variety of viewpoints on the research topic.

The interviewees were asked for characteristics of good organizational management practices in order to successfully involve employees aged 60 and older as well as indicators and methods to measure the characteristics in practice. Supplementing the open question section the interviewees were systematically presented with additional aspects related to good organizational management practices concerning employees nearing retirement age and beyond that had earlier been identified through an analysis of the relevant body of literature, if not mentioned by the expert independently. These aspects were: perception of age/ageing; ways of structuring and designing the work/workplace; procedures in place for retiring or resigning; individual financial situation; methods of guiding older employees; available health

management and promotion resources; and the information on the range of possibilities for continued employment after retirement age. The interviews took 30 to 60 minutes and were audio-recorded, transcribed, and content analyzed by Wöhrmann et al. to derive the initial category system of organizational practices on later life work. To further sharpen the content analysis results the authors conducted an expert workshop with a subset of the original interviewees and further experts with the same backgrounds as the initial interviewees in 2015 leading to the published version of the SWI.

In line with several studies (Armstrong-Stassen, 2008; Armstrong-Stassen & Schlosser, 2011; Cheung & Wu, 2013; Hennekam & Herrbach, 2013) organizational culture and leadership were identified as the two overarching and most important dimensions for successful and motivational work up to and beyond the retirement age supported by several underlying dimensions for working conditions and processes. These were work design, health management, individual development, knowledge management, transition to retirement phase, and employment during retirement phase.

While providing a first version of the SWI Wöhrmann et al. (2018) also reported the need for further improvements to increase inter-coder reliability of the construct. Single index dimensions revealed some improvement potential in clarity of the definitions, wording, and distinction of the indicators. Moreover, the index was based on the German dataset described above, which allowed for a thorough identification of practices important for the German legislation and work culture, but lacked generalizability for other countries. Consequently, the SWI requires further revision and has to be checked against other work environments.

The U.S. Perspective: Age Smart Employer Award

An opportunity to test the SWI in an international context arose from data on organizational practices regarding later life work gathered during the 2014 or 2015 edition of the Age Smart Employer Award, a culture-change strategy to honor New York City businesses

whose practices engage and retain workers of all ages, with a specific focus on older workers (Finkelstein et al., 2013). The award was initiated by the Robert N. Butler Columbia Aging Center and The New York Academy of Medicine in 2012 as an extension of the Age Friendly NYC initiative, a public-private partnership to make New York City a better place to grow old. It honors New York City employers whose policies and practices promote generational diversity in the workforce and highlights the positive contributions of older workers.

The first Age Smart Employer Awardees were chosen from a pool of 20 applicants and honored in 2014. Following the initial cycle of the Awards, it was recognized that the literature-based Compendium of Strategies and Practices that formed the basis of the application and selection process focused almost exclusively on large business practices, while small businesses comprise 98% of all employers in New York City. Based on semi-structured interviews with more than 100 small business owners and 160 employees, as well as several expert interviews the compendium was hence complemented by five industry-specific guides comprising more specific small business practices to recruit, train, and retain older employees. Specific industries were selected based on prevalence of older workers in the sector in NYC, opportunities for older workers to solve business owners' perceived staffing problems (as identified in interviews), and lack of attention to older workforce issues in the sector (Finkelstein et al., 2013). Identified practices included recruitment strategies, training, job restructuring, work flexibility, benefits, and phased retirement.

Based on the practices identified, two different semi-structured free text application forms for the Awards were developed: one for larger businesses including all practices initially identified and one for smaller business based on conducted qualitative primary research. The questionnaires provided a range of domains to consider including recruitment, productivity and performance, engagement, retention of workers, and the transition to retirement, but also explicitly asked for practices not covered by these domains.

Organizations and businesses interviewed for the guides yielded the first tranche of applicants and promoters for the Awards. The outreach strategy included presentations to organizations, webinars, and personal networking. In 2015, due in part to the momentum built in the previous year, 52 businesses applied. The Awards have been well received by the business community and national press leading to interest from other U.S. cities to implement similar strategies.

Basis for this integration were practices identified within 61 New York based organizations that applied for the 2014 or 2015 edition of the award. The sample was widespread across industries and company sizes, so that a holistic view on organizational practices from a variety of viewpoints could be obtained. Among the 61 companies analyzed 23% were active in the food industry, 18% in health care, 15% in social service, 13% in services, 13% in manufacturing, 7% in entertainment, 7% in education, and 5% in retail. 22 companies (36%) were nonprofit and 27 (44%) were family-owned. The years in business differed from 3 to 261 years (M = 59.69; SD = 50.51). The number of employees ranged from 4 to 200.000 (M = 7,819.85; SD = 28,900.19).

The applications were content analyzed by a selection committee resulting in a categorized longlist of practices (see Appendix B for details). The committee included experts in workforce development, human resources, executive education, small business services, diversity, aging, and communications. After the committee identified ‘finalists,’ staff conducted employee surveys at each business, or NYC establishment of a larger firm, to validate the policies and practices described in the applications. The committee then used survey response rates, survey results, and the original applications to identify Award winners. The Award has been repeated in 2018 with a combined application and assessment for both, smaller and larger companies.

Integrating Both Perspectives

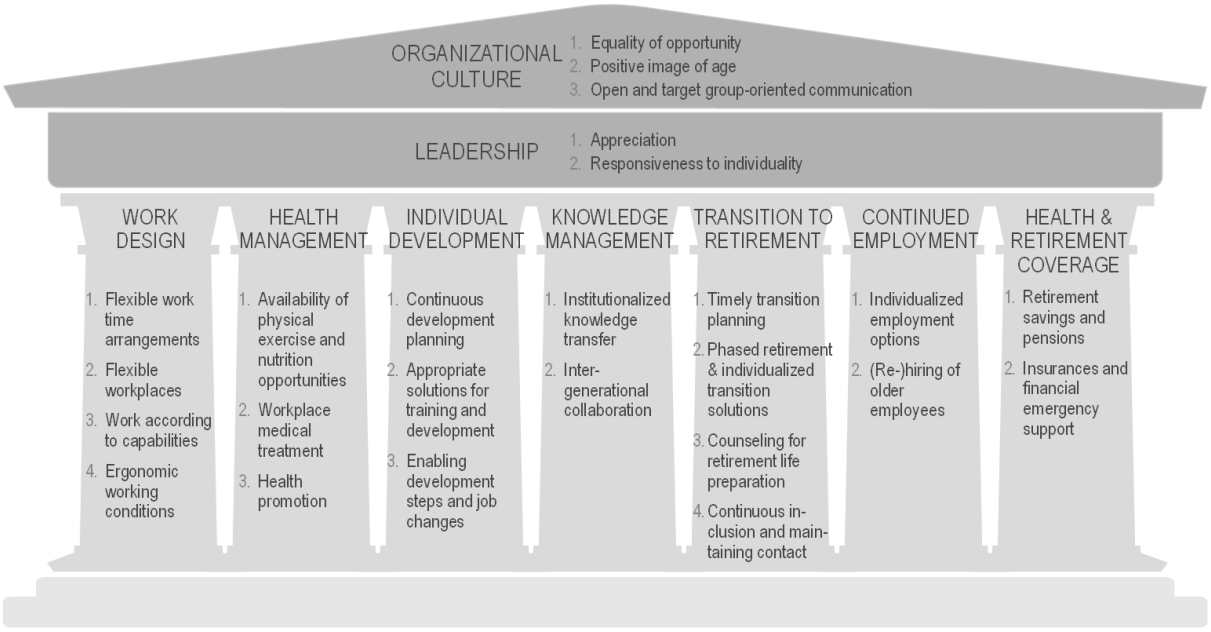
By means of integrating the German and the U.S. findings on organizational practices for later life work we intend to revise the initial SWI from an intercultural perspective and to further sharpen its definitions of dimensions and indicators. For this we joined in an international collaboration. Two of the authors contributed with extensive knowledge on the SWI as they were part of the initial SWI development in Germany. Two further authors contributed with the U.S. perspective as they were initiating and managing the Age Smart Employer Award.

For both the German and the U.S. perspective, the original empirical qualitative data was leveraged for the integration. In a first step, an independent person compared both category systems of organizational practices for later life work in order to identify similarities and differences. For this, experts involved in the Age Smart Employer Award were interviewed regarding the practices identified. Within the 90 minutes semi-structured interview each practice was explained in detail. The interview was transcribed for the comparison with the SWI. For the SWI practices the existing documentation was used for the comparison. Ambiguities for both category systems were clarified in direct conversations with the respective experts involved in the initial content analysis. The individual who was independent to the SWI and the Age Smart Employer Award then systematically identified differences between the practices explained in the interview and the SWI. This step resulted in a matching table highlighting corresponding practices, as well as practices with further alignment need.

In a second step, all five individuals met for a workshop in Germany to jointly discuss identified differences and to align on necessary changes to the SWI in order to reflect both, the German and the U.S. perspective. Following general alignment on required changes to the index, the indication and dimension definitions were iteratively derived in a process of back-and-forth translation between German and English and thereby further improved. Finally, all five individuals agreed to the final definitions of dimensions and indicators within a revised

Figure 1.2

Later Life Work Index (LLWI)



Note. Later Life Work Index (LLWI). Published with kind permission of © Max R. Wilckens, Anne M. Wöhrmann, Jürgen Deller 2019. All Rights Reserved

index of organizational practices for later life work as depicted in Figure 1.2. The revised index is named Later Life Work Index (LLWI). The definitions of dimensions and indicators can be obtained from Appendix A.

Intercoder reliabilities (see Krippendorff, 2013) were determined in order to validate the revised category system. Two individuals who were not involved in the project, but are knowledgeable on human resource management practices independently coded the 27 interviews from the SWI dataset, as well as the interview conducted with the two individuals involved in the Age Smart Employer Award. The revised category system, which resulted from the revision workshop, detailed by the aligned definitions was used as the rating instructions. The two coders were asked to assign each paragraph to all categories that reflect practices mentioned as being important for later life work. Krippendorff's alpha ranged from .65 to .92 for the nine dimensions, which reveals further opportunities for improvements, but can be considered acceptable given the multidimensionality of the category system and the high complexity and amount of rating material.

The revised LLWI differs from the SWI published by Wöhrmann et al. (2018) in two aspects: First, an additional dimension for Health and Retirement Coverage was added to the index. The dimension accounts for organizational support in terms of retirement savings and insurance coverage for the organizations employees, in case not sufficiently provided by public systems. Requirements vary due to different regulations and social systems. In Germany the support may be a direct financial benefit or put into practice as individual planning and assistance. Indicators are retirement savings and pensions, as well as insurances and financial emergency support. The retirement savings and pensions indicator covers offers to employees for retirement savings, if not sufficiently covered by public systems. Organizations may include pensions and retirement saving accounts into their full compensation packages, offer optional saving possibilities to be opened by the employees individually, and support their employees in timely planning and organization of their retirement savings. Insurances and financial emergency support describe offers for health related insurance coverage, if not sufficiently covered by public systems. This includes (additional) health-, disability-, care- or life insurances, which particularly cover risks that increase with age. Additional financial support may be offered in case of family emergencies, as e.g., in a case of nursing care or child sickness.

Secondly, several dimensions and indicators were sharpened and rephrased. Most importantly, the dimension Employment during Retirement Phase has been renamed to Continued Employment and its indicators have been restructured. This was necessary to account for the more flexible retirement age in the U.S., so that it describes the offer of employment options more generally for all employees that would have already been retired in their former job. This includes former employees of the organization as well as external employees looking for continued employment. Rephrased indicators are individualized employment options and (re-)hiring of older employees. The individualized employment options indicator reflects the positive effect of employment opportunities for individuals, who would otherwise be fully retired. To ensure employment options are meaningful for both the

organization and the employee, integration of those employees into the organization should be strategically planned and systematically framed. For example, organizations might define areas and activities suited for continued employment, for which employees might be brought in on a temporary basis at peak workload times. Tasks, working conditions and work time should be adaptable to the individual employee. This can be achieved through alternative contract forms such as consulting or mentoring activities, work on specific projects, or holiday replacement. The majority of the experts held the opinion that the arrangements should be temporary and should involve fewer hours than a fulltime position. The (re-)hiring of older employees indicator reflects a finding especially from the U.S. data. Older individuals, particularly including already and almost retired employees should be specifically addressed by job marketing, hiring and re-employment processes. This is achieved through age-friendly, open and transparent communication of job offers and the use of alternative marketing paths to address external as well as internal individuals. This explicitly includes employees with long careers in other industries or companies.

In other dimensions additional examples were added to include the U.S. perspective. For example, tuition reimbursement and apprenticeships also for older employees were included in the Individual Development dimension as they had not been identified as relevant practices within the German version. Moreover, ‘phased retirement’ and ‘job changes’ were emphasized by explicitly stating them in the indicator titles.

Overall, the revision led to an index of organizational later life work practices that are applicable in both, Germany’s rather regulated work environment with a strong public social system, as well as the rather free market driven U.S. work environment. Although the index does now account for both perspectives, identified differences showed that practices are not equally important and that it is hence required to attach country-specific importance to specific dimensions.

Discussion

This integration provides an additional incremental step towards a well-founded index for good organizational management of later life work. Prior to operationalization and quantitative validation of the LLWI, this study adds an intercultural revision based on the comparison of the initial Germany based qualitative study and a second U.S. dataset from the New York Age Smart Employer Award.

Overall, both perspectives on organizational practices for later life work were very similar. Both datasets show that good organizational management of older employees is a complex, multidimensional subject including both, environmental dimensions as organizational culture and leadership, as well as specific practices regarding work conditions and arrangements. Results confirm that many aspects of good organizational management of employees of this particular age group are not specific to older adults. However, work design and health management for example are dimensions of growing importance with age, while other as knowledge management and individual development need to be tailored towards older employees. Thus, retirement practices are relevant for older employees only.

Findings support the emphasis put on appreciative and individualized leadership as one overarching dimension. As already identified by the initial SWI ‘feeling valued’ has also been identified as one of the most important aspects in U.S. based organizations. Having a ‘family-like’ environment with a “leadership style that makes it seem like it’s not a leadership style” , flexibility in designing work conditions and the willingness to consider individual needs are the most important levers for retaining employees in the organization who are nearing retirement age or older. Moreover, both datasets emphasize the importance of providing the possibilities of reducing work hours and phased retirement, changing the focus of the work content, as for example in projects or by job-rotation.

However, two divergent aspects have been identified by the collaboration between Germany and the U.S.: First, differences between the social security systems between the two

countries result in diverging importance of financial benefits granted by the organizations. While German experts attached only little importance to organizational health and retirement coverage, the U.S. data emphasized that taking care of ageing employees within the organization also in a financial manner is a crucial part of good organizational management of employees nearing retirement age and beyond.

Second, the meaning of 'being retired' differs given the two cultural and social system backgrounds. In Germany the national-wide 'retirement age' will be continuously increased from the age of 65 to 67 in 2029. By that threshold age employees receive public pension, so that being retired at that age is common practice in most employment frame contract agreements. While the first version of the SWI was hence focused on regular employment up to that age and 'during retirement age', dimensions and indicators of the LLWI have now been broadened to reflect continued employment independently of a particular retirement age and explicitly include practices to (re)hire older employees.

Implications

Despite a general improvement and sharpening of dimensions and indicators within the index, this study shows that small amendments qualify the revised index for application beyond Germany. Once operationalized, the LLWI will allow research to holistically compare moderating practices on the effect between workforce demography and organizational outcomes. From a business perspective, it will be important to understand the relative impacts of moderating factors to ensure focused allocation of resources and efforts. By integrating the German and the U.S. perspective, the LLWI may serve as a foundation for further inter-country comparisons of organizational practices for later life work. Certainly, further validations in different work cultures and legal frameworks are required and might potentially match as well as it was possible for Germany and the U.S. However, striving for a common basis for later life work practices simplifies further comparisons.

From a practice perspective, the LLWI will allow organizations to access their own capabilities in terms of practices regarding the employment of older employees. Organizations facing the challenge of an ageing workforce either due to labor force shortages or political requirements may use the LLWI as a diagnostic tool to identify improvement opportunities for the management of older employees. By allowing for industry and regional benchmarks, the LLWI has the potential to raise awareness for later life work not only in business, but also in the political debate. Benchmark results can support the identification of best-practices and thereby support the organizations' management to handle ageing issues. Thereby the importance of each dimension within the entire index might differ from industry to industry, region to region and country to country. Specific recommendations on effective practices will be the results. Within the proposed theoretical moderation model for the effect of workforce age on organizational outcomes, it has hence to be expected that different organizational circumstances require certain combinations of index dimensions as moderating levers to be pulled in order to drive organizational outcomes.

Limitations and Future Research

For the German dataset the selection of a heterogeneous group of interviewees ensured a variety of viewpoints regarding the subject (Wöhrmann et al., 2018). For the U.S. dataset the significant public attention for the Age Smart Employer Award leads to the conclusion that applicants belong to the forefront of good management practices for employees nearing retirement and beyond. Heterogeneity regarding size and industry of the organizations supports the assumption that derived management practices are exhaustive. However, it has to be noted that the U.S. data originates from establishments in the City of New York only.

Data analysis was done systematically. The comparison of practices identified in the U.S. data and the dimensions and indicators of the LLWI were determined in a data-oriented manner and were therefore generated empirically. It can hence be assumed that the revised

dimensions and indicators of the construct are content-valid. However, this study also proves that country-specific differences due to differences in laws and regulations exists and that future studies need to test required amendments for additional countries. Reliability has significantly improved since the first version of the index, but needs to be rechecked in case of further revisions.

Going forward, operationalization and validation of the LLWI in a cross-organizational study is intended. Assessment is planned as a combination of fact based questions on an organizational level (e.g., share of employees offered / working in phased retirement) and questions on culture and leadership in employee samples (e.g., the image of age, equality of opportunity, appreciation). The external validation of the index shall link the index to employee and organizational level outcomes.

2

Assessment of the Reliability of Multi-Valued Data: An Analysis of Krippendorff's Alpha Coefficient

Max R. Wilckens

Abstract

This article discusses Krippendorff's alpha for reliability assessment of multi-valued data from qualitative sources. Multi-valued data exists whenever a phenomenon is described by more than one unidimensional value. Krippendorff's alpha is the first agreement coefficient extended to multi-valued data. It allows qualitative researchers to compute chance-corrected inter-coder agreement for multi-valued data. Monte-Carlo simulation of datasets with up to seven individual binary values and two coders supported the need for multi-valued reliability coefficients. Though, results suggested that Krippendorff's alpha varies depending on whether disagreement occurs in units, for which most of the individual values are present, or in units, for which most of the individual values are absent. For most datasets, the variation appears to be undesired, when assessing reliability. To avoid the variation, this paper proposes a modified version of Krippendorff's alpha for multi-valued data, which is based on the set-theoretic difference metric. Both versions are implemented in R code to ease access for qualitative researchers, who aim to test and proof reliability of their multi-valued data extraction from qualitative sources.

Keywords: content analysis, Krippendorff's alpha, multi-valued data, qualitative research, reliability

Reliability is an inevitable prerequisite for valid findings. Scholars have to ensure reliability of their research; that is, confidence in the extraction of meaning from data on the phenomenon researched. A finding on a stable phenomenon is reliable, if it persists even if gathered at different points in time, by different instruments or by different individuals (Krippendorff, 2011). In qualitative research, scholars extract findings from qualitative sources as texts, videos, images, and observations by coding the content (Krippendorff, 2019). The coding process has to be objective and reproducible to allow reliable conclusions. Agreement coefficients such as Krippendorff's alpha (1970, 2004), Cohen's kappa (1960), or Scott's pi (1955) allow researchers to proof the reliability of their coding process by testing the consistency of the coding results across two or more independent coders (Lacy et al., 2015). However, the predominant agreement coefficients are limited to a single nominal, ordinal, interval, or ratio coding scheme. Many real-world phenomena are multi-valued and are best characterized by sets of values. For example, images may depict more than one motive, objects may be multi-colored, sound patterns in audio data may be overlapping, and medical x-ray scans may show more than one medial diagnose on the same scan. The traditional agreement coefficients cannot deal with multi-valued data, so that reliability assessment for qualitative studies was limited to single-valued phenomena.

Krippendorff (2019) addressed the issue and proposed a version of his alpha coefficient that allows for multi-valued data. However, the proposed coefficient for multi-valued data has not been tested empirically so far. The present study uses simulated data to investigate the statistical properties of Krippendorff's alpha coefficient for multi-valued data and contributes to the body of existing literature in several ways. First, by examining the statistical properties of Krippendorff's alpha for multi-valued data, the article provides strengths and weaknesses of Krippendorff's alpha coefficient for multi-valued data. Researchers get to know the specificities of the alpha coefficient for multi-valued data, which is important to conclude on the reliability

of data from the coefficient. Second, by providing a modification of the alpha coefficient for multi-valued data based on the set-theoretic difference metric, the paper advances our methodological repertoire to compute agreement coefficients for multi-valued data. Although Krippendorff's multi-valued alpha coefficient (2019) is a milestone for reliability analysis of multi-valued data, this paper's amendment to the coefficient based on a set-theoretic difference metric is better suited for many applications in research and practice. Third, this article contains an implementation of the multi-valued alpha coefficient, which provides easy access to multi-valued reliability assessment. Researchers can use the implementation to compute chance-corrected agreement for their multi-valued data. Thus, this article may contribute to more reliable conclusions from data, which does not fit single-valued agreement coefficients. Reliability of conclusions from these data has most likely not been proven sufficiently so far. Consequently, this articles' analyses and methodological advancements may contribute to researchers' ability to correctly and easily reassure reliability of their conclusions in more complex multi-valued research designs and the corresponding data.

Coefficients Assessing Data Reliability

Qualitative researchers have debated, to what extend and how reliability has to be proven in qualitative studies. Methodologists developed processes and standards for good qualitative research (e.g., Gehman et al., 2018), which guide researchers to reproducible, objective conclusions. Some argued that following these processes and standards thoroughly and conscientiously is sufficient to achieve reliable results. Consequently, these researchers opposed agreement coefficients providing an empirical proof of reliability as superfluous and as a surrender towards quantitative research methodologies. However, most of today's high-quality publications presenting qualitative studies reassure reliability of the conclusions by some kind of agreement coefficient. Also, often cited methodologists such as Gioia et al. (2013)

recommended to compute intercoder agreement “to bolster [the authors’] own confidence in […] assertions and findings” (p. 22).

Assessment of reliability from data builds upon multiple different data points characterizing the same phenomenon for the same unit of analysis. Agreement of these multiple data points with each other increases confidence in the meaning extracted from the data. The different data points usually originate from different methods, different points in time, or different individuals. For example, researchers use scales with multiple items to assess the same phenomenon in quantitative research (DeVellis, 2017) and compute inter-item correlation with the Cronbach’s alpha coefficient (Cronbach, 1951) to assess agreement among the items. In qualitative research, data often appears from observations or as artifacts, such as texts, images, or videos. To obtain multiple data points for the assessment of reliability, multiple individuals code the same artifacts or observations resulting in multiple data points characterizing each artefact or observation. Krippendorff (2019) refers to this secondary data as “reliability data”. The reliability data allows researchers to compute the level of inter-coder agreement and, thus, an indication for the reliability of the extraction of meaning from the original qualitative data.

Research broad-up a large number of agreement coefficients to assess reliability (e.g., Hayes & Krippendorff, 2007; Hsu & Field, 2003). The coefficients differ by three basic characteristics: First, coefficients differ in how differences between data points are assessed. For example, internal consistency coefficients, such as Cronbach’s alpha assess differences based on correlation. Inter-coder agreement coefficients, such as Cohen’s kappa (1960) or Krippendorff’s alpha (1970, 2004) assess differences with a difference function on a per unit basis. Second, some of the coefficients correct the observed agreement between the data points for the expected agreement that would occur, if the data is random. Other coefficients, such as percent agreement do not correct for the expected change-agreement between coders. Chance-corrected coefficients have been criticized, because they may show very low values despite high percent agreement, if the distribution is skewed; that is, a high share of the coded units fall

into one of the categories as for example the case, for rare diseases among a population, for which the vast majority of subjects fall into the category ‘not affected’ (Feng, 2013; Gwet, 2008). Yet, Krippendorff (2011) argued that the lack of variability of those values is a more general issue of the measurement itself. He suggested in line with Tinsley and Weiss (2000) to use stratified samples for the reliability assessment, so that each value shows sufficient variation. Third, those coefficients that correct for expected change agreement differ in the assumption underlying the estimation of the expected agreement. For example, Cohen’s kappa (1960) estimates the change for agreement based on each coders’ individual distribution of codes. Contrarily, Scott’s pi (1955), its generalization to more than two coders by Fleiss (1971), and Krippendorff’s alpha (1970, 2004) base the estimation on the joined distribution of all coders. Scott’s pi and Fleiss’ kappa draw chance agreement for each unit with replacement, while Krippendorff’s alpha assumes change agreement between coders without replacement. The differences between coefficients have widely been discussed (e.g., Hayes & Krippendorff, 2007; Hsu & Field, 2003; Krippendorff, 2011) and exceed the scope of this article.

Krippendorff’s alpha has been argued to be the most general coefficient among chance-corrected agreement coefficients, because it equals other coefficients for special cases, such as a very large number of data units (Krippendorff, 1970). Moreover, it is independent of the number of raters, the number of units, and the number of different values assigned (Hayes & Krippendorff, 2007). Krippendorff’s alpha is based on the common form of chance-corrected agreement coefficients:

$$Agreement\ coefficient = \frac{P_o - P_e}{1 - P_e} = 1 - \frac{1 - P_o}{1 - P_e} = 1 - \frac{D_o}{D_e},$$

where P_o and P_e denote the observed and expected agreement, and D_o and D_e the observed and expected disagreement among values assigned to units. Krippendorff (2019) defines the disagreement measures for his alpha coefficient as:

$$D_o = \sum_c \sum_k o_{ck}\, {}_{metric}\delta^2_{ck}$$

with observed coincidences $o_{ck} = \sum_u \frac{1}{m_u - 1} \begin{cases} n_{uc}(n_{uk} - 1) & iff\ c = k \\ n_{uc} n_{uk} & iff\ c \neq k \end{cases}$, and

$$D_e = \sum_c \sum_k e_{ck}\, {}_{metric}\delta^2_{ck}$$

with expected coincidences $e_{ck} = \frac{1}{n_{..} - 1} \begin{cases} n_{\cdot c}\,(n_{\cdot k} - 1) & iff\ c = k \\ n_{\cdot c}\, n_{\cdot k} & iff\ c \neq k \end{cases}$,

where ${}_{metric}\delta^2_{ck}$ denotes the difference between values *c* and *k*, m_u the number of values assigned to unit *u* (i.e., the number of coders that coded unit *u*), and n_{uc} and n_{uk} the frequency unit *u* is assigned value *c* and *k*, respectively. The difference between two values can be calculated for nominal, ordinal, interval, and ratio data. Each difference function ranges between zero and one and equals zero, if *c* is equal to *k*. For nominal data the difference function is:

$${}_{nominal}\delta^2_{ck} = \begin{cases} 0 & iff\ c = k \\ 1 & iff\ c \neq k \end{cases}.$$

Krippendorff's alpha can be assessed for any number of coders and any number of units. It has been argued to be rather independent from sample size (Hayes & Krippendorff, 2007). Moreover, it has been detailed for multiple common use cases in research, such as the unitization of continuous data, agreement with a given standard, and recently, for multi-valued data (Krippendorff, 2019).

Occurrences of Multi-Valued Data

Krippendorff (2019) proposed a generalization of his α-family of agreement coefficients to multi-valued data. Compared to single-valued data, in multi-valued data each unit of analysis is characterized by multiple values. For example in the case of three nominal values, namely the colors *green*, *yellow*, and *red*, apples as the units of analysis may each either be assigned one of the three colors to suit calculation of any of the standard agreement measures including the single-valued alpha by Krippendorff. However, apples may be multi-colored. Multi-valued data characterizes units of analysis by multiple values. To stick to the example,

an apple may be assigned the values *red* and *yellow*, because one side might be red, while the other is yellow.

Moreover, Krippendorff (2019) states two further use cases for multi-valued data. First, multi-valued data may result from multiple valid perspectives on a single-valued phenomenon. For example, a law suit may be evaluated from the victims, the offender's, and the jury's perspective. All three parties may come to different, yet valid assessments. Taken together, the three perspectives result in a multi-valued characterization of the law suit, even though the parties' individual assessment may be single-valued. Second, if the units of analysis are not clearly defined in continuous data, such as long texts, videos, or observations, coders may overlap in their evaluations of interpretable units, resulting in multi-valued data. For example, two coders rating a video-clip on a continuous basis may overlap in their ratings. Coder *A* may code the first three minutes of the video-clip with category *a* and minutes four to six with category *b*. If coder *B* codes the first four minutes of the same video-clip with category *a* and minutes five to six with category *b*, minute three to four is coded with categories *a* and *b*. The overlap results in multi-valued data. Multi-valued computations are required even though the category system for each coder is single-valued.

Clearly, the possibility to research reliability of multi-valued data might mislead researchers to not sufficiently differentiate and define values. Particularly during the design of coding tasks to obtain reliability data, researchers may perceive complex phenomena as multi-valued, even though they might be unidimensional and just not sufficiently understood yet. Thus, researchers should always strive for single-valued solutions to ensure that multi-valued assignments are inevitable necessary in the remaining cases.

Since chance-corrected agreement measures were not available for multi-valued data, researchers (including myself) limited themselves to single-valued category systems when evaluating phenomena. I personally caught myself applying agreement coefficients for single-valued data to individual values of multi-valued data. To stick to the above example: we

calculated three independent levels of agreement between coders for whether apples are *red* or *not red*, *yellow* or *not yellow*, and *green* or *not green*, although those values had been assigned dependently by the same coders at the same time. Thus, the expected agreement should not be based on a single value's distribution. Researchers should either use Krippendorff's alpha for multi-valued data or limit their category system by for example asking the coders to select the most applicable value for each unit.

Krippendorff's Alpha for Multi-Valued Data

To assess chance-corrected agreement among multi-valued data, Krippendorff (2019) computed alpha based on the sets of values present in the data. Each set is treated similarly as a single value in the case of single-valued data. The expected agreement results from the joint distribution of sets across the coders. Thus, the multi-valued agreement coefficient only makes sense, if the individual values composing the sets are dependently allocated to units, jointly as a set or sequentially, yet interdependently. There may be other cases of multi-valued data, in which the values are independently allocated to units. Then, a set-based estimation of the expected agreement is misleading.

The key difference between alpha for single-valued data and alpha for multi-valued data is the difference function, which has to compute differences between sets of values in the multi-valued case. Krippendorff used the following difference function to compare the sets *C* with values c_1 to $c_{|C|}$ and *K* with values k_1 to $k_{|C|}$ assigned to a unit:

$$_{metric}\Delta_{CK} = \sum_{c \in C} \sum_{k \in K} {}_{metric}\delta^2_{ck} - \sum_{c \in C \cap K} \sum_{k \in K \cap C} {}_{metric}\delta^2_{ck}.$$

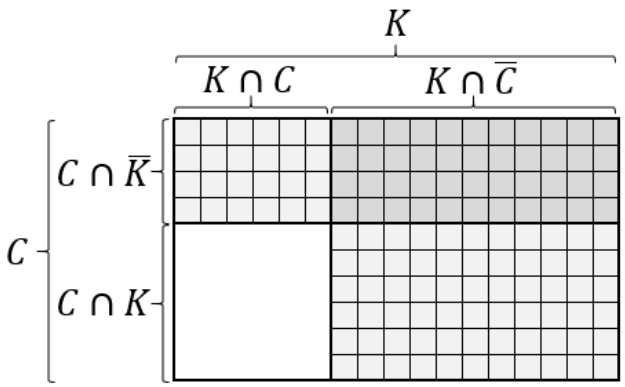
With a nominal difference metric for individual values $_{nominal}\delta^2_{ck} = \begin{cases} 1 & iff\ c \neq k \\ 0 & iff\ c = k \end{cases}$ the nominal set difference metric becomes:

$$_{nominal}\Delta_{CK} = |C||K| - |C \cap K|^2,$$

where $|C|$ and $|K|$ denote the number of individual values in *C* and *K*, $|C||K|$ the number of possible *c*-*k* pairs of values between *C* and *K*, and $|C \cap K|^2$ the number of *c*-*k* pairs of values

Figure 2.1

Visualization of the difference function for two sets of values in multi-valued data.



Note. Each row denotes one value of set *C* and each column one value of set *K*. (Krippendorff, 2019)

that *C* and *K* share. Krippendorff treated empty sets { } as a specific set and defined $|\{\}| = 1$, $|\{\} \cap \{\}| = 1$, and $|\{\} \cap K| = 0$. He visualized the components of the difference function in Figure 2.1. Notably, the function implements a squared difference measure, which's maximum equals the number of possible *c*-*k* pairs of values between *C* and *K*. Thus, the observed and expected disagreement measures are adjusted with a denominator accordingly:

$$_cD_o = \frac{\sum_C \sum_K o_{CK}\ _{metric}\Delta_{CK}}{\sum_C \sum_K o_{CK}|C||K|}$$

with coincidence $o_{CK} = \sum_u \frac{1}{m_u - 1} \begin{cases} n_{uC}(n_{uK} - 1) & iff\ C = K \\ n_{uC} n_{uK} & iff\ C \neq K \end{cases}$, and

$$_cD_e = \frac{\sum_C \sum_K n_C\ n_K\ _{metric}\Delta_{CK}}{\sum_C n_C|C|\ (\sum_C n_C|C| - 1)}.$$

The alpha coefficient for multi-valued data then results from observed and expected disagreement similarly as for single-valued data:

$$_c\alpha_{nominal} = 1 - \frac{_cD_o}{_cD_e} = 1 - \frac{\sum_C n_C|C|\ (\sum_C n_C|C| - 1)}{\sum_C \sum_K o_{CK}|C||K|} \cdot \frac{\sum_C \sum_K o_{CK\ nominal}\Delta_{CK}}{\sum_C \sum_K n_C\ n_{K\ nominal}\Delta_{CK}}.$$

Compared to a former approach by Krippendorff and Craggs (2016) that used the Dice coefficient as the difference measure for nominal sets, the outlined approach by Krippendorff (2019) is advantageous, because it measures differences in absolute terms. The Dice coefficient defined as

$$Dice\ Coefficient = \frac{|C \cap \bar{K}| + |\bar{C} \cap K|}{|C| + |K|} = 1 - \frac{2\,|C \cap K|}{|C| + |K|}$$

with $\bar{K}$ and $\bar{C}$ denoting the complements of *K* and *C*, measures the standardized difference on a scale between zero and one, but not the absolute difference. For example, the coefficient equals one (i.e., maximum difference), if set *C* consists of a single value *a*, and set *K* consists of a different single value *b*. But the coefficient also equals one, if set *C* consists of five values *a, b, c, d, e,* and set *K* consists of five different values *f, g, h, i, j*. In absolute terms, the latter case is more different than the first, which is accounted for in Krippendorff's revised difference function for multi-valued alpha.

Nevertheless, the difference function accounting for differences in absolute terms also has its specificities. Due to the squared nature of Krippendorff's (2019) difference metric for sets (see Figure 2.1), the maximum possible contribution to the observed disagreement by each unit depends on the product of set sizes $|C|$ and $|K|$. The additional denominator that has been added compared to alpha for single-valued data though corrects for the expected effect. However, the coefficient for multi-valued data still treats observed disagreement between two assigned values more severe, if the disagreement occurs in larger sets. For example, two sets $C = \{a\}$ and $K = \{b, c\}$ with no shared value yield a nominal difference of ${}_{nominal}\Delta_{CK} = 2$. Adding an additional shared value to both sets ($C = \{a, d\}$; $K = \{b, c, d\}$) increases the difference metric to ${}_{nominal}\Delta_{CK} = 5$, although the additional value occurred in both sets (i.e., the coders agree on this additional value). Disagreement between individual values is thus hypothesized to lower alpha stronger, if present in larger sets. The severity of this effect on the overall alpha has not been studied empirically so far.

Following this hypothesis, disagreement among values in sets larger than average reduces alpha over-proportionally. This also implies that alpha changes depending on how symmetric values are coded. Symmetric values describe binary attributes, for which assignment of the attribute is as informative as not assigning the attribute (Tan et al., 2004). For example,

gender is generally seen as a symmetric attribute, because knowing that a person is not female is as informative as knowing that the person is male. For those symmetric values, a multi-valued reliability coefficient should be independent of whether *male* is coded 1 and *female* is coded 0, or the other way around. For Krippendorff's multi-valued alpha this is not necessarily the case.

However, tabulating the differences between two sets of values in a squared difference function also has advantages: Since the difference function compares each value of set *C* with each value of set *K*, the differences can be allocated to the sets' individual values (klaus.krippendorff@asc.upenn.edu, personal communication: 10/17/2017). Thus, it is possible to compute the alpha agreement coefficient for an individual value, while accounting for its occurrence in sets for the calculation of the expected agreement. This allows researchers to gain additional information on the reliability of individual values within their multi-valued data.

Krippendorff's alpha for multi-valued data allows researchers to assess the reliability of multi-valued data. However, its statistical properties have not yet been investigated. To understand, how Krippendorff's alpha for multi-valued data can enhance science, it is important to understand, how the multi-valued coefficient relates to the data's characteristics and how it deviates from the average single-valued alpha of the individual values composing the multi-valued dataset.

Resulting from the review of Krippendorff's alpha for multi-valued data, three objectives emerged: First, we hypothesize that Krippendorff's alpha for multi-valued data deviates from the average single-valued alpha for the individual values composing the multi-valued dataset. However, we exploratively question to what extend and under what conditions Krippendorff's multi-valued agreement coefficient differs from its single-valued version. Second, we examine, which characteristics of the data influence Krippendorff's multi-valued agreement coefficient. In particular, we hypothesize (a) that the set size of multi-valued data influences Krippendorff's multi-valued alpha and, (b) that Krippendorff's multi-valued alpha varies depending on how symmetric values are coded. Third, based on our analysis of the

characteristics of Krippendorff's multi-valued alpha, we amend the multi-valued alpha by an alternative difference metric for disagreements between sets. We test, whether our amendment's is able to overcome the shortcomings of Krippendorff's multi-valued alpha and to more accurately assess reliability for certain multi-valued datasets.

Method

Procedure

To answer outlined research questions, Monte Carlo simulation was used. The approach followed recommendations for Monte Carlo experiments by Paxton et al. (2001) and was similar to the approach used by Feng (2013). Simulation of the sample data took place in a three-step process. First, combinations of four different sample sizes (50, 100, 500, and 2000), and four different numbers of individual values (2, 3, 5, and 7) were tabulated. The two parameters resulted in 16 combinations. Second, for each combination 300 sets of probabilities (one probability for each of the 2, 3, 5, or 7 values) were randomly generated. Probabilities ranged between 0.001 and 0.999. Third, binominal data for two coders and the number of individual values as defined for the respective combinations was simulated based on the defined sets of probabilities. For each coder the binominal data can be interpreted as if the coder would assign (1) or not assign (0) a certain value of the two, five, or seven nominal values to a certain unit of the sample. Thus, the simulated data's dimensionality was two coders times the number of values times the sample size. For each of the 300 runs, the third step was replicated 50 times, preserving the exact same prevalence and percentage of agreement between the two coders for each value. However, since the allocation of the values to the individual units changed among the replications, the joint distribution of multi-valued sets and in particular the set size varied among the 50 replications. Thereby, it became possible to investigate the influence of the multi-valued data's characteristics on alpha without confounding effects of varying agreement for

individual values. In total, the dataset comprised 240,000 simulation runs with 50 to 2000 multi-valued units each.

Analytic Strategy

To characterize the data of each simulation run, the parameters used to simulate the data (number of units, individual values, and raters) were used. Moreover, several coefficients were calculated for each simulation run: Prevalence describes how common a certain characteristic (i.e., an individual value within multi-valued sets) is within a population. Since Krippendorff based chance agreement of his alpha coefficient on the joint distribution of all coders, prevalence for each individual value was calculated across the two coders. Prevalence can only be calculated for individual binary values. Thus, aggregated coefficients were computed to validate independence of the multi-valued alpha from prevalence. First, the prevalence of each of the two, five, or seven individual values in each array was averaged to the mean prevalence per simulation run. Similarly, the prevalence standard deviation captured the variability among the prevalence of the individual values. Second, the mean set size denoted the number of values assigned to a certain unit of the sample on average.

Percent agreement measured the share of units, for which the two coders agreed to assign a certain value or agreed to not assign the value. Again, percent agreement is tied to a single individual value. Thus, a mean percent agreement measure and a percent agreement standard deviation measure aggregated the percent agreement for all values within each simulated multi-valued dataset.

To capture the relationship between the individual values of each simulated dataset, a correlation table for the two, three, five, or seven individual values was computed. The average correlation among individual values was used to describe the interdependency between the individual values of each dataset.

Further, the number of sets denoted the number of combinations of individual values that occurred in the simulation run. For the simulation runs with seven values, a maximum number of $2^7 = 128$ unique sets was theoretically possible (including the empty set).

To assess, how agreement was distributed within the simulation run, the number of values per unit, for which the two coders agreed, was correlated with the average set size per unit. This agreement set size correlation was high for simulation runs, in which agreement occurred among values in large sets. Krippendorff's alpha for multi-valued data treats the empty set (i.e., when no values are assigned to a unit by the coder) similarly as a set with a single value with set size of one. To compute the correlation, the set size of empty sets was hence set to one.

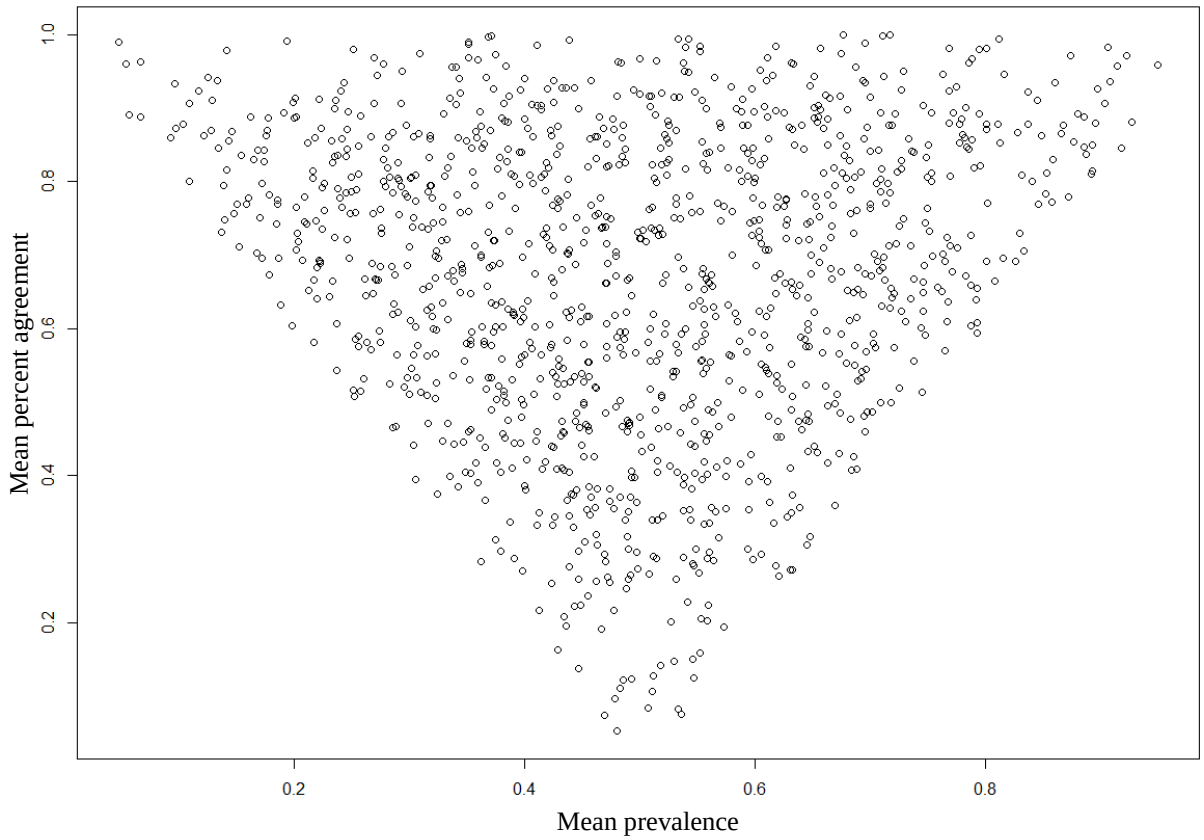
The subject of analysis for this study was Krippendorff's alpha for multi-valued data and the deducted alpha for individual values within multi-valued data (Krippendorff, 2019).

The analysis followed a three-step process: First, descriptive statistics provided insights into correlations among computed coefficients. Second, multi-variate linear regression analysis yielded dependencies of Krippendorff's alpha for multi-valued data on characteristics of the reliability data, such as prevalence and set size. Third, a random intercept model was used to investigate dependencies between alpha and further data characteristics (Gałecki & Burzykowski, 2013). The random intercept model allowed to analyze dependencies on top of the marginal effects of all individual values' prevalence and percent agreement, for example the distribution of disagreements.

The analyses were conducted with the open-source statistical software R 3.6 (Bates et al., 2015; http://www.R-project.org/) and several additional packages, for parallel computing (parallel), computation of Krippendorff's alpha (irr), data transformation (Rfast, dplyr), fitting of multi-level linear models (lme4, Bates et al., 2015), and visualization of results (ggplot2). To compute Krippendorff's alpha for multi-valued nominal data, a precompiled package was not available. Thus, a custom function was programmed and extensively tested. The code can be obtained from Appendix C.

Figure 2.2

Distribution of simulation runs



Note. Simulation runs are well distributed across different levels of mean prevalence and mean percent agreement. Percent agreement has a theoretical lower bound at $|1 - 2 \cdot prevalence|$, because for distributions with $prevalence \neq .5$ two coders must agree for some units.

Results

Descriptive Analysis

The simulated data was well-distributed. The prevalence of the individual values ranged from .001 to .999 (M = .50; SD = .31). The percent agreement ranged from zero to one (M = .67; SD = .27) and covered the theoretically feasible range of $[|1 - 2 \cdot prevalence|; 1]$. Similarly, the mean prevalence of the individual values in each simulation run ranged from .05 to .95 (M = .50; SD = .23) and mean percent agreement from .05 to 1.00 (M = .76; SD = .20). Figure 2.2 provides a graphical representation of the mean prevalence and agreement within the dataset. Correlation results can be obtained from Table 2.1 for the level of sets of multi-

Table 2.1

Means, standard deviations, and correlations

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
1. Number of units (50, 100, 500, 2000)	662.50	791.66											
2. Number of individual values (2, 3, 5, 7)	4.25	1.92	.00										
3. Average prevalence	0.50	0.18	.00	.01**									
4. Prevalence standard deviation	0.16	0.09	.00	.20**	.02**								
5. Average percent agreement	0.76	0.20	.00	.01**	.01**	.40**							
6. Percent agreement standard deviation	0.19	0.11	.00	.17**	-.02**	.04**	-.40**						
7. Average correlation among ind. values	0.06	0.18	.00	.02**	.00	-.06**	.04**	-.04**					
8. Average set size	2.11	1.20	.00	.80**	.55**	.18**	.00*	.13**	.02**				
9. Number of unique sets	26.89	30.75	.19**	.84**	.01**	.08**	-.10**	.18**	-.02**	.67**			
10. Agreement set size correlation	0.00	0.14	.00	-.03**	.72**	-.09**	.02**	-.02**	.20**	.35**	-.01**		
11. Krippendorff's alpha for multi-valued data	0.17	0.40	.00	.02**	.00	.34**	.83**	-.52**	.08**	.01**	-.08**	.04**	
12. Multi-valued alpha with set-theoretic difference metric	0.18	0.40	-.01**	.02*	.03**	.35**	.94**	-.50**	.06**	.03**	-.08**	.00	.99**

Note. N = 240000 simulated distributions of sets of two to seven individual values for two coders each. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

valued data. The number of units, the number of individual values, the mean prevalence, and the mean percent agreement parameters were independent from each other (r = [-.01; .01]). Krippendorff's alpha for multi-valued data was independent from the number of sampled units and the mean prevalence among the individual values, and almost independent from the number of individual values within the data (r = .02) and the average set size assigned to each unit within the simulation runs (r = .01). The number of unique sets, which were simulated in each run, was negatively correlated with the average percent agreement (r = -.10), because an additional disagreement, particularly among large sets, is likely to result in an additional set, which would not be present in the data without that disagreement. Thus, the number of unique sets was also negatively correlated with the multi-valued alpha (r = -.08). Finally, the multi-valued alpha was strongly correlated with the mean percent agreement (r = .83).

Table 2.2

Linear regression. Dependent variable: difference between Krippendorff's multi-valued alpha and the average alpha coefficient for each individual value

Variable	*β*	*SE*	*t*
(Intercept)	0.00	0.00	0.263
Units	0.00 *	0.00	-1.981
Mean prevalence	-0.20 ***	0.00	-70.842
Prevalence standard deviation	-0.01 **	0.00	-2.940
Mean percent agreement	-0.39 ***	0.00	-210.319
Percent agreement standard deviation	-0.70 ***	0.00	-415.635
Correlation between categories	0.19 ***	0.00	124.034
Set size	0.01 ***	0.00	3.641
Unique sets	0.09 ***	0.00	35.136
Agreement set size correlation	0.15 ***	0.00	65.776

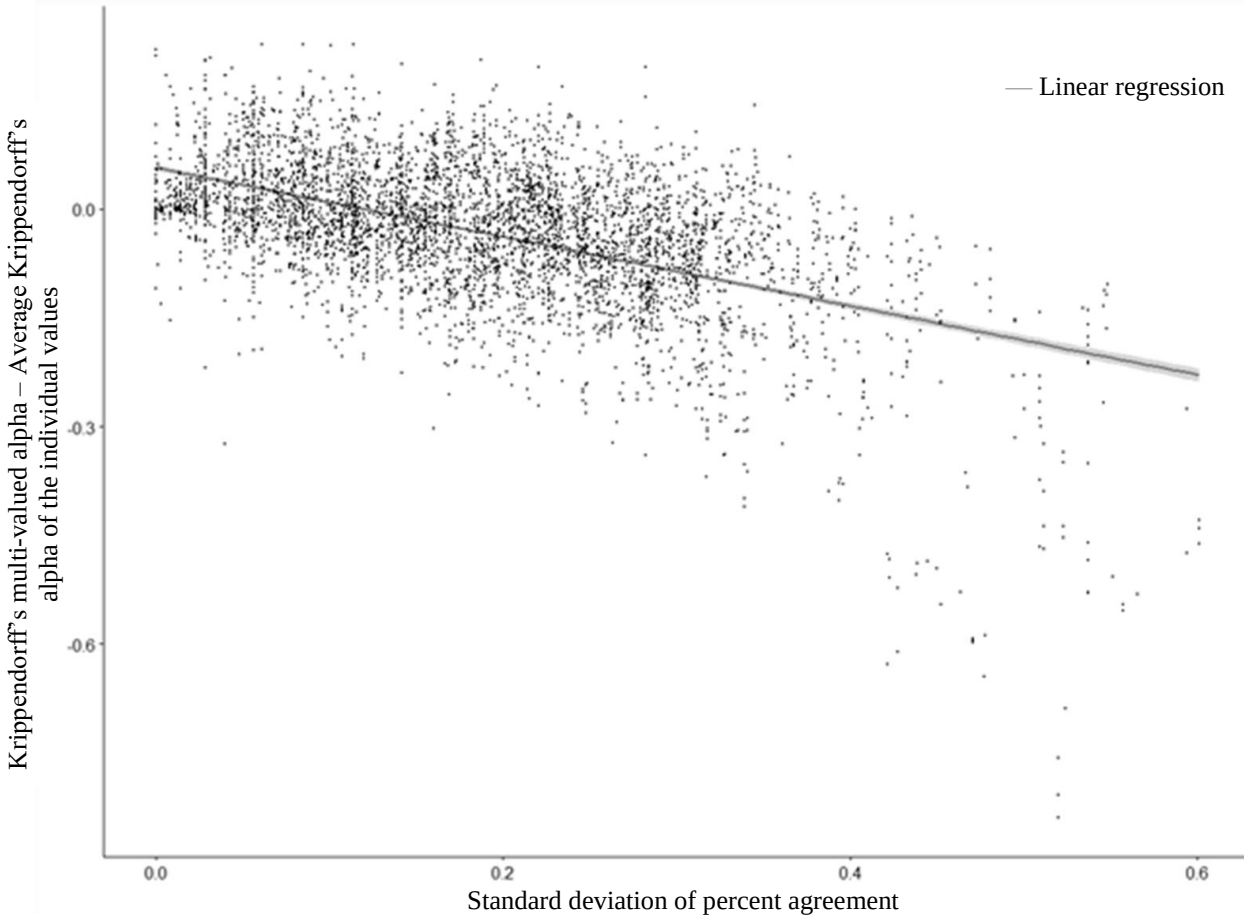
Note. N = 240000. Linear model with mean centered and standardized predictors. * indicates p < .05. ** indicates p < .01. *** indicates p < .001. R^2= .47.

Comparison of Krippendorff's Multi-Valued Alpha With Individual Values' Agreement

Computing reliability for multi-valued data is more resource intense than computing individual values' reliability. Thus, it should be questioned, whether treating multi-valued data with a multi-valued reliability measure results in different conclusions than assessing reliability of the values individually, as if they were independent from each other. Simulation results revealed that the multi-valued alpha was slightly lower than the average Krippendorff's alpha of the individual values ($M_{Difference}$ = -.03; $SD_{Difference}$ = .10). Table 2.2 provides multi-variate regression results with the difference between Krippendorff's multi-valued alpha coefficient and the average of the individual values' alpha. Coefficient estimates revealed strong negative effects for the average percent agreement (β = -.39) and the variation among percent agreement (β = -.70) between the individual values in the data. Consequently, the multi-valued alpha coefficient is particularly lower than the average single-valued alpha for the individual values, if agreement varies strongly between the individual values. The effect is depicted in Figure 2.3. An increase of .1 in the standard deviation of percent agreement among the datasets' individual values led to a decrease of the multi-valued alpha compared to the average single-valued alpha

Figure 2.3

Differences between Krippendorff's alpha for multi-valued alpha and the average Krippendorff's alpha for the individual values (treated as single valued data each) for different levels of variation among the percent agreement of the individual values within the data



Note. N = 5,000 randomly selected multi-valued datasets from the simulation data.

by .06. Thus, the stronger percent agreement varies among values, the more important it is to investigate dependently coded or generated data with multi-valued methods.

Compared to the average single-valued alpha, the multi-valued alpha was even lower for datasets with high average prevalence among the individual values (β = -.20). However, if the individual values were correlated, the multi-valued alpha was higher and closer to the average single-valued alpha (β = .19). This is in line with expectations, because for highly correlated individual values the multi-valued chance agreement converges with the chance agreement of the individual values. The number of units and the average set size had very limited effects.

Sources of Variation in Krippendorff's Multi-Valued Alpha

For single-valued data with a given prevalence, a certain level of agreement leads to a single, defined alpha value. For multi-valued data, however, prevalence of the individual values and the level of agreement are insufficient to derive an exact Krippendorff's alpha. Instead, the distribution of sets, in which single values occur, influences the multi-valued alpha coefficient.

The simulation data comprised groups of 50 simulation runs each. For each of these groups, prevalence and agreement of the individual values were fixed, so that variation in the multi-valued alpha originated from the allocation to sets only. Results showed that the multi-valued alpha varied substantially within groups. The mean range between the highest and lowest observed multi-valued alpha within each group was .14 with a standard deviation of .09. The variation was weaker for large samples (r = -.18), high prevalence (r = -.06), and high agreement (r = -.37). However, even for groups, which's multi-valued alpha was close to the threshold values for acceptable reliability between .60 and .80 (Krippendorff, 2019; Riff, 2014), within-group variation was .08 on average for samples with 500 and 2,000 units and .11 for smaller samples.

Three nested random intercept models with the multi-valued alpha as dependent variable revealed several relevant coefficients, which are listed in Table 2.3. The models were fit with in restricted maximum likelihood estimator and converged normally. First, the empty null model was computed with random intercepts for each group. The groups' random effect explained more than 99% of the variance of the multi-valued alpha. However, we knew that the remaining variance was large enough to account for a mean within-group variation of .14 in the multi-valued alpha. The second model contained eight predictors, of which six varied on group level, and two on the within group level. Group level variables were grant-mean centered, within level variables were group-mean centered. Before, all variables were standardized to allow comparison of coefficients. On the group level, percent agreement showed a strong effect

Table 2.3

Random intercept model. Dependent variable: Krippendorff's alpha for multi-valued data

Predictor	Empty Model			Model 1			Model 2		
	β	*CI*	*df*	*β*	*CI*	*df*	*β*	*CI*	*df*
Fixed effects									
(Intercept)	.00	[-.03, .03]	4799	.01	[-.00, .02]	4638	.01	[-.00, .02]	4638
Units $_{gmc}$				-.02***	[-.04, -.01]	4638	-.02***	[-.04, -.01]	4638
Mean prevalence $_{gmc}$				.04***	[.02, .07]	4638	.04***	[.02, .07]	4638
Prevalence standard deviation $_{gmc}$				-.09***	[-.11, -.08]	4638	-.09***	[-.11, -.08]	4638
Mean percent agreement $_{gmc}$				.75***	[.74, .77]	4638	.75***	[.74, .77]	4638
Percent agreement standard deviation $_{gmc}$				-.26***	[-.27, -.24]	4638	-.26***	[-.27, -.24]	4638
Set size $_{gmc}$				-.07***	[-.09, -.05]	4638	-.07***	[-.09, -.05]	4638
Unique sets $_{bg}$				.07***	[.05, .09]	4638	.07***	[.05, .09]	4638
Unique sets $_{cwc}$				-.00**	[-.01, -.00]	227701	-.01***	[-.01, -.00]	227694
Agreement set size cor.$_{bg}$				-.03***	[-.05, -.01]	4638	-.03***	[-.05, -.01]	4638
Agreement set size cor.$_{wg}$				.14***	[.14, .14]	227701	.14***	[.14, .14]	227694
Agreement set size cor.$_{wg}$: Units$_{gmc}$							-.00***	[-.00, -.00]	227694
Agreement set size cor.$_{wg}$: Mean prevalence$_{gmc}$							.02***	[.02, .02]	227694
Agreement set size cor.$_{wg}$: Prevalence sd$_{gmc}$							.02***	[.02, .02]	227694
Agreement set size cor. $_{wg}$: Mean perc. agreement$_{gmc}$							-.04***	[-.04, -.04]	227694
Agreement set size cor. $_{wg}$: Perc. agreement sd$_{gmc}$							.00***	[.00, .01]	227694
Agreement set size cor.$_{wg}$: Set size$_{gmc}$							-.02***	[-.02, -.02]	227694
Agreement set size cor.$_{wg}$: Unique sets$_{bg}$							.01***	[.00, .01]	227694
Random effects									
σ^2	.00			.00			.00		
τ_{00}	1.00 $_{group}$			.14 $_{group}$			.14 $_{group}$		
ICC	1.00			.99			.99		
N	4800 $_{group}$			4647 $_{group}$			4647 $_{group}$		
Observations	240000			232350			232350		
Marginal R^2 / Conditional R^2	.000 / .996			.854 / .998			.854 / .998		

Note. Random intercept model with group level and within group level predictors. Group level predictors were grand mean centered (*gmc*) and standardized. Within group predictors were standardized across all groups and split up into the between group component (*bg*, i.e., group means) and the within group component (*wg*, i.e., within group variation centered at group mean). Sample size is lower for Model 1 and 2, because cases with full agreement for one of the individual values were excluded. For those cases, the correlation between agreement and set size could not be calculated, because agreement had no variation. ** indicates $p < .01$. *** indicates $p < .001$.

(β = .75), followed by the variation in percent agreement between the individual values (β = -.26). Since both parameters are negatively correlated (r = -.63), these results were as expected. Moderate effects were found for the variation in prevalence among the individual values (β = -.09), the average set size (β = -.07), the average number of unique sets within the data (β = .07), and the average prevalence across individual values (β = .05). Very small effects were found for the number of units (β = -.02) and the average correlation between agreement and each unit's set size within each simulation run (β = -.03).

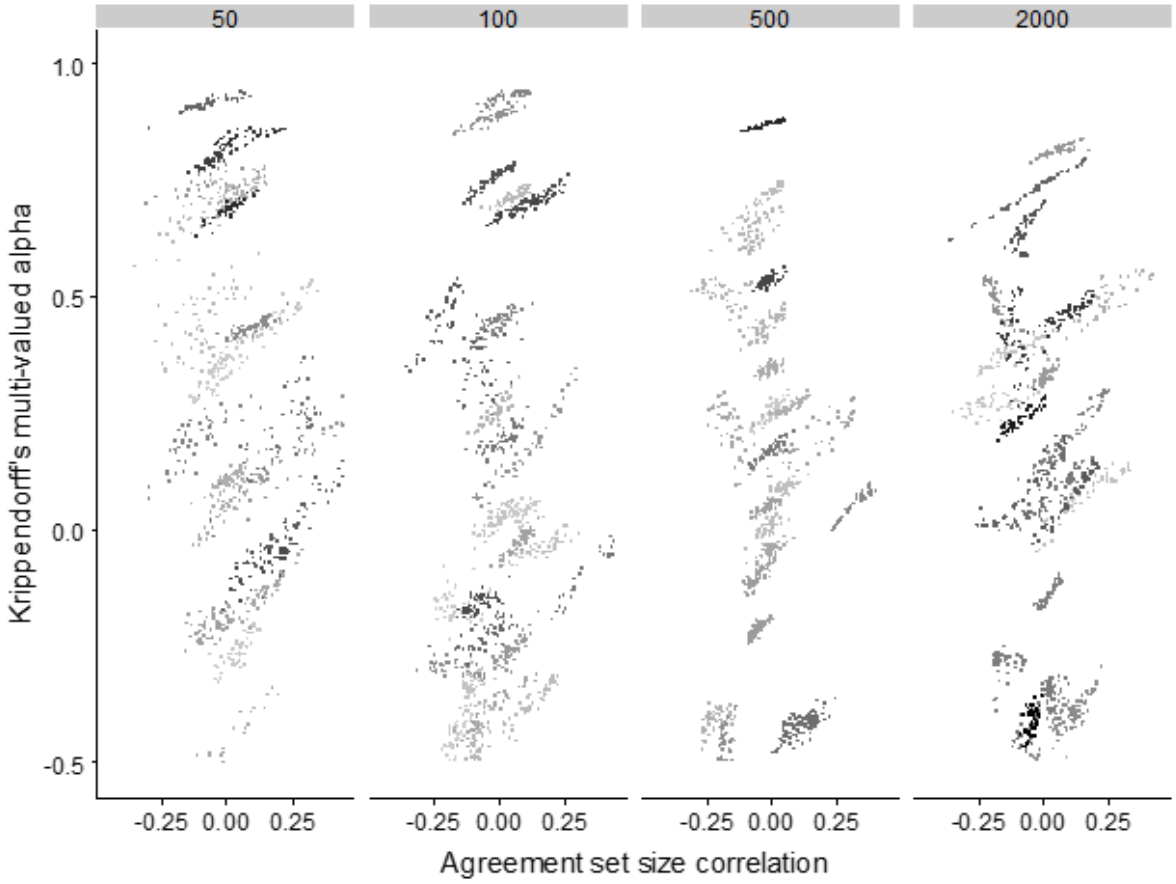
On the within group level, the variation in the number of unique sets with each simulation run had no effect on the multi-valued alpha. However, a rather strong effect was found for the correlation between agreement and each unit's set size within each simulation run (β = .14). Thus, the multi-valued Krippendorff's alpha was significantly lower, if disagreement among individual values occurred in units, which were also coded with other values (i.e., in larger sets). To test, whether this effect is dependent on any of the simulation parameters, the third model contained multiple interaction effects. None of the interaction effects was strong. The strongest, yet relatively small interaction effect was found for the average percent agreement across values within the group (β = -.04). Figure 2.4 depicts 100 randomly selected groups of 50 simulation runs each. The effect of the correlation between agreement and set size becomes clearly visible.

Alternative Set-Theoretic Difference Metric

Krippendorff and Cragg's former approach to calculate multi-valued reliability (2016) used the standardized Dice coefficient as a difference metric for nominal data. The Dice coefficient specializes a general metric, which was previously proposed by Krippendorff

Figure 2.4

Krippendorff's multi-valued alpha in 100 randomly selected groups of 50 simulated multi-valued data sets each.



Note. The datasets within each group share the same marginal distributions and levels of agreement for each individual value. The data sets within each group differ in the distribution of sets and the agreement among sets. Across different sample sizes, level of agreement, and prevalence, Krippendorff's alpha for multi-valued data is higher, if agreement occurs in units with more allocated values (i.e., larger sets).

(1992). The metric is also based on a table comparing the two sets' individual values, but standardized to the range between zero and one:

$$_{metric}\Delta_{CK} = \frac{\frac{\sum_{c\in C}\sum_{k\in K\cap\bar{C}} {}_{metric}\delta^2_{ck}}{|C|} + \frac{\sum_{k\in K}\sum_{c\in C\cap\bar{K}} {}_{metric}\delta^2_{ck}}{|K|}}{|C| + |K|}$$

The components of the metric are depicted in Figure 2.1. For nominal data with $_{nominal}\delta^2_{ck} = \{1\ iff\ c = k; 0\ else\}$ the metric's nominator equals the traditional set-theoretic difference metric

$$_{Set\ Theory}\Delta_{CK} = |K \cap \bar{C}| + |C \cap \bar{K}|,$$

which ranges between zero and $|C \cup K|$, given that the metric between individual values ${}_{metric}\delta_{ck}^{2}$ is between zero and one. Thus, each set could theoretically add to the overall disagreement according to the set's length. Selecting the most appropriate difference metric depends on the data's nature and how sets of values are best compared. First, the standardized metric (see Krippendorff & Craggs, 2016) treats each comparison of two sets equally, independent of the number of values composing the sets. Second, the set-theoretic difference metric weights the difference between two sets with the number of values composing the sets. For example, two sets with three values each can be three times more different than two sets of one value each. And third, the squared difference metric used by Krippendorff (2019) weights the difference between two sets with the product of the number of values in one and the other set. To take up the example, the two sets with three values each can be nine times more different than the two sets with one value each, because with the squared difference metric each value of the first set is compared with each value of the second set. Hence, it is not surprising that the simulation showed lower values for Krippendorff's multi-valued alpha, if disagreement occurred disproportionally in larger sets. Table 2.4 lists a few examples of set comparisons with nominal values for each of the three metrics.

The set-theoretic difference metric offered an opportunity for a multi-valued reliability coefficient, which neither neglects set size by standardizing set differences, nor emphasizes

Table 2.4

Examples of differences between two sets of individual values assessed by the squared difference metric used by Krippendorff's alpha for multi-valued data and the set-theoretic difference metric

Values coder A	Values coder B	Squared difference metric	Set-theoretic difference metric
u	tv	2	3
uv	tv	3	2
suw	tv	6	5
uw	tvw	5	3

disagreements among values in larger sets over those in smaller sets. For many datasets the set-theoretic metric, in which each assigned value can add equally to the disagreement measure, seems to be the most appropriate metric. Application of the set-theoretic difference metric to the multi-valued alpha coefficient required modifications to the denominators of the observed and expected disagreement, in order to fit the difference metric's maximum value of $|C| + |K|$. The modified alpha coefficient implementing the set-theoretic difference metric was:

$$_{Set\ theory\ metric}\alpha_{nominal} = 1 - \frac{D_o}{D_e} = \frac{2\sum_C n_C|C|(\sum_C n_C - 1)}{\sum_C\sum_K o_{CK}(|C|+|K|)} \cdot \frac{\sum_C\sum_K o_{CK}\ {}_{Set\ Theory}\Delta_{CK}}{\sum_C\sum_K n_C\, n_K\ \ {}_{Set\ Theory}\Delta_{CK}}$$

with $_{Set\ Theory}\Delta_{CK} = \frac{\sum_{c\in C}\sum_{k\in K\cap\bar{C}}\ {}_{metric}\delta^2_{ck}}{|C|} + \frac{\sum_{k\in K}\sum_{c\in C\cap\bar{K}}\ {}_{metric}\delta^2_{ck}}{|K|}$, which equals $|K \cap \bar{C}| + |C \cap \bar{K}|$ for nominal data.

The random intercept model was repeated with the same set of simulated data as used previously. Results yielded similar correlation results as obtained for the original multi-valued alpha (see Table 2.5). Also, the alternative set-theoretic multi-valued alpha was .99 correlated with the original multi-valued alpha. However, multi-variate regression analysis with the new multi-valued alpha as dependent variable showed differences: First, the coefficient for the within-group variation of the correlation between agreement and set size was close to zero (β = .01 compared to β = .14 for the original multi-valued alpha coefficient). Thus, disagreement among individual values impacted the coefficient independently of the set size in which they occurred (see also Figure 2.5). This is not to be confused with a standardized difference metric, which ranges between zero (complete disagreement) and one (complete agreement) independently of the sets' length (cf. Krippendorff & Craggs, 2016). Secondly, the mean percent agreement was a slightly stronger predictor for the set-theoretic alpha than for the original multi-valued alpha (β = .78 compared to β = .75). Third, the mean prevalence had a stronger effect on the multi-valued alpha (β = .10 compared to β = .04). However, at the same time, the correlation between agreements and set size on group level, which is highly correlated with the mean prevalence(r = .72) showed a stronger negative effect (β = -.10 compared

Table 2.5

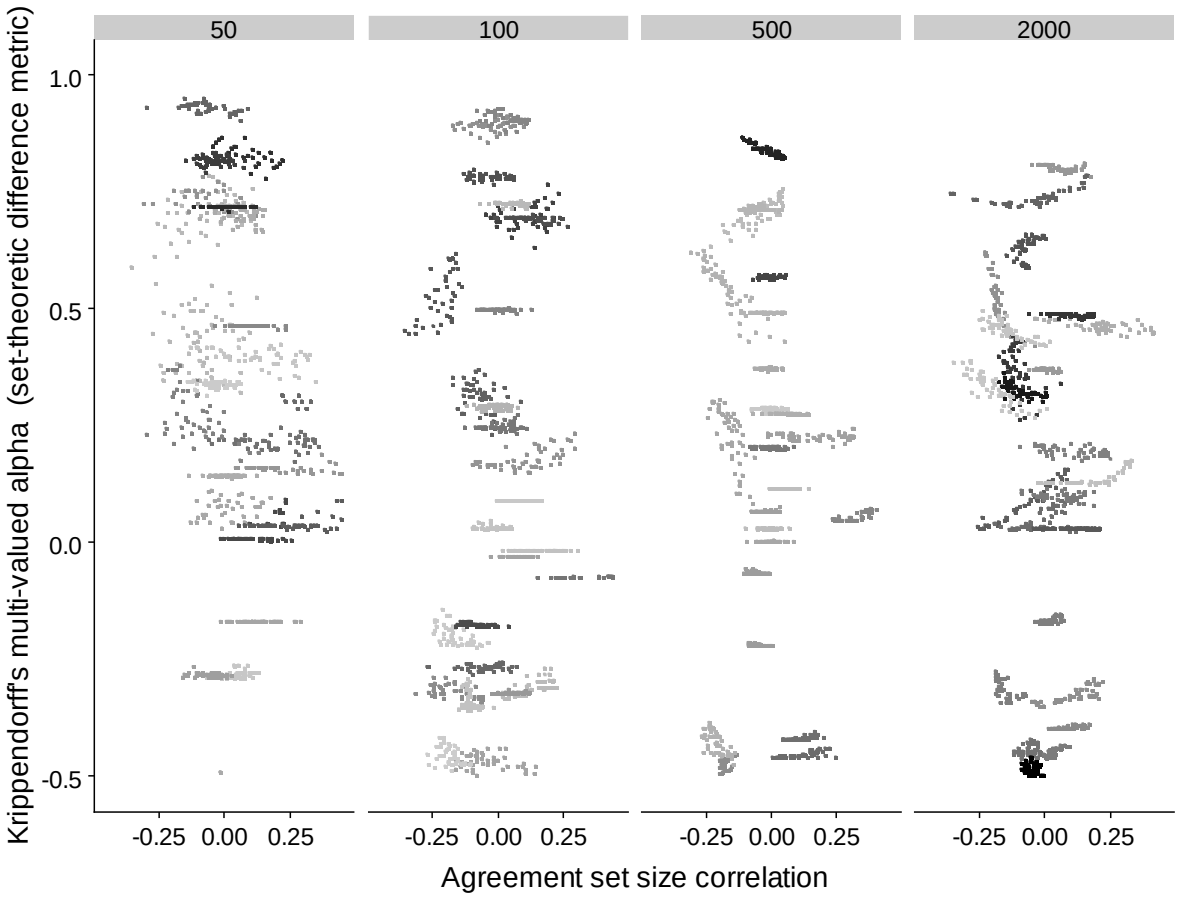
Random intercept model. Dependent variable: alpha with the set-theoretic difference metric

Predictor	Empty Model			Model 1			Model 2		
	β	*CI*	*df*	*β*	*CI*	*df*	*β*	*CI*	*df*
Fixed effects									
(Intercept)	.00	[-.03, .03]	4799	.01	[-.00, .02]	4638	.01	[-.00, .02]	4638
Units $_{gmc}$				-.03***	[-.04, -.02]	4638	-.03***	[-.04, -.02]	4638
Mean prevalence $_{gmc}$				.10***	[.08, .12]	4638	.10***	[.08, .12]	4638
Prevalence standard deviation $_{gmc}$				-.11***	[-.12, -.10]	4638	-.11***	[-.12, -.10]	4638
Mean percent agreement $_{gmc}$				.78***	[.78, .80]	4638	.78***	[.78, .80]	4638
Percent agreement standard deviation $_{gmc}$				-.22***	[-.24, -.21]	4638	-.22***	[-.24, -.21]	4638
Set size $_{gmc}$				-.07***	[-.09, -.05]	4638	-.07***	[-.09, -.05]	4638
Unique sets $_{bg}$				.08***	[.06, .09]	4638	.08***	[.06, .09]	4638
Unique sets $_{cwc}$				.00**	[.00, -.01]	227701	.00***	[.00, -.00]	227694
Agreement set size cor.$_{bg}$				-.10***	[-.12, -.08]	4638	-.10***	[-.12, -.08]	4638
Agreement set size cor.$_{wg}$				-.00***	[-.00, -.00]	227701	-.01***	[-.01, -.00]	227694
Agreement set size cor.$_{wg}$: Units$_{gmc}$							-.00***	[-.00, -.00]	227694
Agreement set size cor.$_{wg}$: Mean prevalence$_{gmc}$							.01***	[.01, .01]	227694
Agreement set size cor.$_{wg}$: Prevalence sd$_{gmc}$							-.01***	[-.01, -.01]	227694
Agreement set size cor. $_{wg}$: Mean perc. agreement$_{gmc}$							.01***	[.01, .01]	227694
Agreement set size cor. $_{wg}$: Perc. agreement sd$_{gmc}$							-.01***	[-.01, .00]	227694
Agreement set size cor.$_{wg}$: Set size$_{gmc}$							-.01***	[-.01, -.01]	227694
Agreement set size cor.$_{wg}$: Unique sets$_{bg}$							.00***	[.00, .01]	227694
Random effects									
σ^2	.00			.00			.00		
τ_{00}	1.00 $_{group}$			.14 $_{group}$			.14 $_{group}$		
ICC	1.00			.99			.99		
N	4800 $_{group}$			4647 $_{group}$			4647 $_{group}$		
Observations	240000			232350			232350		
Marginal R^2 / Conditional R^2	.000 / .996			.852 / .998			.852 / .998		

Note. Random intercept model with group level and within group level predictors. Group level predictors were grand mean centered (*gmc*) and standardized. Within group predictors were standardized across all groups and split up into the between group component (*bg*, i.e., group means) and the within group component (*wg*, i.e., within group variation centered at group mean). Sample size is lower for Model 1 and 2, because cases with full agreement for one of the individual values were excluded. For those cases, the correlation between agreement and set size could not be calculated, because agreement had no variation. ** indicates $p < .01$. *** indicates $p < .001$.

Figure 2.5

Multi-valued alpha with the set-theoretic difference metric in 100 randomly selected groups of 50 simulated multi-valued data sets each.



Note. The datasets within each group share the same marginal distributions and levels of agreement for each individual value. The data sets within each group differ in the distribution of sets and the agreement among sets. Notably, the correlation between agreement and set size does not influence the alpha coefficient.

to β = -.03). For the other predictors such as, the average set size, the number of unique sets, and the number of units no substantial differences were observed between Krippendorff's original difference metric and the set-theoretic difference metric.

Overall, the within-group variation among simulated datasets with identical marginal distributions for each value was almost halved compared to the original multi-valued alpha (*M* = .04; *SD* = .05 compared to *M* = .09; *SD* = .06 for the original multi-valued alpha). Nevertheless, this did not imply that the set-theoretic multi-valued alpha was substantially closer to the average single-valued Krippendorff's alpha of the underlying individual values.

The absolute deviation of the set-theoretic multi-valued alpha from the average single-valued alpha (M = .08; SD = .08) was only slightly lower than for the multi-valued alpha as proposed by Krippendorff (M = .09; SD = .09). Thus, a multi-valued approach is still meaningful. Averaging single-valued alpha coefficients for marginal agreement on single values of multi-valued data is no appropriate alternative.

Multi-Valued Reliability With Symmetric Values

Krippendorff's alpha for multi-valued data is defined for asymmetric values, which contain more information if present (the unit is as defined by the value) than if not present (the unit can be defined by any potential other value). However, for many cases of nominal reliability data, values contain symmetric information. In the single-valued case, Krippendorff's alpha is independent of whether a symmetric nominal binary value is coded one for *true* and zero for *false*, or the other way around (reverse-coded). In the multi-valued case, sets change depending on how each value is coded. Thus, Krippendorff's multi-valued alpha was hypothesized to vary depending on whether single values are reverse-coded or not.

To test the hypothesized effect and to exploratively investigate its magnitude, 2,800 multi-valued datasets from the simulation data were replicated with partially reverse-coded individual values. Comparison of Krippendorff's multi-valued alpha for the datasets with reverse-coded items with the alphas for the original datasets revealed a very limited mean deviation of .03 (SD = .03). For the alternative set-theoretic difference metric the mean deviation between original datasets and the modified datasets, in which selected values were reverse-coded, was even lower (M = .02; SD = .03). However, the deviation was stronger, if the symmetric values were positively or negatively correlated with other values of the dataset. Further, the deviation was higher, if percent agreement deviated stronger between the symmetric values and other values of the dataset.

Discussion

This article investigated the statistical properties of Krippendorff's (2019) alpha for reliability of multi-valued data and outlined an alternative implementation based on the set-theoretic difference metric. Beginning with the theoretical concept of multi-valued data, the analyses provided a detailed assessment of the coefficient's purpose, implementation, and statistical properties.

Simulation results revealed substantial deviation between Krippendorff's multi-valued alpha and its single-valued version. Differences were expected, because the alpha for multi-valued data computes the expected chance agreement based on the distribution of sets rather than the distribution of individual values within the dataset. The multi-valued alpha was lower than the average alpha of the individual values, if percent agreement varied strongly among the individual values of the multi-valued dataset and if correlation among the dataset's individual values was low. Consequently, results showed that treating interdependently coded multi-valued data with coefficients for single valued data may result in serious misjudgment of chance agreement and, in turn, reliability. For example, qualitative researchers who investigate multi-valued phenomena and rely on single valued agreement coefficients such as the standard Krippendorff's alpha run the risk of overestimating the reliability of their conclusions.

Moreover, the simulations provide insights into the characteristics of the data, which influence Krippendorff's multi-valued alpha. As expected, the alpha for multi-valued data was primarily dependent on the percent agreement among the coders. However, the multi-valued alpha was also substantially influenced by the set size of those sets, in which the coders disagreed. The alpha coefficient was significantly lower, if disagreement occurred in larger sets. The effect results from the difference metric applied by Krippendorff (2019). The metric is meaningful, if assigned values should be compared with all values assigned by the other coders on a one-by-one basis. However, for many use cases, a metric that treats disagreements independently of the set size is better suited. For example, if two coders identify motives in art

work, disagreement among a single motive should not be treated less or more severe depending on whether the particular art work depicts just one or several motives. However, disagreement should also not be standardized to the range of zero to one for each unit. Instead, an art work with a total of three unique motives, should also contribute exactly three times, if the coders disagree on all three values. Results showed that a modified alpha implementing the set-theoretic difference metric is better suited for those cases than Krippendorff's original multi-valued alpha. That modified alpha was shown to be independent of whether disagreement takes place in larger or smaller sets. For Krippendorff original multi-valued alpha this was not the case.

Furthermore, the design of Krippendorff's alpha for multi-valued data suggested that the reliability assessment changes depending on whether symmetric values are reverse-coded or not. Results confirmed differences in the multi-valued alpha depending on whether a single value was reverse-coded or not. For most cases the deviation was negligibly small. Nevertheless, since reliability of the data should not depend on how symmetric values are coded, these results suggest that Krippendorff's (2019) multi-valued alpha may not serve as an appropriate indicator for reliability of multi-valued data.

Implications

Results of the statistical analysis of multi-valued datasets supported the need for specific reliability coefficients for multi-valued data, if the data's individual values were coded dependently, for example simultaneously by the same coders or in consideration of the other values. Researchers that apply agreement coefficients designed for single-valued data to individual values of multi-valued data run the risk to significantly misjudge the reliability of their data. Depending on the frequency of individual values in sets and the sets' distributions, the expected agreement estimated by the coefficients for single-valued data may be significantly above or below chance agreement. Consequently, researchers should use coefficients for multi-

valued data, whenever multiple values are coded by the same coders or dependently by any other means.

In the case of single-valued data, disagreements among data points are characterized by the value's difference metric (i.e., nominal, ordinal, interval, or metric difference). However, for multi-valued data, disagreement is not only characterized by the difference metric between individual values, but also by the metric between sets. Krippendorff's (2019) agreement coefficient for multi-valued data uses a squared metric to compare sets, which weights disagreements among larger sets more strongly than among smaller sets. Thus, this article proposes to carefully select the difference metric based on the data.

For qualitative researchers, the modified version of Krippendorff's alpha for multi-valued data implementing the set-theoretic difference metric provides the opportunity to prove reliability of the extraction of meaning from data, if a set-theoretic difference metric better fits the data than the squared difference metric applied by Krippendorff (2019). The most appropriate difference metric depends on the data. However, in many cases including the examples of multi-valued data provided by Krippendorff (2019), and Krippendorff and Craggs (2016), the set-theoretic difference metric seems to be the more natural metric. In most nominal datasets there is no reason, why assigning and not assigning a single nominal value should be more different, if the respective unit is assigned to other values as well. Moreover, simulation results revealed that the modified alpha with set-theoretic difference metric is less sensitive to reverse-coding of symmetric values. Further, the R code provided in Appendix C makes it very easy to compute alpha for multi-valued data with either the original squared difference metric by Krippendorff (2019) or the set-theoretic difference metric. The code is efficiently written and quickly computes alpha for up to 1.000 unique sets on a standard computer.

Excursus: Application of the Set-Theoretic Difference Metric to the Later Life Workplace Index

The development of the Later Life Workplace Index (LLWI, see Article 1) marks a typical example of qualitative research based on multi-valued data. We applied the set-theoretic difference metric and Krippendorff's (2019) metric to compute alpha of the multi-valued codes that two coders allocated to 2559 paragraphs of the expert interviews, which we conducted with researchers and professionals in the field of aging at work to develop the index (see Figure **2**). The coders were asked to decide for each paragraph of the interviews which of the LLWI's nine dimensions were addressed. In many cases, the coders allocated more than one dimension per paragraph, for example, because a mentoring program mentioned by one of the interviewees aimed at both, improving the knowledge transfer between older and younger employees and a reduction of age stereotypes within both groups. Thus, the reliability data was multi-valued. The coders allocated 73 unique sets including the empty set (the average non-empty set size was 1.28 codes). Alpha values for the two difference metrics can be obtained from Table 2.6.

For the LLWI, application of the set-theoretic difference metric resulted in higher alpha values compared to the multi-valued alpha based on Krippendorff's difference metric. Following from this study's Monte-Carlo simulation the difference may result from the fact that within the LLWI's multi-valued reliability data, disagreements between the two coders were more present in units with multiple applicable dimensions. The first dimension applicable to a certain unit may be clear to both coders, while the second or third dimension may be vaguer, thus, resulting in more disagreements. However, whether disagreements occur in larger or smaller sets does not affect the LLWI's reliability, so that alpha based on the set-theoretic difference metric that is independent of the set size in which disagreements occur is the better

Table 2.6

Krippendorff's alpha values for the Later Life Workplace Index

	Krippendorff's alpha for multi-valued data	
	Krippendorff's difference metric	*Set-theoretic difference metric*
Overall Later Life Workplace Index	.75	.82
Individual values:		
Organizational culture	.62	.67
Leadership	.73	.79
Work design	.76	.83
Health management	.79	.83
Individual development	.80	.84
Knowledge management	.80	.88
Transition to retirement	.70	.75
Continued employment	.66	.72
Health and retirement coverage	.62	.63
Empty set	.88	.88

suited indicator to assess reliability. Furthermore, because the LLWI's reliability data is multi-valued, a multi-valued alpha is more appropriate than computing single-valued alphas for each dimension as done in Article 1. The single-valued alpha ranged from .65 to .92 (see Article 1), thus being somewhat higher than the multi-valued alpha. Nevertheless, the multi-valued alpha still indicates good overall reliability of the index (alpha = .82). Only, the health and retirement coverage domain showed an insufficient multi-valued alpha (alpha = .62) indicating further improvement potential on the dimension's definition.

To conclude, for the LLWI's interview data, in which a single thought or sentence may refer to multiple index dimensions, alpha for multi-valued data assesses change agreement between the coders more accurately than single valued alphas and thus, provides a more accurate indication for the reliability of the index' framework. We used obtained reliability information to further improve the index before operationalization.

Limitations and Future Research

First, the simulation analyses were based on nominal data only. Both, Krippendorff's multi-valued alpha, and the alternative alpha applying the set-theoretic difference metric can also be applied to ordinal, interval, and metric data. However, the need for multi-valued reliability assessment originated from qualitative content analysis (Krippendorff, 2019), in which most coding schemes are nominal. Thus, the present article covers a major use case for multi-valued data. None of the analyses was specifically tied to nominal data, wherefore results should be transferable to other data types. In particular, the multi-valued alpha using the set-theoretic difference metric is applicable for non-nominal data, too.

Second, this study is based on simulated data with two coders only. Two coders were chosen, because percent agreement—one of the predictors of this study—is clearly defined for two coders, whereas it becomes more difficult to interpret for three or more coders. However, results of this study should also hold for three or more coders, because Krippendorff's alpha is computed via contingency tables, in which additional coders have the same effect as additional units. Results were consistently found 50 to 2,000 units and, thus, also apply for more coders.

Third, this article provides an implementation of the set-theoretic difference metric for multi-valued reliability assessment. The set-theoretic difference metric arguably is a very common and natural difference metric for unordered sets. However, other metrics such as for

example the Hausdorff metric (see Rodríguez-López & Romaguera, 2004) may suit certain data even better. Future studies may implement multi-valued reliability assessment based on further metrics and provide guidance on how to choose the correct metric.

In conclusion, this article advances the understanding of how to reassure reliability with a chance-corrected agreement coefficient in multi-valued data. It provides an alternative alpha for multi-valued data based on the set-theoretic difference metric, which is more appropriate for certain datasets than the original alpha by Krippendorff (2019). In any case, the opportunity to compute a chance-corrected level of agreement for multi-valued data allows researchers to reassure reliability of their data, even if the research design and its data is not single-valued and, thus, not suitable for one of the standard single-value agreement coefficients, such as Krippendorff's alpha, Cohen's kappa, and Scott's pi. Hopefully, these advancements further encourage qualitative researchers to independently reassure the reliability of their conclusions and present reliability statistics in their publication.

3

Organizational Practices for the Aging Workforce: Development and Validation of the Later Life Workplace Index

Max R. Wilckens, Anne M. Wöhrmann, Jürgen Deller, and Mo Wang

Published as: Wilckens, M. R., Wöhrmann, A. M., Deller, J., & Wang, M. (2020). Organizational practices for the aging workforce: Development and validation of the Later Life Workplace Index. *Work, Aging and Retirement*. Advance online publication. https://doi.org/10.1093/workar/waaa012

Abstract

The present three studies focused on the development and validation of a multifaceted measure of organizational practices for the aging workforce, the Later Life Workplace Index (LLWI). The first study developed a comprehensive item pool based on expert interview evidence from Germany and the United States. Two further studies among workers across industries in Germany (N = 609, N = 349) provided psychometric evidence. The LLWI comprises nine distinct domains of organizational practices for the aging workforce, namely an age-friendly organizational climate and leadership style, certain work design characteristics, health management, individual development opportunities, knowledge management, the design of the retirement transition, continued employment opportunities, and health and retirement coverage. The final LLWI consists of 80 items in total. In addition, the studies demonstrated that the LLWI measures correlated with older workers' work outcomes such as stress level, workability, person-organization fit, and post-retirement work intentions in meaningful ways. Applications for the LLWI in research and practice are discussed.

Keywords: aging at work, aging workforce, bridge employment, construct validation, human resource practices, later life work, older workers, organizational practices, scale development

In wake of increased retirement ages and an aging ‘baby boom’ generation, most developed countries face an aging workforce (OECD, 2017). Extended working lives require organizations to employ older individuals successfully, that is healthy, motivated, and productive. Researchers have named a variety of different organizational practices that may drive successful employment of an older and increasingly age-diverse workforce (e.g., Armstrong-Stassen & Lee, 2009; Göbel & Zwick, 2013; Parker & Andrei, 2020; Wöhrmann et al., 2018). However, validated measures with a thorough conceptual coverage of organizational practices relevant to the context of aging at work remain limited (Boehm et al., 2014).

Given that age-related organizational practices are often multifaceted (e.g., Kooij et al., 2014; van Dalen et al., 2015), multidimensional measures with a thorough conceptual coverage are required in order to facilitate better understanding of these organizational practices in research and to allow for differentiated analyses of organizations in practice. Organizational practices relevant to the context of aging differ from general organizational practices in that employees’ individual needs and circumstances change throughout the lifespan due to individual life courses and aging processes (Kooij et al., 2014; Zacher, 2015a). Lifespan theory thereby adds to the general and widely supported reasoning that organizational practices—particularly those attributable to strategic human resources management—influence employees’ attitudes and behaviors, which in turn affects work outcome gains (e.g., Huselid, 1995).

In this article we report three studies focusing on the development of a comprehensive measure assessing organizational practices relevant in the context of aging at work. This *Later Life Workplace Index (LLWI)* is a multifaceted measure to facilitate disentangling and understanding the variety of relevant organizational practices. Our aim was to develop a measure that is suitable for field research and for identifying potential improvements in practice. The measure stems from a conceptual framework that we recently developed and published and

that is based on qualitative evidence from Germany and the U.S. (Wilckens, Wöhrmann, Adams, et al., 2020; Wöhrmann et al., 2018). This framework comprises nine domains of organizational practices relevant to the context of aging at work. The present article describes the LLWI's operationalization and its validation.

The new measure contributes to the body of existing literature in several ways. First, by building upon qualitative evidence from 27 expert interviews in Germany and assessments of 61 companies in the U.S., the LLWI is characterized by a broad and thorough conceptual coverage. The measure is suitable for a wide range of contexts and applications, because it is neither tied to a specific context nor focused on a specific selection of practices related to the aging workforce (e.g., developmental practices only). Second, the LLWI is multidimensional and thereby facilitates a more granular understanding of organizational practices for the aging workforce. Existing measures have suggested positive effects of organizational practices in general (e.g., Boehm et al., 2014; Taneva & Arnold, 2018) but suffer from a unidimensional factor structure or limited psychometric evaluation. The multidimensionality of the new measure, in contrast, allows different organizational practices to be distinguished. Third, the measure not only focuses on explicit human resources programs (e.g., existence of a mentoring program) but also addresses informal practices and norms (e.g., older and younger employees pass on their knowledge to other generations), which is important for capturing the work environment as experienced by the workers (Boselie et al., 2005; Wright, 2002). Depending on the respondents, the measure assesses both practices as designed or implemented if assessed via human resource managers or general managers and practices as experienced if assessed via affected workers. Workers' experiences of the practices capture to what extent offered practices reach the individual worker. This is a key prerequisite as most practices are effective either by shaping the work environment for the workers or by influencing the workers' attitudes and behavior to improve their aging (Nishii et al., 2018; Zacher & Yang, 2016). The measure can be easily administered, which is of particular importance when conducting research on

organizations and employees during their work time. Finally, the new measure not only supports research but also contributes to workplace improvement. Results generated by this measure can serve organizations as a foundation for interventions on various practices. A detailed evaluation of the status quo helps management identify specific areas for improvement and allocate resources effectively. This cannot be achieved with a lump-sum assessment that does not differentiate between practices.

Existing Measures in the Literature

We conducted a review of the existing literature and identified several measures of organizational practices for the aging workforce. The existing measures are of three types. First, several measures assess organizational practices in a unidimensional manner. Unidimensional measures of organizational practices facilitate research on practices in general. However, they are too broad to disentangle organizational practices and thus do not allow for diagnoses regarding specific practices. For example, Taneva and Arnold (2018) developed an 8-item scale on organizational practices based on qualitative interviews among older workers. Their scale includes items on whether employees have "challenging and meaningful tasks" and whether the "significant role mature employees can play" is recognized. Each of the eight items covers a different content aspect, but reliable assessment of specific practices is not possible. Another unidimensional measure of "age-inclusive human resources practices" was developed by Boehm et al. (2014). The 5-item measure primarily addresses age-inclusive recruiting activities and development opportunities. The authors showed positive organizational level effects of the practices on a 4-item "age-diversity climate" measure assessing inclusion, good management of people of different ages, and equal opportunities for developmental growth, regardless of age. Similarly, Zacher and Yang (2016, p. 2) proposed the construct of an organizational climate for successful aging, defined by them as "employees' shared perceptions of the extent to which their organization enables successful aging." The authors operationalized the construct by

assessing respondents' perceptions of the organization's understanding for age-related changes, responsiveness to age-related changes, and supportiveness of all age groups. All these measures assess practices in a unidimensional way. However, given the wide range of organizational practices, work outcomes are not unidimensionally affected by all organizational practices in the same way (Kooij et al., 2010). Moreover, within an organization certain factors may be present and effective, while others are not. Unidimensional measures are capable of assessing the overall nature of an age-friendly organizational work environment, but they are incapable of differentiating specific practices.

Second, human resources management research investigated organizational practices using multifaceted measures (e.g., Armstrong-Stassen & Templer, 2006; Kooij et al., 2014). These measures incorporate a wide range of different practices and most often assess their availability within the organization (Boselie et al., 2005). For example, Kooij (2014) proposed bundles of human resource practices, following Baltes and Baltes's (1990) lifespan theory of selection, optimization, and compensation. However, the practices were measured by a dichotomous response format, which does not allow the intensity, saturation and quality of the practices to be assessed (Boselie et al., 2005; Vandenberg et al., 1999). Moreover, these kinds of measures are lists of practices, which assess each practice with a single item. But the implementation and framing of practices differ from organization to organization. Thus, a single item per practice seems insufficient for reliably capturing the constructs of interest (DeVellis, 2017). For instance, a single item assessing ergonomic adjustments of the workplace may be conceptualized very differently by the respondents. Using several items rather than one to address the main components of ergonomic workplace adjustment in a multi-item scale would provide more adequate conceptual coverage and more reliable assessment of the construct.

Third, our review revealed two measures that assess organizational practices for the aging workforce in a multifaceted, multi-item manner. However, these measures fall short either in terms of the evaluation of the psychometric measurement quality or in terms of

thorough conceptual coverage. One measure was proposed and used by Armstrong-Stassen (2008), who listed 28 organizational practices and grouped them into seven strategies (flexible working options, training and development, job design, recognition and respect, performance evaluation, compensation, pre- and post-retirement options). Armstrong-Stassen asked the participants to rate their employer's engagement in listed practices and obtained acceptable coefficient alphas for the seven strategies. However, she did not examine the factor structure underlying the 28 items and whether the strategies were sufficiently distinct, thereby neglecting to evaluate the measurement quality. Another measure that was proposed and used by Armstrong-Stassen and Lee (2009) assessed four organizational practices (training and development for older workers, training for the managers, recognition of older workers, and pre- and post-retirement options) with two to five items each. This measure was jointly tested with three further constructs (contribution to the organization, perceived respect shown by work group members, and whether workers were treated with respect) in a confirmatory factor analysis (CFA), which revealed good fit. However, this four-dimension measure does not comprehensively cover organizational practices for the aging workforce: It does not, for instance, cover job design.

In conclusion, the measures obtained from the literature are either too broad, lack comprehensive conceptual coverage, or have not been sufficiently shown to be of sound psychometric quality. Our aim was to overcome these shortcomings while also responding to calls from the literature such as that by Zacher and Yang (2016, p. 9), who emphasized the need for developing a "multidimensional model […], which includes shared perceptions of more specific age-related organizational policies, norms, practices, and procedures related to topics such as recruitment, training, performance appraisal, and promotion."

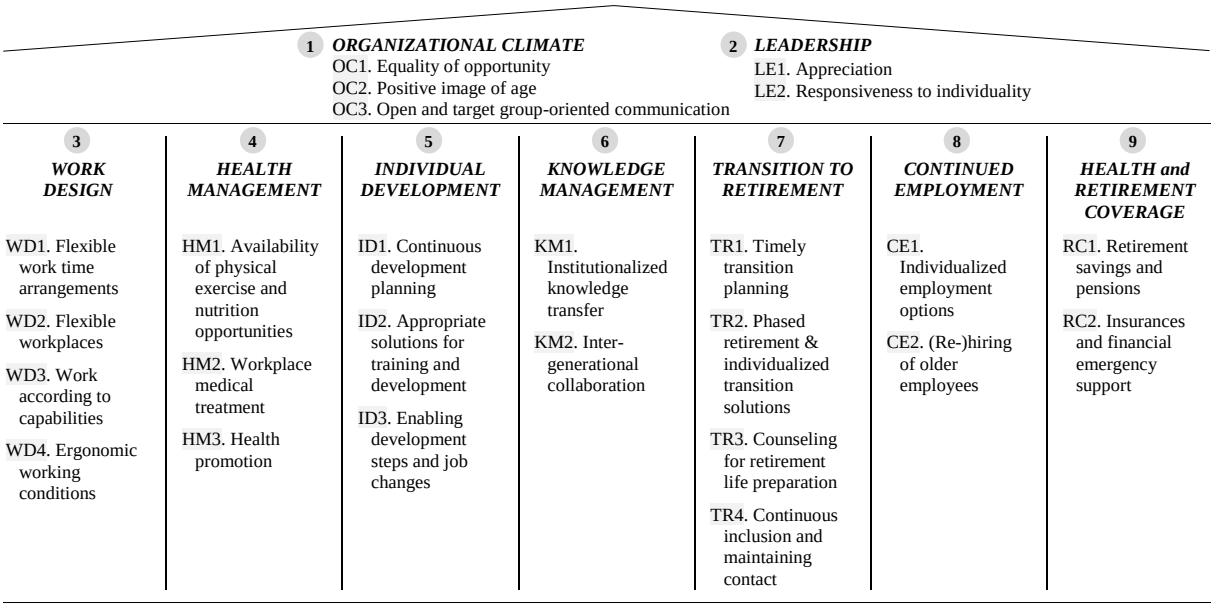
Conceptual Framework of the Later Life Workplace Index

The conceptualization of the LLWI is rooted in qualitative, empirical evidence on organizational practices for the aging workforce from Germany and the United States (Wilckens, Wöhrmann, Adams, et al., 2020). In particular, the initial qualitative framework was developed by Wöhrmann et al. (2018), based on 27 semi-structured expert interviews in Germany. These interviewees were asked to "elaborate on aspects that they thought were characteristic of good organizational management practice involving employees aged 60 and older" (p. 79). Interviewees had a wide range of expertise and various backgrounds in research, practice, and politics. Researchers covered the fields of demographics, economics, gerontology, human resources management, and psychology. Practitioners were human resource executives; older workers including some who had already reached retirement age; and representatives of strategic and operational management in various industries. The experts had either dealt with an aging workforce as part of their job responsibilities or could share personal experience from later life employment. Wöhrmann et al. (2018) systematically derived the initial taxonomy of the LLWI using qualitative content analysis (Mayring, 2010). Subsequently, Wilckens et al. (2020) amended the taxonomy to incorporate organizational practices relevant in the United States. These practices had previously been identified by the *Age Smart Employer Award* honoring businesses in New York City that successfully engage and retain older employees (Finkelstein et al., 2013). Thus, the LLWI comprises an interculturally validated and comprehensive set of practices for "good organizational management of later life work" (Wilckens, Wöhrmann, Adams, et al., 2020, p. 70).

As shown in Figure 3.1, the taxonomy of the LLWI consists of nine domains covering age-inclusive organizational climate and leadership as well as age-related practices and age-friendly organizational conditions pertaining to work design, health management, individual development, knowledge management, transition to retirement, continued employment options, and health and retirement coverage. Each of the nine domains is further broken down into two

Figure 3.1

Domains and underlying facets of the Later Life Workplace Index



Note. First published by Wilckens et al. (2020). Published with permission of © Max R. Wilckens, Anne M. Wöhrmann, Jürgen Deller 2019. All Rights Reserved

to four facets (e.g., *the organizational climate* by *equality of opportunities* for all age groups, a *positive image of age*, and *open and target group-specific communication*). Despite the complexity of the model, intercoder reliability results confirmed that the individual aspects are well differentiated (Wilckens, Wöhrmann, Adams, et al., 2020). A detailed definition of each facet can be obtained from Appendix D.

Given the broad range of disciplines contributing to its conceptualization, the LLWI comprehensively reflects the diverse spectrum of research on "how [employers can] make an aging work staff work" (Henkens et al., 2018, p. 809). Some researchers, for example, explored which organizational practices encourage older employees to prolong their careers (e.g., Armstrong-Stassen, 2008). Others focused on gains in older employees' work outcomes from an organizational perspective. For instance, Göbel and Zwick (2013) found positive effects of specific trainings for older employees, adaptations of the workplace, and mixed-age teams on labor productivity; the effect increased with age. More generally, research identified organizational practices from a broad range of domains including leadership, organizational

climate, health promotion, knowledge management, and work design that positively affect work outcomes (Armstrong-Stassen & Lee, 2009; Armstrong-Stassen & Templer, 2006; Börsch-Supan & Weiss, 2016; Klaffke, 2014; Kunze et al., 2013; Naegele & Walker, 2006; Schuett, 2014; Tisch, 2015; Zacher & Yang, 2016).

Organizational practices for the aging workforce are not limited to practices explicitly addressing older employees (e.g., specific trainings *for older workers*). Rooted in personnel diversity research, Boehm et al. (2014) found that age-inclusive practices (e.g., equal access to training *for all age groups*) positively influenced work outcomes independent of the worker's age. Qualitative research on organizational antecedents of older workers' work outcomes supports the notion that both practices specific to older employees and those generally supporting age-inclusiveness in the organization are relevant (Taneva & Arnold, 2018). A sole focus on age-specific practices may even negatively impact the organizational climate by devaluing older employees, even if those practices are also implemented to accommodate older employees (Hennekam & Herrbach, 2015). The LLWI comprehensively integrates both age-specific and age-inclusive practices. Moreover, qualitative studies identified not only these two forms of practices, but also aspects attributable to organizational climate and leadership style (e.g., leaders' recognition of work outcomes and supervirsor support; Silver et al., 2019; Taneva & Arnold, 2018). This triad of organizational practices, climate, and leadership, has previously been proposed by Boehm and Dwertmann (2015) and is covered by the framework of the LLWI.

In summary, the conceptual framework of the LLWI builds upon qualitative evidence and covers the breadth of the existing research streams on organizational practices for the aging workforce. Thereby, the framework provides a thorough and precise conceptualization as emphasized by scale development research to achieve a substantively valid measure (Clark & Watson, 1995; Hinkin, 1998; Worthington & Whittaker, 2006). We conducted the current studies to operationalize the measure as conceptualized.

Table 3.1

Content of each study

Study 1 (Item generation)	Study 2 (Scale development)	Study 3 (Cross-validation)
Item style and response format definition Initial item pool development Item revision based on topic level experts' assessment Initial item selection based on pre-tests	Item selection for the final scales based on item distributions, explorative and confirmatory factor analysis Construct validity assessment Convergent validity assessment regarding organizational practices and climate measures Discriminant validity assessment regarding positive and negative affect	Construct validation based on confirmatory factor analysis Criterion validity assessment regarding older employees' work outcomes

The Current Studies

To create a valid and reliable measure, we followed widely applied and theoretically derived recommendations for the scale development and score validation process by Clark and Watson (1995) and Hinkin (1998). The three studies reported in this article are outlined in Table 3.1.

Study 1 addressed the generation of an initial item pool based on the qualitative framework of the LLWI. We explicitly set the goal of maintaining the comprehensive content coverage of this qualitative framework within the operationalized measure. The multifaceted structure of the LLWI should enable researchers and practitioners to disentangle organizational practices for the aging workforce. Thus, the operationalization closely followed the qualitatively derived definitions of the nine LLWI domains and the underlying facets. For the item development, we opted for a Likert-scale type measure, consulted topic-level experts to review proposed items, and pre-tested the item set on several small samples.

Study 2 covered the development of the scale. We administered the item set to employees in Germany to derive the factor structure, built several subscales accordingly, and iteratively removed items not fitting proposed scales. Worthington and Whittaker (2006) presented empirically derived best practices for the combined use of explorative and confirmatory factor analysis techniques, which we incorporated throughout the study. The study

also provided initial evidence for discriminant and convergent validity of the LLWI scales. To ensure wide applicability of the new measure in research and practice, the sample comprised responses across industries of various organizational size.

Study 3 cross-validated previous results regarding the fit of the scales to the LLWI model in a second sample of older workers across industries in Germany. Moreover, the study provided validity evidence on the basis of several criterion variables, such as older workers' health, workability, well-being, work engagement, perceived stress, and person-job fit.

Study 1: Item Generation

In their recommendations and outline for a thorough scale development process, Clark and Watson (1995) emphasized the importance of an exhaustive item pool. The qualitative framework of the LLWI supplies detailed definitions of each domain and the indicating facets relevant in the context of aging at work (see Appendix D for the final scale including the conceptual definitions; Wilckens, Wöhrmann, Adams, et al., 2020). These definitions formed the starting point for our approach. The purpose of this first study was to review the literature for each domain, identify measurable indicators, and develop a comprehensive item pool as a foundation for the LLWI scales.

Method

Procedure. To develop an item pool for the LLWI measure, we followed a four-step process for each of the nine LLWI domains. First, we identified relevant scales from the literature that measure content areas similar to those described by the LLWI construct definitions (see Appendix F for a list of identified instruments). For many content areas, however, the literature did not provide adequate measures. That was particularly the case for the age-specific aspects related to retirement and continued employment. Moreover, none of the items within the identified measures were directly suitable to the LLWI. Thus, we developed new items based on the content areas covered by existing measures. Second, we consulted the

LLWI construct definitions to add to the item pool. To enrich the content base, we also consulted the original qualitative interview transcriptions on which the LLWI framework is based. We then compared these aspects stated in the definitions and the interviews with the items developed in the first step. For aspects not yet covered by the items, we developed additional items. This resulted in an overall item pool of about 200 items. Following recommendations by Chan (1998), we selected the organization as referent of the LLWI measure (e.g., "In our organization…"). When assessing the practices, we intended to not rely on the policies that are officially in place within an organization, but instead set out to capture each of the practices by means of its level of implementation throughout the organization. In particular, we asked participants to rate the availability of these practices to them and their colleagues. The LLWI thereby captures the availability and participants' awareness of the practices. Assessing respondents' awareness of organizational practices surpasses the assessment of practices as officially offered by the organization. Officially offered practices are most likely not as effective for the workers as originally intended (Boselie et al., 2005; Wright, 2002). Organizational barriers such as an adverse institutional context or a lack of resources for implementing the practices can significantly shape the extent to which these practices are available for the workers, even if those practices are officially espoused (Nishii et al., 2018). Moreover, workers' awareness of the practices also captures informal work arrangements between older workers and their (local) managers, which might not be officially offered. Thus, to identify both needs for improvement as well as the antecedents and effects of practices, workers' awareness of practices can more validly reflect the actual practices and workers' exposure to them. Third, we asked four experts from the field of human resources management and organizational age management to revise and amend the item pool for overall comprehensiveness. Two of the experts provided detailed written feedback; the other two were interviewed while responding to the questionnaire. On the basis of their valuable input, we revised the item pool and agreed on the item style and response format. In the fourth step, we

tested the item pool to revise items and subscales with poor psychometric characteristics (e.g., skewness and reliability). Given the large number of items, we split the item pool and administered it to four different samples (see next section for details).

Participants. We used four samples to test the items in the item pool. The first sample consisted of 174 employees from three medium-sized organizations in the service sector, logistics industry, and the public sector, respectively. Using this sample, we tested the items developed for o*rganizational climate* and *leadership*. The sample was 57% female; 57% of participants were at least 45 years of age. For these two overarching domains (o*rganizational climate* and *leadership*), it was acceptable to test them in a small number of organizations because we expected within-organization variability between different departments. However, the other domains covering more specific practices can be assumed to be more strongly and directly affected by managerial decisions (Wöhrmann et al., 2018) and are likely to show less within-organization variance. Thus, a second sample of 76 human resource representatives and general managers from different companies was used to test items developed for the domains *work design, knowledge management, continued employment,* and *health and retirement coverage.* The sample was 62% female; 36% of participants were aged 40 and above. It covered service organizations (45%) and industrial organizations (55%), with organizations ranging in size from 23 to 28,000 employees. Third, we administered the items for the *health management* and *individual development* domains to a sample of 38 human resource and health management managers. The sample was 45% female; 39% were aged 40 and above. Again, the sample covered small to very large organizations from the service sector (66%) and industrial sector (34%). Finally, the fourth sample contained 42 human resources managers, among whom we tested the items developed for the *transition to retirement* domain. The sample was 38% female, 67% of the participants were aged 40 and above, and 43% of the participants worked for service organizations, with the remaining 57% in the industrial sector. Again, the sample covered small to very large organizations.

Results

Item style and response format. After acquiring subject matter experts' opinions on the initial questionnaire, we discussed potential response formats. To capture different levels of quality or saturation of the practices within the organization (Boselie et al., 2005), we adopted a multiresponse format. At the same time, to keep the cognitive load for respondents as low as possible (in particular because the experts emphasized that the broad scope of the LLWI can be demanding in itself), we opted for a Likert-type response format that can be repeated across all LLWI domains (DeVellis, 2017). In consideration of an expected skewness of the items for the *organizational climate* and *leadership* domains (cf. other age-related climate measures by e.g., Boehm et al., 2014; Zacher & Yang, 2016), these items had a 7-point response format to capture sufficient detail despite agglomeration of responses on the upper half of the scale (Garner, 1960; Green & Rao, 1970). The other seven domains had a 5-point response format to limit respondents' cognitive load (Weijters et al., 2010)[3]. The response categories ranged from "does not apply at all in our organization" to "does fully apply in our organization." In the preface of the questionnaire we instructed participants to think of both the intensity and the coverage within their organization (Boselie et al., 2005).

Initial item selection. Administration of the initial item pool to the four samples yielded initial evidence on item quality and on necessary modifications. We excluded selected items according to three principles. First, several pairs of items showed intercorrelations in excess of .80. For each of these pairs, we either developed a new overall item replacing the pair or dropped one of the two items to eliminate redundant items in the pool if the item content was

[3] Based on the subject matter experts' response and our experience when seeking for acceptance from organizations to take part in employee surveys a 5-point response format is more accepted than a 7-point format. As the LLWI is intended to be used in organizational level research and as a self-assessment tool in practice, we generally seek for a 5-point response format and consequently conducted the validation studies accordingly.

very similar. Second, items with a high share of "don't know" answers were removed from further analysis. For the LLWI, the share of "don't know" answers is important because it indicates how well practices are communicated within the organization. During scale development, however, items with an extraordinarily high share of "don't know" answers may also indicate unclear wording and irrelevant practices. Third, analysis of reliability for each facet and item-total correlations provided input for the authors' iterative discussions on reducing the pool to a manageable number of items. However, in the process we gave priority to the comprehensive content coverage of the instrument over any gains in reliability. The final LLWI inventory consisted of 102 items.

Study 2: Scale Development

The purpose of Study 2 was twofold: First, we conducted descriptive and exploratory analyses to assess the psychometric properties and the factor structure underlying the developed inventory. From these analyses we selected a subset of items to form the LLWI scales. Second, we provided initial construct validity evidence for the newly developed scales by assessing both convergence with existing age-friendly organizational climate and human resource practices measures and divergence from participants' positive and negative affect as a key source of common rater variance (Podsakoff, Ahearne, & MacKenzie, 1997). Because the LLWI is an intentionally broad construct, we also assessed discriminant validity among the index domains and underlying facets.

Method

Procedure. To achieve as highly diverse a sample as proposed for scale development (Clark & Watson, 1995), we administered the 102 LLWI items obtained from the previous study in an online questionnaire through a panel provider in Germany. Employees aged 25 to 65 were invited to participate. However, we invited employees aged 50 and above with higher frequency to ensure that about half of the sample was of an age at which they could be affected

by the practices researched. Participants had to be working at least 10 hours a week for a single employer with more than 30 employees. We defined the lower limit of 30 employees in line with previous research (e.g., Shaw et al., 1998) to ensure availability of human resource management in general, a prerequisite for meaningful assessment of the LLWI items. Participants were asked to evaluate the newly developed LLWI items with regard to their particular employer. Participants received a three euro incentive for taking part in the research.

Participants. We received responses from 34% of the invited employees, resulting in 609 usable questionnaires[4]. The sample was 55% female, and 32% of participants had a management or supervisory position. Almost half of the participants (48%) were aged 50 and older. Accordingly, the majority of participants (81%) had been working for more than five years for their current employer. The sample was well distributed across small, medium, and large organizations[5] and represented economic sectors in Germany, with almost 50% service organizations, 24% public institutions, and 19% industrial organizations.

Measures. In addition to the 102 LLWI items, age-diversity climate and age-friendly human resource practices with German item sets by Boehm et al. (2014) were measured as organizational level scales. To show convergent validity of the new LLWI measure, we expected the climate measure of these scales to be particularly strongly correlated with the LLWI climate and the leadership domains (i.e., Pearson correlations stronger than .5; Cohen, 1988). For the human resources practices measure, we hypothesized strong positive correlations

[4] We included five attention check questions and conduced outlier analysis to remove participants with careless response patterns as those are frequent, in particular in online surveys (Meade & Craig, 2012). Participants, who answered one of the three easiest questions incorrectly were removed from further analysis (33%). Assessment of the response times revealed that, on average, excluded participants answered 32% faster than participants who passed the attention checks.

[5] The sample included 39% small organizations of 30 to 499 employees, 29% medium sized organizations with 500 to 4,999 employees, and 23% large organizations of 5,000 and more employees.

in particular with the *individual development* domain of the LLWI because most of the items in the practices measure by Boehm et al. (2014) address developmental practices (sample item: "Our company offers equal opportunities to be promoted, transferred, and to make further career steps irrespective of one's age."). But we also expected moderate to strong correlations with the other LLWI domains because organizations typically engage in multiple practices to achieve the same goal, meaning that the LLWI domains of practices should be correlated among each other.

Additionally, we assessed positive and negative affect to show discriminant validity of the newly developed items from participants' individual mood. We used the Positive and Negative Affect Schedule (PANAS) by Watson and Clark (1988) in a shortened 10-item version by Thompson (2007), translated to German by Breyer and Blümke (2016). Because LLWI items are self-rated and assess organizational practices that are generally considered positive (no reverse-coded items), we did not expect the LLWI items to be independent of affect. However, we hypothesized positive affect to be weakly positively correlated and negative affect to be weakly negatively correlated with the LLWI domains ($r < \pm .3$). To measure affect least influenced by the measurement itself, we administered the affect measure first, followed by the LLWI domains and, finally, the scales for convergent validity (Podsakoff et al., 2003). For all measures, reliability was acceptable and can be obtained from Table 3.3.

Analytic strategy. We analyzed data in a four-step process. First, we reviewed the item distributions to drop items showing high skewedness, kurtosis, or a high share of "don't know" answers. Second, we analyzed the factor structure of the remaining items. We tentatively allocated the items to the nine LLWI domains for which they were developed and iteratively conducted explorative factor analyses (EFA) to identify and improve the within-domain factor structure for each domain. To identify and resolve cross-factor structure coefficients between LLWI domains, we conducted EFAs with the overall modified item sets. We then performed a CFA per domain and computed modification indices to further improve derived models and the

unidimensionality of identified factors (cf. Gerbing & Anderson, 1988). Third, we conducted an overall hierarchical CFA to ensure sufficient fit of the overall model prior to cross-validation in Study 3. Fourth, we computed correlations between the developed LLWI scales and the scales for convergent and discriminant validity.

For the EFAs, the factor structure was assessed with oblique rotation and minimum residuals extraction[6]. Oblique rotation was appropriate because the domains and facets measured were expected to be correlated (Worthington & Whittaker, 2006). To determine the number of factors, we used parallel analysis and retained factors with eigenvalues in excess of the 95th percentile of eigenvalues in randomly resampled data (Humphreys & Montanelli Jr., 1975; Longman et al., 1989). In the cases of eigenvalues being close to the cut-off value, we also investigated the scree plot to verify the determined number of factors. Appropriateness of the correlation matrices for factor analysis was ensured by the Kaiser-Meyer-Olkin test and the Bartlett's test of sphericity. To promote unidimensionality of the factor coefficients and to improve reliability, we iteratively removed items that showed either factor structure coefficients above .33 for multiple factors or coefficients below .35 for all factors (Stanton et al., 2001). We further identified strongly intercorrelated items within each factor ($r > .80$ and at least .15 above the average inter-item correlation among the respective factor's items) to avoid redundant items that might affect validity (Clark & Watson, 1995). Taking into account the content coverage of the factor, the item characteristics, and the factor structure coefficients for the items, we retained only one item per pair in these cases. On the basis of the CFA results, we systematically analyzed within-factor or between-factor covariance of residuals and iteratively solved the cases of insufficient model fit by dropping selected items.

[6] Since the data is partially skewed and non-normally distributed, a minimum residuals extraction was more appropriate than a maximum likelihood extraction (Briggs & MacCallum, 2003; Zygmont & Smith, 2014).

Results

Item distributions. For the 102 items, skewness ranged from -1.09 to 1.10 and kurtosis from -1.44 to 1.15. Results did not exceed recommended thresholds (Curran et al., 1996), so that all items were retained for further analysis. Furthermore, missing value analysis revealed that on average 57 participants (9%) chose the "don't know" answer option. The share of "don't know" answers was higher for the retirement-related domains *transition to retirement* and *continued employment* (19%) than for the more general domains (6%). This indicates that—particularly for retirement-related practices—organizations lack proper communication of the practices so that workers are partially unaware of their organization's offerings. For three items the number of "don't know" answers was extraordinary high (larger than three times the innerquartile range above the median number of missing values in the items' domain). To promote applicability and ease of completion of the measure, we dropped these three items[7] from further analysis. A full list of item characteristics can be obtained from Appendix G.

Domain level factor analyses. Following our analytic strategy, we analyzed each of the nine LLWI domains individually. After multiple iterations of EFA and CFA per domain, we then removed a further 19 items in total to achieve obliquely rotated factor solutions without cross-factor structure coefficients in excess of .33 and acceptable model fit in CFA for each domain.

Organizational climate for the aging workforce. The developed item pool contained 12 items assessing an age-friendly organizational climate. Parallel analysis revealed three factors explaining 24%, 24%, and 20% of the variance, respectively. In the rotated EFA

[7] The three items covered to what extend managers are specifically prepared for dealing with older employees (e.g., training), whether employees may take additional unpaid leave at certain intervals, and whether the organization offers its employees immediate financial support in case of family and private emergencies (e.g., advance on salary).

solution, four items loaded primarily on the first factor, four items on the second, and three items on the third. One item showed factor structure coefficients in excess of .33 for the first and the third factor (.47 and .37) and was therefore dropped from further analysis. Moreover, two items from the first factor were highly correlated (r = .87; factor structure coefficients of .93 and .95) and addressed a very similar aspect, so that we dropped the item with a lower content contribution. The remaining 10 items were simply structured, with primary factor structure coefficients ranging from .78 to .95.

CFA yielded good fit of the three-factor model (χ^2 = 99; df = 32; *RMSEA* = .06; *CFI* = .98). Content-wise, the three-factor structure complies with the three facets proposed by the qualitative framework of the LLWI: *positive image of age* (first factor, three items), *open and target group-oriented communication* (second factor, four items), and *equality of opportunity* (third factor, three items). We formed three scales, which showed good internal consistency of .88, .90, and .90, respectively.

Leadership for the aging workforce. The item pool contained eight items describing a leadership style characterized by appreciation of all age groups and responsiveness to workers' individual needs. Parallel analysis suggested two factors. However, four pieces of deviating evidence led us to determine a one-factor solution as most appropriate. First, the scree plot showed a flat plateau starting at the second factor. Second, the two factors identified by oblique rotation showed a strong correlation of .87. Third, the second factor explained four percent of the variance only, compared with 74% for the first factor. Fourth, only two items loaded highest on the second factor. Contrary to the qualitative framework that proposed two facets, we thus proceeded with a one-factor solution. In addition, two items were highly correlated (r = .90), so that we dropped the one with lower content contribution. The subsequent EFA with the remaining items revealed sufficiently high factor loadings for the single-factor solution, which ranged between .77 and .90. CFA for the one-factor solution revealed further need of improvement (χ^2 = 93; df = 14; *RMSEA* = .10; *CFI* = .98). Systematic analysis of the

residuals disclosed positive covariance of residuals for two items (indicating redundancy), so that we again dropped the item with the lower content contribution. An additional CFA with the remaining items showed acceptable model fit (χ^2 = 31; *df* = 9; *RMSEA* = .06; *CFI* = .99). With the remaining 6 items, we formed a scale that showed an excellent internal consistency of .95.

Work design for the aging workforce. The work design domain was represented by 16 items from the item pool. Parallel analysis revealed four factors explaining 19%, 15%, 14%, and 6% of the variance, respectively. In the rotated EFA solution, five items loaded primarily on the first factor, four items on the second, four items on the third, and three items on the fourth. The rotated solution did not reveal any cross-factor structure coefficients in excess of .33. The primary factor structure coefficients ranged from .43 to .89.

However, the EFA of the overall item set across all nine LLWI domains revealed substantial covariance of two items with other domains of the LLWI. The first, an item addressing ergonomic work design, showed covariance with the *health management* domain. The second, an item addressing managers' consideration of older workers' individual capabilities while designing their work, showed covariance with the *leadership* domain. To support discriminant validity among developed scales, we removed both items from further analysis.

Reassessment of the EFA with the remaining 14 items supported the four-factor solution. We thus formed four scales accordingly and conducted a CFA. Results showed good fit of the four-factor model (χ^2 = 257; *df* = 71; *RMSEA* = .07; *CFI* = .96). Moreover, the four-factor structure complies with the four facets proposed by the qualitative framework of the LLWI content-wise: *ergonomic working conditions* (first factor, four items), *work according to capabilities* (second factor, three items), *flexible work time arrangements* (third factor, four items), and *flexible work places* (fourth factor, three items). The scales showed good to adequate internal consistency of .86, .86, .78, and .84, respectively.

Health management for the aging workforce. We obtained 12 items for the health management domain from Study 1. Parallel analysis revealed three factors explaining 28%, 19%, and 7% of the variance, respectively. In the rotated EFA solution, six items loaded primarily on the first factor, four items on the second, and two items on the third. The primary factor structure coefficients ranged from .49 to .88. The rotated solution did not reveal any cross-factor structure coefficients in excess of .33. However, the CFA yielded an insufficient model fit (χ^2 = 303; *df* = 51; *RMSEA* = .09; *CFI* = .95) and the third factor incurred primary factor loadings for two items only. To resolve the issue, we removed three items from the scales. First, we removed one redundant item that was .84 correlated with a second item from the same factor (coefficients of .88 and .86) and contributed less content to the overall scale. Second, systematic analysis revealed covariance of residuals of three items from the first factor with the third factor. We dropped two of these items, reallocated the third item to the third factor, where it better fit content-wise, and increased this factor's number of items to three. CFA reassessment yielded an acceptable fit of the three-factor model (χ^2 = 114; *df* = 24; *RMSEA* = .08; *CFI* = .97). The three-factor structure complies with the three facets proposed by the qualitative framework of the LLWI content-wise: *health promotion* (first factor, three items), *availability of physical exercise and nutrition opportunities* (second factor, three items), and *workplace medical treatment* (third factor, three items). The scales showed good to adequate internal consistency of .90, .82, and .77, respectively.

Individual development for the aging workforce. Parallel analysis of the *individual development* domain (13 items) suggested four factors explaining 16%, 13%, 13%, and 11% of the variance, respectively. However, the third and fourth factor incurred rotated structure coefficients in excess of .35 for only two items each. Both item pairs were intercorrelated by .87 and .72, respectively, indicating potentially redundant items and an overly narrow operationalization (Clark & Watson, 1995). Moreover, two items showed cross-factor structure coefficients in excess of .33 for the first and second factor, impeding the achievement of simple

structure. All this evidence suggested tentatively pursuing a single-factor solution for the *individual development* domain. Factor structure coefficients ranged from .67 to .81.

Moreover, factor analyses of the overall item set across all nine LLWI domains revealed substantial covariance of two items with other domains of the LLWI. One item addressing the organization's development support for older workers compared with younger workers showed covariance with the *equality of opportunities* factor of the *organizational climate* domain. Another item, addressing managers' engagement in individual development planning, showed covariance with the *leadership* domain. To support discriminant validity among developed scales, we removed both items from further analysis.

CFA for the one-factor solution revealed further need of improvement (χ^2 = 706; *df* = 44; *RMSEA* = .16; *CFI* = .84). Systematic analysis of the residuals showed positive covariance of residuals for three pairs of highly correlated items (indicating redundancy). For each pair, we dropped the items with the lowest content contribution. An additional CFA with the remaining eight items showed acceptable model fit (χ^2 = 84; *df* = 20; *RMSEA* = .07; *CFI* = .97). Using these items, we formed a scale that showed a good internal consistency of .90.

Knowledge management for the aging workforce. The item pool contained eight items for the operationalization of age-friendly knowledge practices. Parallel analysis disclosed two factors explaining 27% and 24% of the variance, respectively. In the rotated EFA solution, four items loaded primarily on the first factor and four items on the second. One item loading primarily on the second factor also showed structure coefficients in excess of .33 for the first factor. However, removal of the item caused cross-factor structure coefficients for two further items. Thus, we formed two scales allocating the items according to their primary structure coefficients and the CFA results. Systematic analysis of the residuals revealed positive covariance among two items from the second factor, of which one also had a very low communality of .31 in the EFA ("younger and older employees work together a lot"). Reassessment of the EFA without that item did not show any cross-factor structure coefficients

exceeding .33. The same four items as in the initial EFA loaded highest on the first factor (structure coefficients of .40 to .78), the other three items on the second (structure coefficients of .69 to .88). Reassessment of the CFA indicated acceptable model fit (χ^2 = 67; *df* = 13; *RMSEA* = .08; *CFI* = .97). The two-factor structure complies with the two facets proposed by the qualitative framework of the LLWI content-wise: *institutionalized knowledge transfer* (first factor, four items) and *intergenerational collaboration* (second factor, three items). The scales showed good internal consistency of .80 and .88, respectively.

Transition to retirement for the aging workforce. The *transition to retirement* domain was represented by 16 items from the item pool. Parallel analysis revealed four factors explaining 20%, 14%, 14%, and 13% of the variance, respectively. In the rotated EFA solution, four items loaded primarily on the first factor, four items on the second, four items on the third, and three items on the fourth. One item showed factor structure coefficients in excess of .33 for the second and the third factor. We thus dropped that item from further analysis. Moreover, one item was removed following the subsequent factor analysis of the overall item set across all nine LLWI domains. This item, loading on the third *transition to retirement* factor, showed substantial covariance with the *continued employment* domain and thus impeded discriminant validity between developed scales. Reassessment of the EFA with the remaining 14 items supported the four-factor solution found previously. Thus, we formed four scales. CFA results show good fit of the four-factor model (χ^2 = 202; *df* = 71; *RMSEA* = .06; *CFI* = .98). Moreover, the four-factor structure complies with the four facets proposed by the qualitative framework of the LLWI content-wise: *continuous inclusion and maintaining contact* (first factor, four items), *counseling for retirement life preparation* (second factor, three items), *phased retirement and individualized transition solutions* (third factor, four items), and *timely transition planning* (fourth factor, three items). The scales showed good to excellent internal consistency of .94, .92, .86, and .89, respectively.

Continued employment for the aging workforce. The item pool contained eight items for the *continued employment* domain. Parallel analysis revealed three factors explaining 31%, 23%, and 4% of the variance, respectively. Given the limited explanatory contribution of the third factor and that none of the items showed primary structure coefficients for the third factor in the rotated solution, we nevertheless determined a two-factor solution to be more appropriate. In the two-factor rotated solution, four items loaded primarily on the first factor (structure coefficients of .59 to .90) and four items on the second (structure coefficients of .48 to .87). No items showed cross-factor structure coefficients in excess of .33. Accordingly, we formed two scales and computed CFA. Results revealed further need of improvement (χ^2 = 114; *df* = 19; *RMSEA* = .09; *CFI* = .94). Systematic analysis of the residuals showed positive covariance of residuals for one item of the second factor with the first factor. Because the item does not contribute indispensable content, we dropped it from further analysis. An additional CFA with the remaining items showed acceptable model fit (χ^2 = 68; *df* = 13; *RMSEA* = .08; *CFI* = .96). Content-wise, the two-factor structure complies with the two facets proposed by the qualitative framework of the LLWI: *individualized employment options* for workers at retirement age (first factor, four items) and *(re-)hiring of older workers* (second factor, three items). The scales showed good to adequate internal consistency of .86 and .76, respectively.

Health and retirement coverage for the aging workforce. The *health and retirement coverage* domain was operationalized by six items. Parallel analysis revealed two factors explaining 33% and 28% of the variance, respectively. In the rotated EFA solution, three items loaded primarily on the first factor (structure coefficients of .60 to .90) and two items on the second (structure coefficients of .72 and .97). We removed one item showing factor structure coefficients in excess of .33 for both factors. Accordingly, we tentatively formed two scales and conducted CFA. Results revealed an unacceptably high root mean squared error (*RMSEA* = .11). After item removal during EFA, the two-factor submodel comprised five items only. Analysis of residuals showed error covariance between two out of three items from the first

factor. To retain at least three items for the factor's subscale while also resolving the error covariance, one of the covaried items was replaced by an item that was initially dropped because of cross-factor structure coefficients. The cross-factor structure coefficients for the picked-up item did not persist when re-examining EFA with the modified item set[8]. These modifications made in response to the CFAs required us to re-examine the EFAs with the reduced item set, but the previously reported simple factor structure was not affected. An additional CFA with the revised item set showed acceptable model fit (χ^2 = 8; *df* = 4; *RMSEA* = .04; *CFI* = .998). However, because the second factor now only comprised two items, CFA yielded a wide confidence interval for the *RMSEA*. Content-wise, the two-factor structure complies with the two facets proposed by the qualitative framework of the LLWI: *retirement savings and pensions* (first factor, three items) and *insurances* (second factor, two items). Financial emergency support, as described by the qualitative framework, could not be operationalized because the respective item did not fit the scale. The scales showed good internal consistency of .88 and .86, respectively.

Overall confirmatory factor analysis. In the wake of our analyses, 80 items remained in the LLWI item set (see Appendix E for the selected German items and Appendix D for the English translation). We integrated the developed models for the nine LLWI domains into an overall hierarchical model with second-order latent variables for the nine domains and the 22 identified factors as first-order latent variables. CFA yielded an acceptable model fit (χ^2 = 6310; *df* = 3024; *RMSEA* = .04; *CFI* = .91). Furthermore, we computed scale means for each first-order construct and estimated a CFA with the second-order model only. CFA likewise supports

[8] We acknowledge potential impairments for discriminant validity between the two factors within the *health and retirement coverage* domain due to including an item, which initially showed cross-factor structure coefficients.

Table 3.2

Confirmatory factor analysis and reliability results Study 2

Model	Number of subscales	Number of items	α (all items)	α (first order scales)	Chi-Square	*df*	*RMSEA*	*RMSEA* 90% *CI*	*CFI*	*SRMR*
1. Organizational climate	3	10	.92	.88 - .91	99.5	32	.06	[.05, .07]	.98	.02
2. Leadership	1	6	.95		37.2	9	.07	[.05, .10]	.99	.01
3. Work design	4	14	.91	.78 - .86	256.9	71	.07	[.06, .07]	.96	.04
4. Health management	3	9	.92	.77 - .90	114.1	24	.08	[.06, .09]	.97	.03
5. Individual development	1	8	.90		84.4	20	.07	[.06, .09]	.97	.03
6. Knowledge management	2	7	.89	.80 - .88	67.0	13	.08	[.06, .10]	.97	.03
7. Transition to retirement	4	14	.94	.86 - .94	202.5	71	.06	[.05, .07]	.98	.03
8. Continued employment	2	7	.84	.76 - .86	68.0	13	.08	[.07, .11]	.96	.05
9. Health and retirement coverage	2	5	.90	.86 - .87	7.6	4	.04	[.00, .08]	1.00 (.998)	.01
10. Overall hierarchical model		80			6309.6	3024	.04	[.04, .04]	.91	.07
11. Second-order model with first-order scale means					1587.1	491	.06	[.06, .06]	.92	.06

Note. N = 609. α = Cronbach's alpha; *RMSEA* = root mean squared error of approximation; *CI* = confidence interval; *CFI* = comparative fit index; *SRMR* = standardized root mean square residual.

the overall model's fit (χ^2 = 1587; *df* = 491; *RMSEA* = .06; *CFI* = .92). As a tentative initial configuration, we thus formed the LLWI as a set of nine measures accordingly. All subscales yielded coefficient alphas above .76 and item-total correlations above .72[9]. We computed nine latent variables by averaging the factor means for each of the nine LLWI domains. The variance inflation factors of the computed variables ranged from 1.79 to 3.66 (M = 2.67; SD = 0.60); multicollinearity thus did not appear to be a major concern[10]. All this evidence suggested that the qualitatively derived model of the LLWI with nine domains was most appropriate. Table 3.2 summarizes the fit statistics, including *RMSEA* confidence interval and *SRMR* for the overall model, the second order model, and each submodel, as well as reliabilities of the scales (for detailed reliability results, see Appendix I).

[9] Due to short scales of two to nine items we did not correct the item-total correlation by dropping the respective item prior to averaging the scale. Doing so results in a minimum item-total correlation of .54.

[10] The highest VIFs were observed for the domains *transition to retirement* and *individual development*.

Convergent and divergent validity evidence. For the present study, our validity goals were to assess divergence of the LLWI scales from affect and their convergence with existing measures for age-friendly organizational climate and human resource practices. A correlation matrix among the second-order LLWI measures and all validation measures appears in Table 3.3.

To assess convergent validity of LLWI measures, we correlated them to two established measures: age-diversity climate and age-diversity human resource practices (Boehm et al., 2014). With regard to age-diversity climate, we found an average correlation of .55 and a median correlation of .58 to the nine LLWI measures. Correlation was weakest for *health management* (.33) and, as hypothesized, strongest for *organizational climate* and *leadership* (both .73). With regard to age-diversity human resource practices, we found an average correlation of .58 and a median correlation of .60 to the nine LLWI measures. The weakest correlation was again observed for *health management* (.43), and the strongest correlation for the *individual development* (.73) domain. All this evidence indicates good convergent validity for the new measures.

We then evaluated divergent validity of our measures by inspecting their correlations with positive affect and negative affect. All nine LLWI measures were sufficiently independent of both positive affect (.24 average correlation; .26 median correlation) and negative affect (-.13 average correlation; -.15 median correlation). The highest correlation (.29) was observed between the LLWI measure *knowledge management* and positive affect. Results suggest sufficient discriminant validity regarding neurotic traits (Clark & Watson, 1995) and resilience regarding common method biases (Podsakoff et al., 1997).

Study 3: Cross-Validation

The purpose of Study 3 was to cross-validate results from Study 2 and provide additional validity evidence for the developed LLWI measurement model. We performed CFAs to further

Table 3.3

Means, standard deviations, and correlations Study 2

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Avg. *r* with LLWI (1-9)
1. Organizational climate	5.15	1.21	(.92)																			
2. Leadership	4.82	1.35	.71**	(.95)																		
3. Work design	2.76	0.84	.40**	.59**	(.91)																	
4. Health management	2.53	1.04	.26**	.40**	.60**	(.92)																
5. Individual development	3.24	0.91	.65**	.72**	.64**	.55**	(.90)															
6. Knowledge management	2.98	1.01	.49**	.60**	.56**	.56**	.72**	(.89)														
7. Transition to retirement	2.37	1.02	.43**	.55**	.66**	.70**	.67**	.65**	(.94)													
8. Continued employment	3.19	0.87	.55**	.50**	.38**	.35**	.57**	.50**	.55**	(.84)												
9. Health and retirement coverage	2.63	1.13	.33**	.44**	.57**	.67**	.56**	.53**	.75**	.41**	(.90)											
10. Positive affect	3.32	0.76	.26**	.29**	.26**	.17**	.23**	.29**	.28**	.20**	.21**	(.84)										.24
11. Negative affect	1.42	0.56	-.24**	-.23**	-.15**	-.05	-.17**	-.17**	-.08	-.07	-.04	-.17**	(.82)									-.13
12. Age-diversity climate	3.72	0.89	.73**	.73**	.43**	.33**	.72**	.58**	.48**	.58**	.38**	.20**	-.23**	(.87)								.55
13. Age-inclusive HR practices	3.18	0.95	.69**	.66**	.49**	.43**	.73**	.60**	.55**	.60**	.46**	.21**	-.16**	.77**	(.89)							.58
14. Participants' chronological age	47.23	10.14	-.06	-.09*	-.12**	-.02	-.15**	-.03	-.23**	-.14**	-.11**	.04	-.08	-.06	-.10*	-						-.10
15. Participant is manager/ supervisor	0.32	0.47	.10*	.18**	.14**	.08	.17**	.23**	.18**	.11*	.10*	.30**	-.03	.11**	.11**	-.02	-					.14
16. Organizational size	6.43	1.60	-.00	.00	.06	.25**	.09*	-.01	.17**	-.05	.22**	.04	.01	.02	.08	-.01	-.07	-				.08
17. Industry dummy: service organization	0.42	0.49	.02	.02	.03	.00	.02	.04	-.03	.11*	.05	.01	.03	-.02	.01	.03	.04	-.08	-			.03
18. Industry dummy: industrial organization	0.20	0.40	-.04	.04	.10*	.10*	.03	.08	.12*	-.03	.08	-.05	-.03	.03	.00	-.01	.05	.09*	-.42**	-		.05
19. Industry dummy: public organization	0.25	0.43	-.03	-.07	-.06	-.04	-.05	-.09*	-.09	-.15**	-.10*	-.03	.02	-.05	-.06	.02	-.12**	.01	-.49**	-.29**	-	-.07

Note. $N = 609$. *M* and *SD* are used to represent mean and standard deviation, respectively. *Organizational size* is measured as number of employees on an ordinal nine point scale ranging from "less than 10" to "50,000 and above". Internal consistency coefficients, Cronbach's alphas are reported in the parentheses on the diagonal. * indicates $p < .05$. ** indicates $p < .01$.

improve confidence in the measurement models identified in Study 2. To assess criterion validity of the LLWI scales, we included individual level outcome measures such as health status, stress level, and work engagement, and narrowed the studied population to the target group of measured practices, employees aged 50 and older.

Method

Procedure. An online questionnaire was administrated to employees aged 50 and beyond by a panel provider in Germany to capture a highly diverse sample of different organizations. To lower the risk of sampling participants that had already participated in Study 2, we selected a different panel provider. Analysis of potential duplicates revealed duplicate sociodemographic and employer characteristics for three percent of participants only. Participants received a three-euro incentive after completion of the questionnaire.

Participants. We received 349 useable responses[11] at a response rate of 35%. Participants were 57% male, and 40% had a management or supervisory position. With an age range of 50 to 67 (67 is the standard retirement age in Germany), the sample comprised employees nearing retirement age and likely to benefit from the LLWI practices. Moreover, we restricted participants to those working at least 32 hours per week for the evaluated employer to ensure sufficient exposure to the LLWI practices and relevance to their everyday life. Over half of the participants had been working for their employer for more than 20 years, which is typical for that age group in Germany. Similar to Study 2, the sample represented economic

[11] Data was cleaned as in Study 2. Twenty percent of participants failed to answer the attention checks correctly. Analysis of response times supported removal of those participants as they answered 32% faster than attentive participants on average.

sectors in Germany (service sector: 45%, public sector: 24%, industry: 17%) and comprised a broad range of organizational sizes[12].

Measures. In addition to the LLWI scales, criterion measures were administered that captured various criteria for successful integration of older employees into the workforce. We hypothesized moderate correlations between those criterion measures and the LLWI domains (i.e., Pearson correlations stronger than .3; Cohen, 1988) because the LLWI domains were supposed to capture a small part of all potential antecedents of our criteria. First, person-organization fit and person-job fit were measured by three items each, developed by Cable and DeRue (2002; sample item: "The job that I currently hold gives me just about everything that I want from a job."). These measures are suitable criteria for the assessment of the LLWI measure because person-environment fit has been argued to be an important driver of post-retirement work (cf. theory of work adjustment; Dawis & Lofquist, 1984; Harper & Shoffner, 2004) and also because this fit is an important outcome for sustainable employment and job satisfaction of older workers (Kooij, 2015; Rauvola et al., 2019). Second, we measured general well-being with five items from the World Health Organization (Topp et al., 2015; sample item: "Over the past two weeks I have felt cheerful and in good spirits.") and work engagement with a three-item version of the Utrecht Work Engagement scale (Sautier et al., 2015; Schaufeli et al., 2019; sample item: "I am enthusiastic about my job."). Self-rated perceived health status was assessed with four items from Adams and Beehr (1998; sample item: "My health is better than most people my age."). In addition, we assessed work ability, because it is an important mediator between work characteristics and work outcomes (Cadiz et al., 2019). We used a self-rated perceived work ability measure comprised of four items by McGonagle et al. (2015). The four items address work ability with respect to physical, mental, interpersonal, and overall work

[12] The sample included 36% small organizations of 30 to 499 employees, 28% medium sized organizations with 500 to 4,999 employees, 36% large organizations of 5,000 and more employees.

demands on a 10-point scale (sample item: “Thinking about the physical demands of your job, how do you rate your current ability to meet those demands?”). In addition, we asked participants about the number of days they had been absent from work due to illness in the last twelve months (WAI-Netzwerk, 2015). For the number of sick days, we did not expect the correlations with the LLWI to be high. However, even a weak effect (i.e., Pearson correlations stronger than .1; Cohen, 1988) reducing the number of sick days is of high importance. Third, we assessed occupational future time perspective (Zacher & Frese, 2009; sample item: “My occupational future is filled with possibilities.”), for which we hypothesized a moderate correlation with the LLWI domains. Fourth, researchers have emphasized the role of work-related stressors in the employment of older workers (Barnes-Farrell, 2005). To validate a negative correlation of the LLWI measures with stress, we administered the Stress in General Scale measuring two facets of work-related stress, pressure and threat (Stanton et al., 2001; sample item: “How is your job most of the time? Nerve-wracking.”). We hypothesized moderated negative correlations with threat because high levels of the LLWI domains—in particular the leadership domain—should correspond to less stressors for older workers (Boehm & Dwertmann, 2015). For pressure, our expectations were indecisive. On the one hand, organizational practices may support older workers in coping with stressful situations. On the other hand, they might induce additional stress if perceived as an additional burden to the worker.

In addition to situational perception criteria, we included two behavioral intention measures. First, we measured turnover intentions with three items (sample item: “I occasionally think about leaving this organization.”). These three items were selected by Hansung and Stoner (2008) from a four-item scale initially developed by Nissly, Mor Barak, and Levin (2005). Given the study background with the relatively strict standard retirement age of 65 to 67 in Germany, we then administered three items developed by Wöhrmann, Deller, and Wang (2013) to measure participants’ intention to continue working for their current employer after

becoming eligible for retirement (sample item: "I would like to continue to work for my current employer in retirement."). We hypothesized the LLWI domains to be negatively associated with turnover intentions and positively associated with the intention to continue working after reaching retirement age. However, we expected these to be only weak correlations because the relation between organizational practices and attitudinal reactions such as work engagement is more proximal and is a prerequisite for subsequent behavioral intentions and, ultimately, actions (Nishii et al., 2018).

Finally, we administered the Nordic Age Discrimination Scale (Furunes & Mykletun, 2010) to complement convergent validity assessments started in Study 2 (sample item: "Elderly workers do not have equal opportunities for training during work time"). Here we expected the measure to correlate strongly with the climate domain of the LLWI but also moderately with the other domains, as age-friendly climate is interrelated with organizational practices (Boehm et al., 2014; Zacher & Yang, 2016). All measures not available in German language were translated using back-translation (Brislin, 1970). For all measures, reliability was acceptable and can be obtained from Table 3.5.

Results

Building on our results from Study 2, we formed nine hierarchical measurement models with 80 items in total. The first-order scales showed alpha reliabilities between .73 and .93, very similar to the values obtained from Study 2 (for detailed reliability results, see Appendix I). The means of the first-order scales were again averaged per domain to form the nine domain-level latent measures. Analysis of missing value pattern showed that a larger share of participants was able to answer items in the retirement-related domains (88% versus 81% in Study 2).

Confirmatory factor analysis. To cross-validate the measurement models, we conducted individual CFAs for the nine LLWI domains and an overall hierarchical model as in

Table 3.4

Confirmatory factor analysis and reliability results Study 3

Model	Number of subscales	Number of items	α (all items)	α (first order scales)	Chi-Square	*df*	*RMSEA*	*RMSEA* 90% *CI*	*CFI*	*SRMR*
1. Organizational climate	3	10	.93	.89 - .91	86.0	32	.07	[.05, .09]	.98	.02
2. Leadership	1	6	.95		48.7	9	.11	[.08, .14]	.98	.02
3. Work design	4	14	.90	.77 - .86	180.9	71	.07	[.05, .08]	.96	.04
4. Health management	3	9	.91	.76 - .87	89.9	24	.09	[.07, .11]	.96	.04
5. Individual development	1	8	.92		71.7	20	.09	[.06, .11]	.97	.03
6. Knowledge management	2	7	.90	.83 - .87	38.5	13	.07	[.05, .10]	.98	.02
7. Transition to retirement	4	14	.94	.84 - .93	228.1	71	.08	[.07, .09]	.96	.04
8. Continued employment	2	7	.86	.72 - .89	30.3	13	.06	[.03, .09]	.98	.03
9. Health and retirement coverage	2	5	.91	.86 - .89	7.4	4	.05	[.00, .10]	1.00 (.998)	.01
10. Overall hierarchical model		80			5348.1	3024	.05	[.04, .05]	.89	.07
11. Second-order model with first-order scale means					1240.4	491	.07	[.06, .07]	.92	.05

Note. $N = 349$. α = Cronbach's alpha; *RMSEA* = root mean squared error of approximation; *CI* = confidence interval; *CFI* = comparative fit index; *SRMR* = standardized root mean square residual.

Study 2. The fit statistics appearing in Table 3.4 were acceptable for six out of the nine submodels (*RMSEA* < .08; *CFI* > .95). However, the single-factor models for *leadership* (6 items) and *individual development* (9 items) yielded an *RMSEA* of .11 and .09, respectively, indicating some redundancy among the items. Nevertheless, the good *CFI* (> .97) and *SRMR* (< .03) of both models suggested that the proposed scales did not need to be modified. Moreover, the *health management* domain showed an *RMSEA* of .09, indicating opportunity for improvement in future studies. All 22 subscales showed alpha coefficients above .72, and item-total correlations were above .71. Cross-validation of the overall hierarchical model yielded acceptable fit statistics as well. An *RMSEA* of .05 and a *CFI* of .89 suggested that the nine-domain model generally holds across studies.

Criterion validity evidence. Clear patterns emerged from the correlation matrix that appears in Table 3.5. As expected, all nine LLWI domains were moderately positively correlated with person-job fit (r = [.38; .61]), person-organization fit (r = [.43; .65]), participants' work engagement (r = [.32; .47]), and their well-being (r = [.31; .48]).

Table 3.5 *Means, standard deviations, and correlations Study 3*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	Avg. *r* with LLWI (1-9)
1. Organizational climate	5.22	1.24	(.93)																								
2. Leadership	4.73	1.44	.75**	(.95)																							
3. Work design	2.68	0.82	.45**	.59**	(.90)																						
4. Health management	2.57	1.02	.30**	.45**	.59**	(.91)																					
5. Individual development	3.20	0.99	.69**	.76**	.67**	.55**	(.92)																				
6. Knowledge management	3.13	1.01	.51**	.63**	.55**	.57**	.81**	(.90)																			
7. Transition to retirement	2.26	0.98	.46**	.59**	.63**	.62**	.66**	.65**	(.94)																		
8. Continued employment	3.14	0.97	.46**	.47**	.42**	.29**	.49**	.50**	.56**	(.86)																	
9. Health and retirement coverage	2.61	1.20	.28**	.44**	.54**	.65**	.50**	.53**	.66**	.39**	(.91)																
10. Person-job fit	4.94	1.47	.44**	.61**	.43**	.38**	.54**	.55**	.47**	.43**	.40**	(.94)															.47
11. Person-organization fit	4.52	1.55	.50**	.65**	.48**	.46**	.61**	.57**	.53**	.43**	.46**	.75**	(.97)														.52
12. Well-being	3.04	1.18	.31**	.45**	.38**	.41**	.45**	.47**	.48**	.35**	.36**	.59**	.52**	(.93)													.41
13. Perceived health	4.08	1.03	.24**	.26**	.24**	.29**	.32**	.31**	.27**	.25**	.21**	.35**	.33**	.56**	(.90)												.27
14. Perceived workability	8.07	1.77	.34**	.33**	.27**	.24**	.31**	.28**	.28**	.27**	.16**	.40**	.34**	.56**	.69**	(.87)											.28
15. Sickness absence days	15.28	38.41	-.21**	-.20**	-.18**	-.10	-.19**	-.13*	-.14*	-.20**	-.07	-.19**	-.15**	-.25**	-.41**	-.57**	-										-.16
16. Stress: threat	1.67	0.54	-.39**	-.47**	-.32**	-.26**	-.38**	-.33**	-.37**	-.34**	-.22**	-.56**	-.44**	-.57**	-.35**	-.47**	.20**	(.80)									-.34
17. Stress: pressure	2.27	0.64	-.27**	-.35**	-.30**	-.16**	-.24**	-.15**	-.24**	-.17**	-.09	-.27**	-.27**	-.38**	-.23**	-.32**	.11*	.67**	(.86)								-.22
18. Work engagement	4.98	1.03	.35**	.46**	.32**	.34**	.45**	.47**	.45**	.39**	.40**	.72**	.59**	.59**	.33**	.39**	-.15**	-.41**	-.11*	(.82)							.40
19. Occupational future time perspective	2.60	1.14	.19**	.31**	.25**	.25**	.33**	.33**	.37**	.38**	.35**	.33**	.29**	.35**	.23**	.17**	-.16**	-.19**	-.06	.39**	(.80)						.31
20. Turnover intentions	2.48	1.64	-.33**	-.34**	-.25**	-.25**	-.32**	-.22**	-.24**	-.24**	-.21**	-.37**	-.37**	-.25**	-.22**	-.36**	.13*	.34**	.17**	-.27**	-.08	(.74)					-.27
21. Post-retirement work intentions	2.51	1.23	.20**	.28**	.19**	.10	.21**	.23**	.22**	.43**	.15*	.38**	.38**	.21**	.10	.18**	-.08	-.26**	-.08	.35**	.29**	-.28**	(.90)				.22
22. Positive affect	3.43	0.69	.21**	.26**	.23**	.29**	.33**	.36**	.33**	.36**	.30**	.39**	.32**	.51**	.37**	.33**	-.19**	-.19**	-.02	.51**	.23**	-.14**	.21**	(.80)			.30
23. Negative affect	1.35	0.49	-.17**	-.16**	-.07	-.06	-.10	-.19**	-.17**	-.09	-.06	-.26**	-.20**	-.37**	-.20**	-.30**	.20**	.37**	.25**	-.18**	-.11*	.08	-.05	-.09	(.77)		-.12
24. Nordic age discrimination scale	2.10	0.87	-.56**	-.47**	-.30**	-.20**	-.48**	-.34**	-.35**	-.36**	-.20**	-.31**	-.32**	-.18**	-.20**	-.25**	.11	.30**	.27**	-.22**	-.11*	.25**	-.17**	-.09	.09	(.81)	-.36
25. Participants' chronological age	55.87	4.12	.04	.06	.11*	.11*	.11*	.10	.08	-.03	.11	.06	.07	.06	-.08	-.07	.03	.04	-.02	.01	-.21**	.08	.11*	.04	-.03	-.01	.08

Note. $N = 349$. *M* and *SD* are used to represent mean and standard deviation, respectively. Internal consistency coefficients, Cronbach's alphas are reported in the parentheses on the diagonal.
* indicates $p < .05$. ** indicates $p < .01$.

Furthermore, we found positive correlations with participants' perceived health status (r = [.21; .32]) and their occupational future time perspective (r = [.19; .38]). However, the hypothesized moderate relation was not found for all nine domains. As expected, the behavioral intention criteria showed weak to moderate correlations: The LLWI domains were positively correlated with participants' intentions to continue working for their employer after becoming eligible for retirement (r = [.10; .43]) and negatively correlated with participants' intention to quit their job (r = [-.21; -.34]). For the number of days that participants had been absent from work due to illness we found a weak negative relationship (r = [-.10; -.21]), which is nevertheless important. Finally, threat from the job was moderately negatively related with most of the domains (r = [-.22; -.47]). Only for *health management* and *health and retirement coverage* was the association weaker than expected. Moreover, the LLWI domains were weakly negatively related to pressure from the job (r = [-.15; -.35]).

Convergent and divergent validity evidence. Study 3 provided additional convergent validity evidence. Results showed that all nine LLWI scales were negatively correlated with the Nordic Age Discrimination Scale (NADS) by Furunes and Mykletun (2010). Among the LLWI domains, age discrimination was most represented by *organizational climate*, which comprises the *equality of opportunity* factor. Thus, we expected the NADS to be most strongly correlated with that domain. Results confirmed our expectations (r = -.56). However, the NADS was also strongly correlated with *individual development* (r = -.48) and *leadership* (r = -.47). Moreover, results of Study 3 confirmed convergent validity evidence regarding the human resources practices and age-diversity climate measure by Boehm et al. (2014), as had already been shown in Study 2.

Study 3 results also confirmed independence of the LLWI scales from negative affect (r = [-.06; -.19]), which supported the LLWI measures' discriminant validity. For positive affect, correlations were slightly higher than in Study 2 (r = [.21; .36]), indicating a minor impairment due to common method variance within the study. Moderate correlations above .3

were found between positive affect and the domains *individual development*, *knowledge management*, and the three retirement-related domains. However, even for *knowledge management* (r = .36) the variance shared with positive affect was below 15%. The correlation table can be obtained from Table 3.5.

Discussion

This article describes the development of the LLWI, a comprehensive, multifaceted measure of organizational practices for an aging workforce. Beginning with its theoretical roots in the aging at work literature, we built on qualitative evidence to conceptualize the measure. The constructs were operationalized by strictly following a rational and widely applied procedure. The LLWI scales comprise 80 items in a two-order hierarchical measurement model and thus allow for a differentiated assessment of relevant organizational practices. We provided extensive, repeated tests of the LLWI measure, its psychometric properties, the factor structure, and initial validity evidence.

Building the LLWI on extensive qualitative research (Wilckens, Wöhrmann, Adams, et al., 2020; Wöhrmann et al., 2018) laid the foundation for a comprehensive measure. Its comprehensiveness was further enhanced by the extensive item pool generated in Study 1 and revised with the support of multiple human resource and age management experts. The samples for Study 2 and 3 comprised individuals from the entire work population, covering a range of job titles and management levels. Respondents had highly diverse occupational backgrounds and were employed in small to very large organizations in various industries. Thus, the LLWI scales appear to be applicable across organizations with a variety of characteristics.

Being multifaceted, the LLWI is intended to contribute to the measurement of organizational practices for the aging workforce because answers to the question of how organizations can support an aging workforce remain limited without understanding the factorial structure. Thus, the LLWI measures nine distinct domains comprising several factors

each. However, we were not able to successfully distinguish all facets that had been found and defined in the qualitative LLWI studies. The two LLWI domains covering aspects of age-friendly *leadership* and aspects of *individual development* initially comprised multiple facets, but the operationalization was not able to distinguish these facets. Instead, we operationalized each of the two domains using a single unidimensional scale. For the remaining domains, the EFAs and CFAs we conducted supported the multifactor solutions, although we did retain a few items with cross-factor structure coefficients when these items made important content contributions to the scales. Studies 2 and 3 also provided evidence for construct validity of the overall hierarchical, multidimensional model covering all nine content domains with 22 subscales in total. This finding is particularly important because previous studies largely measured organizational practices with unidimensional scales.

Results showed that the LLWI is a reliable measure. Alpha coefficients and item-total correlations yielded acceptable values for all 22 subscales of the LLWI. At the same time, reliability was sufficiently low, indicating low levels of redundancy among items in the set. Study 2 further provided evidence for the LLWI's independence from positive and negative affect. Thus, the LLWI measure appears to be sufficiently distinguished from respondents' mood, a major source of common method variance. Various pieces of criterion validity evidence from Study 3 showed that the factors measured by the LLWI were positively correlated with a number of work outcomes present in the aging at work literature, such as work ability, person-job fit, work engagement, occupational future time perspective, and post-retirement work intentions. Furthermore, the LLWI scales were negatively correlated with stress due to perceived threat and pressure, turnover intentions, and illness-related days absent.

Although the nine domains identified were positively correlated among each other and thus share common variance, we did not form a single latent LLWI variable (such as *the organization's age-friendliness*). In line with Jarvis, MacKenzie, and Podsakoff (2003), an overall single latent variable for the nine domains had to be formed formatively because the

LLWI domains (e.g., age-friendly leadership, work design, health management) are not indicators for an underlying causing factor but are instead independent fields of managerial decisions that jointly induce age-friendliness. Depending on the outcome of interest, one or the other domain may be of higher importance, which could not be modelled by a single reflective latent variable. Each of the nine LLWI domains, however, is measured reflectively by a set of indicators—the measured practices and circumstances within the organization, which causally follow from the LLWI domain. As a consequence, we created the LLWI measure as a set of reflective measurement models, which can be integrated into an overall formative measurement model.

Implications

The operationalization of the LLWI created a multifaceted and comprehensive set of psychometrically sound measures to assess organizational practices for the aging workforce. The nine domains covered by the LLWI enable different organizational practices to be distinguished. The domains comply with a general managerial understanding of organizational levers, such as *health management* or *work design*, which eases application in practice. Practitioners and organizational researchers may find the sets of scales and subscales provided to be useful tools for deepening their understanding of processes and contextual factors of aging at work and for identifying organizational improvement potential to better facilitate aging at work. The LLWI enables organizations to assess their capabilities in managing, engaging, and retaining an aging and age-diverse workforce. Clearly, organizations differ in terms of room for change and resources to invest into practices. Small organizations with limited resources, for example, may be restricted to a certain number of practices. Manufacturing organizations with shift work may have less flexibility for work time and workplace arrangements than others. However, the LLWI can support decision makers in setting priorities, which is particularly important under scarce resources. Benchmarking is recommended, where the peer

organizations need to be carefully selected to match characteristics of the organization in question. Thereby, the LLWI assessment illuminates deficits, which may in turn trigger innovative and often low-cost solutions. Future benchmark studies based on the LLWI should address different organizational characteristics and identify domains of particular relevance for particular organizational settings.

Besides benchmarking among comparable organizations, the LLWI can also unfold within-organization differences in the practices. Different rating sources may perceive the availability of practices differently. This may be due to different subunits within the organization, but also due to different levels of knowledge regarding the practices. For example, managers and human resources representatives may be well aware of certain practices, whereas the organization's older workers are not. Assessing the LLWI from the perspective of human resource representatives is more likely to capture policies or practices as intended by the management, while the assessment from the perspective of older workers captures how practices reach the workers' level. Consequently, LLWI results can inform communication issues regarding the practices in the organization by leveraging different rating sources. Moreover, above average shares of "don't know" answers by older workers may indicate insufficient communication of practices on, if these are actually offered.

The multifactor, Likert-scaled LLWI measure appears to improve on existing measures in two ways. First, unlike unidimensional measures, the LLWI differentiates multiple organizational facets and provides better construct coverage. Second, unlike assessment by lists of practices, the LLWI improves the measurement by providing thoroughly developed items, construct validity, and internal consistency of each subscale.

For researchers in the field of aging at work, the LLWI provides opportunities to tap into specific organizational practices. For example, the concept of successful aging at work, conceptualized by Kooij et al. (2015, p. 309) as "the maintenance of workers' health, motivation, and working capacity or work ability now and in the future," upholds the

importance of organizational contextual factors for individual aging processes and coping strategies (e.g., Kooij, 2015; Zacher, 2015a). Those factors may facilitate, trigger, or obstruct individuals' constant adaption to age-related changes and may even reduce the need for resourceful coping actions (Rudolph, 2016). However, to further understand the influence of particular organizational practices on individual and organizational level outcomes, it is important to disentangle the practices of interest. As such, the LLWI creates the opportunity for sound empirical examination of the relationship between practices and successful aging in the future.

Applying the LLWI measuring all nine dimensions allows researchers to compare different domains of practices comprehensively. Results showed that all nine domains are correlated. However, it is important to capture relative differences between domains in order to understand how individual domains contribute to successful aging at work. For example, both research and practice can benefit from examining which of the nine domains are particularly relevant for older workers' health, commitment, or performance. The relevance of individual domains is also likely to depend on organizational conditions such as size or industry sector. To support practitioners in prioritizing the different domains of practices under unique organizational conditions, evidence-based findings on the relative relevance of different domains of practices are required as well. Thus, the entire set of nine domains with 80 items can offer highly comprehensive information for both research and practice purposes.

At the same time, the developed scales for each domain may also be used separately in future research. Each of the nine scales showed satisfactory psychometric properties. Several of the scales operationalize organizational practices that have not been studied extensively in the literature. For example, the *transition to retirement* scale allows researchers to tap into and disentangle practices for the retirement transition, whose characteristics, processes, and effects are still relatively blurry (Henkens et al., 2018). Other scales, such as the ones for individual development and knowledge management, assess general organizational practices in light of

workplace aging. In this respect, the scales are well-suited to further investigate how and to what extent specific domains of organizational practices influence aging processes and older workers' work outcomes. Consequently, the LLWI measure and its subscales promise to serve future research on aging at work research through thorough measurement.

Given that the 80-item measure is quite long, a shortened version of the LLWI would provide additional value. A short version with a compressed factor structure should be sufficient for capturing the overall construct of organizational practices for the aging workforce and providing research and practitioners with an overview of the status quo within the assessed organizations. As part of the initial diagnostic and need-analysis tool, norms and benchmarks could also be developed based on the short version of the measure that offer organizations comparative information regarding their peers.

Limitations and Future Directions

First, because cross-sectional data was used, presented studies are potentially limited by variance from a common-method and common-source bias. However, these biases do not affect identified low correlations between constructs, which result in identified factor structures and construct validity among the nine LLWI domains, because the biases generally elevate correlations within single-method or single-sourced samples (Campbell & Fiske, 1959). Hence, measured correlations tend to be too high than too low. To limit a potential common-source bias in the assessment of convergent and criterion validity, we clearly separated the LLWI items from the validation scales in the questionnaires (Podsakoff et al., 2003). Moreover, the limited correlations between affectivity and the LLWI suggest that the results are not affected by substantial common-method and common-source variance (Podsakoff et al., 2003).

Second, the quantitative evidence for the applicability of the LLWI is limited thus far to Germany, where we conducted the present studies. Nevertheless, the qualitative framework underlying the measure was initially developed on the basis of evidence from Germany and the

United States (Wilckens, Wöhrmann, Adams, et al., 2020). Thus, the framework is not tied to a single cultural or legal context. Moreover, we carefully sought to avoid legally or culturally specific items when developing items. International validations of the LLWI are planned for several countries and will provide opportunities for further research. This is particularly important because organizational practices for an aging workforce are subject to regulatory and cultural differences (Barnes et al., 2009). Additionally, the relevance of certain organizational practices may be subject to the organizational context. Study 2 results showed small but significant correlations between some of the LLWI domains and both the organizations' industry sector and the number of employees (see Table 3.3). The organizational context should thus be considered when using and interpreting the LLWI measure.

Third, building on the consensus referent model (Chan, 1998), the LLWI intends to abstract from respondents' individual experience to their perception of the organization in general. However, we did not obtain multiple respondents per organization. The present studies therefore do not provide any evidence for within-organization consistency of the measurement. Study 2 showed criterion validity regarding an age-diversity measure and an age-inclusive human resources practices measure, which had previously been validated with good intra-class correlations on the organizational level (Boehm et al., 2014). For this reason, we expect the LLWI to show within-organization consistency, although further research is required to assess the organizational nature of the LLWI in greater detail.

Fourth, our rigid scale development process resulted in nine distinct yet moderately to highly correlated domains of organizational practices. In particular, two sets of domains with high interrelations emerged. The second-order constructs of the retirement-related domains—transition to retirement and health and retirement coverage—showed correlations around .7, as did the domains of organizational climate, leadership, and individual development. In both cases, however, the distinctiveness of the constructs was supported not only by the exploratory and confirmatory factor analyses but also by the correlations between the first-order constructs

and items composing the domain scores. Generally, the correlations between the lower-order constructs across these domains did not exceed .7 (see Appendix H), which supports the validity of the individual constructs. Theoretically, the close association between these constructs and also their distinctiveness are well justified. For example, developmental practices have been shown to influence age-diversity climate (Boehm et al., 2014) and are dependent on the leaders who implement the developmental practices. Leadership style may thus also influence the developmental practices offered. Overall, despite the interdependencies between the LLWI domains, distinguishing domains of practices in assessment is important in order to generate focused effort within organizations.

Drawing upon the limitations of the LLWI measure just outlined, we conclude with the scale development notion that a measure is never complete but requires constant refinement (Clark & Watson, 1995). This article provides initial reliability and validity evidence for the newly developed LLWI measure, without proposing a final measure. Additional research is needed to establish the efficacy of the measurement on the organizational level. In addition, organizational level research on aging at work—which is still limited and which we aim to foster by providing the LLWI—will likely provide improvements to the scales in future.

4

Health and the Intention to Retire: Exploring the Moderating Effects of Organizational Practices

Max R. Wilckens, Anne M. Wöhrmann, Julia S. Finsel, and Jürgen Deller

Acknowledgement: We thank Verena Böhne, Luisa Cadonau, Shawnee Dierks, Tino Glumm, Kristina Spieker, Elisa Stahl, Katharina Ullmann, and all other students who supported us during data collection.

Abstract

Although employees' health is among the strongest predictors of retirement timing organizational effects on their relationship are largely unknown. Based on the theory of work adjustment and socioemotional selectivity theory, this study explores the role of organizational practices in the relation between older employees' health and retirement intentions—specifically, their preferred retirement age and intention to engage in a late career after being eligible for pension. Three groups of practices are distinguished: individual development practices (e.g., life-long learning and career development), practices tailoring the transition to retirement (e.g., phased retirement), and practices providing opportunities to continue working in later life (e.g., individualized employment forms). We tested our model with multi-level data from 556 older employees and 661 managers from 101 organizations in Germany. Results suggest that healthy older employees intend to retire later, if individual development and continued employment practices are present, while employees with poor health intend to retire earlier, if transition-to-retirement practices are present. The positive relation between health and the intention to engage in a late career was stronger in organizations with opportunities to continue working than in organizations engaging in individual development and transition-to-retirement practices. Our findings contribute to a better understanding of retirement intentions and offer practical implications for organizations to shape later-life work to the benefit of both organizations and employees.

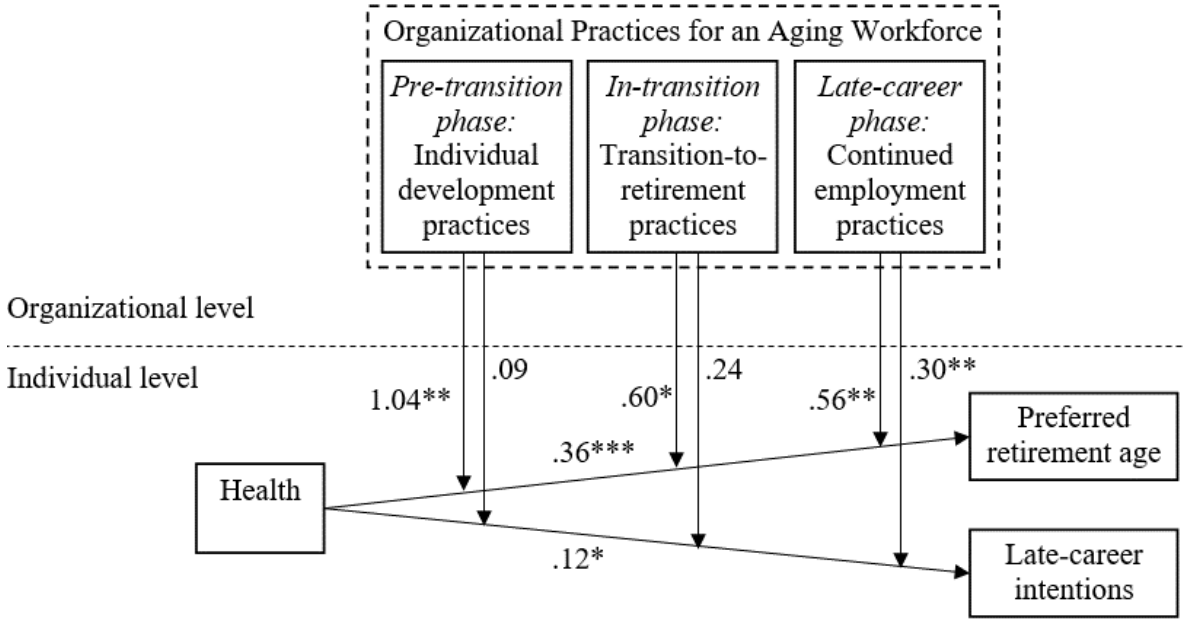
Keywords: older workers' health, person-environment fit, retirement intentions

In the wake of increased life expectancy and low birth rates, the proportion of working people in the population as a whole is shrinking in most developed countries (OECD, 2017). To sustain productivity and a sufficiently large labor force, society and organizations need to motivate older employees to continue working instead of retiring early. Employees' health and contextual factors of the work environment are among the most important antecedents of prolonged careers (Fisher et al., 2016; Henkens et al., 2018; S. Kim & Feldman, 2000; Topa et al., 2009; van den Berg et al., 2010; Wang et al., 2008). Healthy employees generally intend to retire later than employees in poor health (e.g., McGarry, 2004). Moreover, organizational practices such as flexible work arrangements and developmental practices have been found to increase older employees' intention to continue working (e.g., Armstrong-Stassen, 2008; Armstrong-Stassen & Schlosser, 2008; Bal et al., 2012; Polat et al., 2017). Because individual and environmental factors jointly affect employment relations, the intention to continue working most likely also results from an interaction between organizational practices and employees' health. Yet studies on the moderating effects of organizational practices on the relation between health and retirement intentions are limited. To further develop organizational practices for older employees, a finer-grained understanding of how organizational practices interact with older employees' health is required.

Drawing on the theory of work adjustment (Dawis et al., 1964; Dawis & Lofquist, 1984) and socioemotional selectivity theory (Carstensen, 1987, 1991), we argue that organizational practices can strengthen the effect of employees' health on their intention to continue working. We expect this effect for three groups of organizational practices that support an individualized retirement timing, namely for individual development practices, including life-long learning and career development for older employees; for practices that tailor the transition to retirement to older employees' individual needs (e.g., phased retirement); and for practices that allow older employees to continue working by offering individualized forms of employment and

Figure 4.1

Schematic depiction of results from random intercept, random slopes models.
*Organizational level predictors were grand-mean centered. Individual level predictors were centered at the organization's mean. Hypotheses were tested in separate models. * indicates p < .05; ** indicates p < .01; *** indicates p < .001.*



appropriate jobs. We hypothesize that these organizational practices moderate the positive relationships between older employees' health and (a) their intended retirement age and (b) their intention to engage in a late career (in their current or a different job) for their current employer post pension eligibility (see Figure 4.1). Thus, these practices can support earlier or later retirement, depending on employees' individual health resources and respective needs.

This article contributes to the body of existing knowledge in several ways. First, we shed light on the relationship between employees' health and their retirement intentions. With this understanding, it is possible to create opportunities to extend working lives. Second, by investigating the interactions between organizational practices and health as an individual resource we respond to calls for a more integrated person-environment fit perspective on the adjustment process towards retirement (Fisher et al., 2016; Wang & Shultz, 2010). To account for the individual and organizational levels methodologically, this study applies a multi-level, multi-organization, and multisource design, which has rarely been used in retirement research

thus far. Third, following a life-span perspective, researchers have argued that organizational practices beneficial to older employees differ from general human resource practices (e.g., Kooij et al., 2014). For that reason, this article assesses organizational practices specifically tailored to older employees. Our results provide additional empirical evidence for the structure and nature of organizational practices that are relevant in the context of aging at work. Finally, from a practice perspective, our findings on the role organizational practices can play in the retirement intentions of employees with different levels of health may support a more individualized retirement transition—motivating and enabling employees to continue working, while simultaneously allowing for an early exit from work if needed.

Theoretical Framework

Health and Retirement

Health has been widely found to be one of the most important predictors of employees' retirement age and retirement intentions (Fisher et al., 2016; Sullivan & Al Ariss, 2019; van den Berg et al., 2010; Wang & Shultz, 2010). Using data from the Health and Retirement Study, McGarry (2004) in fact revealed that changes in perceived health affect the expected retirement age much more strongly than do changes in income and wealth. This effect can be explained by the theory of work adjustment (Dawis et al., 1964; Dawis & Lofquist, 1984), which conceptualizes intentions to discontinue working based on a lack of fit between the individual and the work environment. Good fit and, respectively, a sustainable employment relation are achieved if an individual employee's needs are fulfilled by the work environment and if demands at work are satisfied by the employee's abilities (Edwards & Shipp, 2007).

An adequate level of health is a prerequisite for workers to satisfy their organization's demands at work. Not only are employees' abilities affected by poor health but employees in poor health have additional needs (e.g., an ergonomic workplace) that cannot always be supplied by the organization. Consequently, poor health impedes achievement of sufficient

person-environment fit at work. To nevertheless maintain such a fit, organizations need to adjust employees' work accordingly and cater to employee's individual needs. Yet this is often difficult and costly because of the limited number of jobs that are both low in demands and economically meaningful while also matching the employee's capabilities and needs (see also Dingemans et al., 2016; Johnson et al., 2009).

According to the theory of work adjustment, a lack of person-environment fit at work results in mutual dissatisfaction. Ultimately, this dissatisfaction leads to higher employee turnover: The employee either quits the job or is forced to leave it by the organization. In fact, poor person-environment fit at work has been found to induce turnover intentions because a different job might better fit the employee's abilities and needs (Edwards, 1991; Verquer et al., 2003). For older employees nearing retirement age, a decrease in person-environment fit at work motivates them to end their career and retire (Beier et al., 2020; Feldman, 2012). Wang and Shultz (2010) argued that "we can conceptualize the decision to retire as a result of misfit between employees' personal characteristics [such as, health,] and the characteristics of their job, group, organization, and supervisors" (p. 195). A recent study by Lahlouh et al. (2019) supported that the fit between work demands and employees' abilities (demands-ability fit) was negatively related to employees' full retirement intentions and that both demands-ability fit and the fit between employees' needs and supplies at work (i.e., organizational resources and opportunities for the employee; needs-supplies fit; see Kristof, 1996) were positively related to older employees' bridge employment intentions. Moreover, Oakman and Wells (2016) found that perceptions of demands-ability fit, operationalized as work ability (V. Ilmarinen et al., 2015), were related to older workers' retirement intentions.

Recent theoretical advancements by Hesketh et al. (2011, 2015) extended the fit framework of the theory of work adjustment to the retirement phase. The authors argued that individuals proactively and reactively behave in a way that improves their fit with the environment. Thus, misfit with the retirement environment may also result in intentions to

engage in a late career (Hesketh et al., 2011; Wang & Shultz, 2010). For instance, healthy individuals in retirement age might achieve higher fit with the work environment, for example, if they do not find the demands of the retirement environment challenging enough. Consequently, good health can result not only in intentions to retire later but also in a late career after the normal retirement age.

A life-span developmental perspective provides a second argument for the relationship between health and retirement intentions that is tailored more strongly to the situation of older employees. Socioemotional selectivity theory (Carstensen, 1987, 1991, 1992) suggests that individuals select their goals in alignment with their future time perspective. When the perceived remaining time in life becomes scarce, individuals strive less for new endeavors and exploration. Instead, they focus on emotional goals, savoring their remaining time in life. Consequently, they deprioritize work-related goals such as prolonging one's career and may thereby form the intention to retire early. Employees' lifetime horizon has been shown to be positively related to their intention to continue working (Akkermans et al., 2016; Kooij et al., 2018) and to their intended retirement age, even after controlling for income, education, health, and marital status (Henry et al., 2017; van Solinge & Henkens, 2010). How individuals perceive their remaining lifetime depends on their health. Kooij and van de Voorde (2011) and Bal et al. (2010) found that perceived health is positively related to an open-ended perception of one's remaining lifetime. Similarly, Fasbender et al. (2019) found negative effects of physical losses on workers' occupational future time perspective. Thus, employees' perceived health influences their perceived remaining time and, in turn, their retirement intentions.

To study employees' retirement intentions we focus on two interrelated motivational constructs: employees' intended retirement age and their intention to engage in a late career for their current employer after being eligible for pension (Lahlouh et al., 2019; Topa & Alcover, 2015; Wang & Shultz, 2010). Early pioneers in research on retirement intentions and decisions focused on the antecedents and consequences of early retirement (Barnes-Farrell, 2003), while

more recent studies also investigated the intention to continue working and to engage in a late career (Armstrong-Stassen & Schlosser, 2008; Bal et al., 2012; Gobeski & Beehr, 2009; Salminen et al., 2019; Wöhrmann et al., 2013). The intention to stop working before the normal retirement age differs from the intention to pursue a late-career because, in most countries, a normal retirement age is deeply rooted in the regulatory framework of, for example, the social security system (Calvo et al., 2013; Radl, 2012). Wang and Shultz (2010) noted that the decision to retire early is dominated by health and financial factors, while the decision to continue working is often motivated by adjustment processes to retirement life. Nevertheless, both constructs have been found to depend on employees‘ health resources (e.g., Bal et al., 2012; Barnes-Farrell, 2003). Drawing on the theoretic approaches just outlined and in keeping with the previous research, we hypothesize:

Hypothesis 1: Older workers' health is positively associated with (a) their intended retirement age and (b) their intention to engage in a late career after being eligible for pension.

Organizational Practices and the Effect of Health on Retirement Intentions

Taking a person-environment fit perspective, organizational practices can help sustain a sufficient fit at work by shaping the work environment. Organizations can—to some extent—react to a decrease in employees' abilities by adjusting demands (e.g., through work design practices) and by catering to employees' individual needs (e.g., Kooij et al., 2014). In turn, sustained fit increases employees' intention to continue working. Previous research has found that specific practices, such as flexible work arrangements—which allow employees to schedule their work flexibly and thereby better reconcile their work with their personal needs—are positively related to the intention to retire later (Armstrong-Stassen, 2008; Bal et al., 2012).

Nevertheless, some employees in exceptionally poor health would potentially achieve the best possible fit of their abilities and the demands posed on them by retiring early. To achieve the best person-environment fit—regardless of whether at work or in retirement—

organizations may therefore offer, support, and promote individualized retirement transitions that allow employees to retire early or later, depending on their own abilities and needs (not to be confused with the usually inadvisable practice of demotion, cf. van Dalen & Henkens, 2018). Life-span developmental theorists have argued that aging processes do not follow discrete time-boxed stages but instead individual trajectories subject to personal and contextual factors (e.g., Rudolph, 2016). Individuals of the same age differ in terms of health constraints affecting their abilities and their needs within their work environment. Empirical evidence supports that this heterogeneity increases with age (e.g., J. Ilmarinen, 2001). Thus, researchers have argued that organizational practices that flexibly account for older employees' diverging needs and ability profiles are most beneficial for their person-environment fit (Kooij et al., 2014; Wilckens, Wöhrmann, Deller, et al., 2020). Such practices support each employee individually in their career development, retirement planning, and retirement transition as well as in extending their career beyond the normal retirement age. They also enable and motivate employees with sufficient ability—such as those in good health—to delay retirement, while supporting those employees who would benefit from earlier retirement in terms of person-environment fit (i.e., those in poor health) to strive for an earlier exit from work. Consequently, we expect organizational practices that foster individualized retirement transitions to strengthen the relationship between workers' health and their retirement intentions.

We conceptualize organizational practices for an individualized transition to retirement in three categories: 1) practices providing older employees with individual opportunities for development at work, 2) practices preparing and tailoring the transition to retirement, and 3) practices facilitating continued employment even beyond the normal retirement age. We selected these practices as they create opportunities for an earlier or later retirement and allow employees to adjust their late career and/or retirement plans to their individual health resources. Since these practices tailor work environments specifically for the employees' current

employer, we also focus on employees' intentions to engage in a late career after normal retirement age at their particular employer.

Individual development practices. This group of organizational practices supports employees in their professional and personal development. It includes opportunities for lifelong learning and for a continuous career development, where the work is adjusted to the employee's capabilities and developmental interests. In the process, employees participate in planning their careers, which may include not only promotions but also horizontal job changes, demotions, further training, and modifications to their tasks and responsibilities, among others (Wilckens, Wöhrmann, Deller, et al., 2020). Previous research by Oostrom et al. (2016) found that such practices allowing employees to influence their tasks and responsibilities at work affect their work ability in terms of increasing their future time perspective. Employees are provided with a feeling of control over their work and with opportunities to develop their capabilities in adjustment to new work-related challenges. As a result, employees gain confidence in their ability to sustain the correspondence between their work environment and their own needs and abilities, even if the latter may change in the future. This effect of individual development practices is not specific to a particular age group. However, to motivate older employees to prolong their careers, it is important that organizations also engage in developmental practices targeting older employees. Bal et al. (2012) found that possibilities of individual development for older employees were positively related to their motivation to continue working, if the organizational climate was characterized by openness to personal development. The authors controlled for health, which was negatively associated with the motivation to continue working. Drawing on work adjustment theory, however, we argue that person-environment fit requires both sufficient health as a personal resource and individual development practices to provide a perspective for future time at work, be it within the current job or in a late career after retirement. Consequently, we posit:

Hypothesis 2: Individual development practices in the organization strengthen the positive relation between older employees' health and (a) their intended retirement age and (b) their intention to engage in a late-career for their current employer.

Transition-to-retirement practices. This group of organizational practices includes a timely planning of the transition to retirement and individualized solutions for the process, such as phased retirement and counselling, to prepare retirees for the life-changing event of their retirement. These practices directly affect employees *in transition* to retirement. According to the theory of work adjustment, opportunities for a tailored transition to retirement can enable employees in poor health to retire earlier if there is a respectively poor fit between their work environment and their abilities and needs. At the same time, such practices also allow employees in good health and with a correspondingly higher congruence between their abilities and the job demands to prolong their careers and retire later. Counselling practices that support older workers in the transition to retirement have been argued to support employees' engagement in late careers for the same or a different employer (Harper & Shoffner, 2004; Wöhrmann et al., 2014; Zacher et al., 2019), but also their adjustment to post-retirement life (e.g., Osborne, 2012). Such practices prior to retirement can also reduce employees' stress and anxiety (Osborne, 2012) that might be attributable to the imminent changes in life after retirement, such as the loss of social contacts with colleagues at work, a significant increase in leisure time, and a lack of structure in everyday life. Thus, such practices may reduce barriers to retire, which is relevant for employees with a poor person-environment fit at work who would benefit from early retirement. Previous research revealed that employees who made use of phased retirement programs intended to retire earlier than those working full-time (van Solinge & Henkens, 2014). However, results from the Health and Retirement Study provided some evidence that such programs actually prolonged employees' careers, because the use of phased retirement practices increased with employee age after controlling for health (Cahill et al., 2015). The set of practices offering flexible retirement timing, phased retirement options, and

counselling for retirement likely supports both the intention to retire early in the event of misfit between the employee and the work environment and the intention to retire later in the event of good fit. We therefore do not expect a direct effect of more tailored transition practices on employees' retirement intentions, but instead an interaction effect with the employees' personal (health) resources. Consequently, we posit:

Hypothesis 3: Transition-to-retirement practices in the organization strengthen the positive relation between older employees' health and (a) their intended retirement age and (b) their intention to engage in a late-career for their current employer.

Continued employment practices. Organizations that implement practices for continued employment and hiring older workers (*late-career phase)* create a work environment in which older employees can work even beyond normal retirement age. This environment includes individualized forms of employment that provide additional resources (e.g., flexibility in how employees conduct their work and social support) and a modified demand profile (e.g., reduced number of weekly work hours or a modified job profile) tailored to the employees' abilities and needs (Rudolph et al., 2015; Wilckens, Wöhrmann, Deller, et al., 2020; Wöhrmann et al., 2018). According to the theory of work adjustment, such a work environment tailored to the abilities and needs of older employees supports late careers by promoting good fit of the individual with the work environment. We expect such organizational practices to be particularly effective in increasing healthy employees' intention to engage in late careers because a good demands-abilities fit and needs-supplies fit also require employees to be sufficiently healthy. Moreover, from a life-span perspective, organizational practices targeting employment beyond the normal retirement age and explicitly hiring older employees may also provide employees with a perspective for future time at work (Zacher & Frese, 2009). Employees within the company may take newly employed older employees as a role model and discover new perspectives to prolong their working lives. This in turn could positively affect

not only employees' intentions for a late career but also their general intentions to delay their retirement in case of sufficient individual resources. Consequently, we posit:

Hypothesis 4: Continued employment practices in the organization strengthen the positive relation between older employees' health and (a) their intended retirement age and (b) their intention to engage in a late-career for their current employer.

Method

We tested our hypotheses upon a unique multilevel dataset from a diverse sample of 101 organizations in Germany. Per organization at least two older employees aged 50 and above, and three managers participated. By including multiple respondents per organization, we were able to obtain the constructs relevant in this study from different sources, limiting the effect of potential common method bias on the results (Podsakoff et al., 2003). Health and retirement intentions were self-rated by the employees. Organizational practices were assessed by the managers and aggregated on the organizational level. To ensure validity of our results, we followed the widely acknowledged principles for multi-level research by Kozlowski and Klein (2000) throughout the study.

Participants

To test our hypotheses we needed a highly diverse sample of organizations in order to achieve sufficient variance in the HR practices of interest. Yet, representativeness was less a concern, because we did not aim at researching prevalence of the phenomena. We followed three complementary approaches to recruit organizations for the study. First, we contacted industry associations that forwarded an invitation to take part in the study to their member organizations. Second, we approached an association of human resources representatives to invite their organizations to take part in the study. Third, we directly invited organizations via our personal networks. In line with previous research, participating organizations were required

to employ at least 30 employees in order to ensure existence of organizational HR practices in general (Shaw et al., 1998).

Participating organizations were asked to invite at least three managers (one human resources manager and several line managers) and at least two older workers to respond to the questionnaire. Some larger companies took part with multiple sites of very different kind—for example, a production facility and the company's headquarter. Thus, for the analyses in this study we treated each site as a different organization. In exchange we offered participating organizations a detailed analysis of the results with specific recommendations for their organizations on how to improve organizational practices for an aging workforce. We received responses from at least three managers and two older employees from 106 organizations (6% of the invited organizations), yielding responses from 641 older employees and 679 managers in total. Responses from 85 older employees (13%) were removed from further analysis due to missing values for the two dependent variables or one of the independent variables. Most of the excluded participants (62 older employees) did not respond to the demographic control variables. The remaining responses were obtained from 101 organizations.

Organizational characteristics. The 101 organizations were widely spread across sectors and organizational size, thereby providing a highly diverse sample of work environments. The organizations were active in the industrial sector (36%), the services sector (36%), or the public sector (29%). The median organization had 4,300 employees, of whom about 20% were aged 55 and older. One fourth of the organizations (25%) were family businesses. Consequently, the sample of organizations well covered the spectrum of organizations with sufficient size to actively implement organizational practices (at least 30 employees).

Managers' characteristics. We received 661 usable responses from managers, who assessed the organizational practices for their organizations. Respondents were aged 49 on average and had a supervisory position. The sample was largely male (65%) and the majority

of managers (63%) had worked for their organization for more than 15 years. In line with their supervisory tasks, 94% of the managers had a white-collar job. Most managers (72%) had at least a college degree.

Older employees' characteristics. We received 556 usable responses from older employees. This amounts to 5.5 participating older employees per organization on average. Respondents were aged 56 years on average. The sample was largely male (55%) and the majority of participants (74%) had worked for their organization for more than 15 years. Most of the participants (77%) worked in white-collar jobs, which is in line with the overall share of white-collar jobs of older employees in Germany (Statistisches Bundesamt, 2019). The older employees were less educated than the managers surveyed. A minority of 35% had a college degree and a further 24% of participants had multiyear vocational training, which is a common educational track following secondary school in Germany.

Measures

Preferred retirement age. The preferred retirement age was measured with the question "If it was your free choice, at what age would you like to retire?" Assessing respondents' preferred retirement age in years and with a single question is common practice in research on retirement decision-making (Vignoli et al., 2019; Wöhrmann et al., 2017; Zaniboni, 2015). Here, participants preferred to retire at an age between 50 and 70 years (M = 56.05, SD = 3.71). However, responses were not evenly distributed: 32% of respondents preferred to retire at the age of 60 years and 24% at the age of 63 years.

Late-career intentions. Employees' intention to continue working beyond the normal retirement age was measured with three items from Wöhrmann et al. (2013). This measure assesses respondents' willingness to continue working for their current employer on a 7-point Likert scale. The items were "I would like to work for my current employer after retirement," "I would like to continue paid work after retirement," and "If the organization asked me to

return to work for it after my retirement, I would." Reliability analysis revealed that the three items were sufficiently congruent to compute the scale mean (*Cronbach's alpha = .86*). We applied predictive mean matching (Kleinke, 2017) to impute missing data before computing the scale mean for those participants, who answered at least half of the items of the scale. For the remaining participants a scale mean was not computed. We followed this procedure also for the following multi-item scales.

Health. Respondents' health was assessed with four items from Adams and Beehr (1998; sample item: "My health is better than most people my age"). Respondents answered on a 7-point Likert scale. The scale showed very good reliability (*Cronbach's alpha = .89*), so that we computed employees' health as the scale mean.

Organizational practices. To assess the three facets of organizational practices—individual development practices, transition-to-retirement practices, and continued employment practices—we used scales from the Later Life Workplace Index (Wilckens, Wöhrmann, Deller, et al., 2020). The scale measuring individual development practices comprised eight items with good reliability (*Cronbach's alpha* = .82). Sample items were "In our organization, employees, regardless of age, know about their potential for development" and "In our organization, employees move to a different job or position if it better suits their specific skills and abilities." The scale measuring transition-to-retirement practices comprised 14 items with good reliability (*Cronbach's alpha* = .84). Sample items were "In our organization, managers take time to plan the transition to retirement for individual employees" and "In our organization, the transition to retirement is flexibly shaped according to employee needs." Finally, the scale assessing the possibility to continue working comprised seven items with good reliability (*Cronbach's alpha* = .80). Sample items were "In our organization, employees may work beyond the conventional retirement age if they wish so" and "In our organization, older applicants are hired as well." Respondents answered on a 5-point Likert

scale ranging from "does not apply at all in our organization" to "fully applies in our organization."

We obtained the assessments of the organizational practices from the managers of each participating organization to gain as objective an evaluation as possible (Khilji & Wang, 2006). Results from a confirmatory factor analysis (CFA) confirmed the quality of the measurement model and showed that the measured practices were sufficiently distinct (*RMSEA = .06; CFI = .89*). Despite limited shared variance between managers of the same organization fair to good interrater reliability (Bliese, 1998) confirmed that the assessments within each organization were sufficiently congruent (individual development practices: *ICC1 = .14, ICC2 = .48,* $r_{wg(j)}$ *= .89*; transition-to-retirement practices: *ICC1 = .22, ICC2 = .61,* $r_{wg(j)}$ *= .83*; continued employment practices: *ICC1 = .36, ICC2 = .75,* $r_{wg(j)}$ *= .85*). Thus, scale means from the managers per organization were aggregated to assess the organizational-level constructs (Kozlowski & Klein, 2000).

Controls. In our analysis, we included several control variables to account for confounding factors of employees' retirement intention that had been identified in previous research. First, employees' preferred retirement age had been found to be positively related to their chronological age (e.g., Zaniboni, 2015). The older employees are and the nearer their retirement respectively is, the longer employees are willing to work. Thus, we expected a positive relation between employees' chronological age and their preferred retirement age. Second, gender had shown mixed effects on employees' retirement intentions and timing in previous research (Fisher et al., 2016). We therefore included gender as a control variable to account for potential biases in the sample. Third, employees with higher levels of education were related to later retirement and to late-career intentions (e.g., Mermin et al., 2007). Fisher et al. (2016) argued that the effect may be partly explained by the fact that higher levels of education are also related to more attractive jobs with higher income and better working conditions. Thus, we included participants' education as a control variable. Education was

measured with a single question regarding the participants' highest degree. Fourth, we included two organizational-level control variables to mitigate potential confounding of our results by structural between-organization differences in employees' retirement intentions: first, the number of employees per organization, and second, whether the organization is a public-sector organization.

Results

Descriptive Results

Descriptive results and correlations of all variables used in the present study can be obtained from Table 4.1. Employees' health was significantly and positively correlated with their preferred retirement age (r = .18, p < .001) and their late-career intentions (r = .20, p < .001). Health was also significantly and positively correlated with individual development practices (r = .14, p < .01) within the employees' organization. The three groups of organizational practices (individual development practices, transition-to-retirement practices, and continued employment practices) were significantly and positively correlated among each other (r = [.24; .39], p < .001). The correlation among practices of the Later Life Workplace Index is in line with previous research by Wilckens et al. (2020). The authors argue that low to medium level correlations result from the fact that employers usually do not implement a single practice but bundles of practices to support the aging workforce. Continued employment practices were significantly and positively related to both employees' preferred retirement age (r =.12, p < .001) and their late-career intentions (r = .18, p < .001). Moreover, this group of practices was significantly more pronounced in small organizations (r = .49, p < .001) and less evident in public-sector organizations (r = -.20, p < .001). The other two groups of organizational practices did not show significant relations to any of the study's outcome variables. Furthermore, employees' preferred retirement age was significantly related to employees' age (r = .29, p < .001), their education (r = .13, p < .001), and the number of

Table 4.1

Means, standard deviations, and correlations of study variables

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
1. Chronological age	56.02	3.72	-										
2. Gender (female = 1; male = 2)	1.55	0.50	-.04	-									
3. Education (no degree = 1; college degree = 7)	5.07	1.69	.04	.08.	-								
4. Organization's number of employees	7.15	1.29	.02	.08.	.03	-							
5. Industry dummy: Public sector	0.44	0.50	.13**	-.13**	-.12**	-.03	-						
6. Health	5.44	1.17	.01	-.03	.11**	-.04	-.14**	(.89)					
7. Individual development practices[a]	3.55	0.33	.03	.01	.03	.06	-.08*	.14**	(.82)				
8. Transition-to-retirement practices[a]	2.53	0.42	-.01	.04	.04	-.09*	-.08	.06	.39***	(.84)			
9. Continued employment practices[a]	3.01	0.48	-.01	-.06	.00	-.49***	-.20***	.08.	.27***	.24***	(.80)		
10. Preferred retirement age	61.72	2.93	.29***	-.01	.13**	-.12**	.08.	.18***	.04	-.06	.12**	-	
11. Late-career intentions	2.23	1.09	.06	.03	.11*	-.21***	-.16***	.20***	.05	.03	.18***	.34***	(.86)

Note. $N = 556$. *M* and *SD* are used to represent mean and standard deviation, respectively. Internal consistency coefficients, Cronbach's alphas are reported in the parentheses on the diagonal. [a] indicates variables assessed by the managers and averaged on the organizational level. * indicates $p < .05$. ** indicates $p < .01$. *** indicates $p < .001$.

employees within their organization (r = -.12, p < .001). Employees' intention to engage in a late career was significantly lower in large organizations (r = -.21, p < .001) and in public sector organizations (r = -.16, p < .001).

Hypothesis Testing

To test our hypotheses, we computed two sets of linear mixed-effect models using the lme4 package in R (Bates et al., 2015; http://www.R-project.org/). The first set of models was used to test our hypotheses regarding employees' preferred retirement age. The second set of models was used to test our hypotheses regarding employees' late-career intentions. Results can be obtained from Table 4.2 and Table 4.3, respectively. We used mixed-effect models with a random-slope random-intercept design because they allowed us to simultaneously compute both the hypothesized direct effects and the cross-level interactions.

The control variables showed different effects for the two outcome variables. Participants' preferred retirement age had a significant and positive relation with their chronological age (γ = .19, p < .001), a significant and positive relation with their educational level (γ = .21, p < .01), and a significant yet negative relation with the organizations' number of employees (γ = -.31, p < .001). Participants' late-career intentions were neither related to chronological age (γ = .02, n.s.), nor to education (γ =.05, n.s.). However, they were significantly and negatively related to the organizations' number of employees (γ = -.20, p < .001) and to whether the organization belongs to the public sector (γ = -.28, p < .05). Thus, employees' late-career intentions were significantly lower in large and public-sector organizations.

Hypothesis 1 stated that employees' health positively relates to (a) their preferred retirement age and (b) their late-career intentions. Results confirmed this relationship for both outcome variables (γ = .36, p < .001 and γ = .12, p < .01).

Table 4.2

Random intercept random slopes model. Dependent variable: preferred retirement age

Predictor	Empty model			g	Controls only			g	Model without controls			g	Model with controls			g	ID interaction only			g	TR interaction only			g	CE interaction only			
	γ	*T*	*SE*		*γ*	*T*	*SE*		*γ*	*T*	*SE*		*γ*	*T*	*SE*		*γ*	*T*	*SE*		*γ*	*T*	*SE*		*γ*	*T*	*SE*	
	CI				*CI*				*CI*				*CI*				*CI*				*CI*				*CI*			
Fixed effects																												
(Intercept)	61.75***	378.44	0.00		61.42***	239.16	0.00		61.73***	386.77	0.00		61.32***	237.84	0.00		61.33***	237.71	0.00		61.34***	237.74	0.00		61.32***	237.51	0.00	
	[61.43; 62.07]				[60.92; 61.92]				[61.41; 62.03]				[60.83; 61.83]				[60.83; 61.83]				[60.84; 61.84]				[60.82; 61.82]			
Chronological age					0.19***	5.10	0.00						0.19***	5.19	0.00		0.19***	5.18	0.00		0.18***	5.03	0.00		0.19***	5.16	0.00	
					[0.12; 0.26]								[0.12; 0.26]				[0.12; 0.26]				[0.11; 0.25]				[0.12; 0.26]			
Gender (male)					0.21	0.86	0.39						0.27	1.09	0.28		0.27	1.09	0.27		0.27	1.10	0.27		0.28	1.13	0.26	
					[-0.29; 0.70]								[-0.23; 0.74]				[-0.23; 0.74]				[-0.23; 0.74]				[-0.22; 0.75]			
Education					0.21*	2.48	0.01						0.17*	2.08	0.04		0.18*	2.18	0.03		0.19*	2.23	0.03		0.17*	2.05	0.04	
					[0.04; 0.37]								[0.01; 0.33]				[0.02; 0.34]				[0.02; 0.35]				[0.01; 0.33]			
Organization's number of employees					-0.31**	-2.63	0.01						-0.19	-1.42	0.16		-0.19	-1.40	0.17		-0.19	-1.38	0.17		-0.20	-1.46	0.15	
					[-0.55; -0.08]								[-0.46; 0.07]				[-0.45; 0.08]				[-0.45; 0.08]				[-0.46; 0.07]			
Industry dummy: Public sector					0.51	1.58	0.12						0.62.	1.87	0.07		0.61.	1.85	0.07		0.59.	1.79	0.08		0.62.	1.88	0.06	
					[-0.12; 1.14]								[-0.02; 1.27]				[-0.03; 1.26]				[-0.05; 1.24]				[-0.02; 1.27]			
Health									0.39**	3.32	0.00		0.36**	3.13	0.00		0.34**	2.90	0.01		0.32**	2.72	0.01		0.34**	2.94	0.01	
									[0.15; 0.62]				[0.13; 0.59]				[0.10; 0.57]				[0.08; 0.56]				[0.10; 0.57]			
Individual development practices									0.37	0.74	0.46		0.47	0.95	0.34		0.49	0.99	0.32		0.37	0.75	0.45		0.39	0.80	0.43	
									[-0.59; 1.33]				[-0.47; 1.41]				[-0.46; 1.43]				[-0.58; 1.33]				[-0.56; 1.36]			
Transition-to-retirement practices									-0.53	-1.30	0.20		-0.49	-1.23	0.22		-0.50	-1.26	0.21		-0.44	-1.11	0.27		-0.51	-1.28	0.21	
									[-1.31; 0.27]				[-1.26; 0.27]				[-1.27; 0.26]				[-1.22; 0.33]				[-1.28; 0.26]			
Continued employment practices									0.73*	2.27	0.03		0.63.	1.70	0.09		0.60	1.62	0.11		0.59	1.60	0.11		0.65.	1.75	0.08	
									[0.10; 1.35]				[-0.08; 1.35]				[-0.11; 1.31]				[-0.12; 1.31]				[-0.07; 1.37]			
Health x Individual development practices									0.80*	2.00	0.05		0.86*	2.23	0.03		1.04**	3.03	0.00									
									[0.02; 1.58]				[0.11; 1.61]				[0.37; 1.71]											
Health x Transition-to-retirement practices									0.20	0.58	0.56		0.10	0.30	0.77						0.60*	1.99	0.05					
									[-0.48; 0.89]				[-0.57; 0.77]								[0.01; 1.21]							
Health x Continued employment practices									0.40	1.52	0.13		0.42	1.65	0.10										0.56*	2.28	0.02	
									[-0.11; 0.91]				[-0.08; 0.92]												[0.08; 1.05]			
Random effects																												
σ^2	7.69				7.17				7.30				6.83				6.82				6.87				6.90			
τ_{00}	0.91 $_{Organization}$				0.88 $_{Organization}$				0.88 $_{Organization}$				0.89 $_{Organization}$				0.90 $_{Organization}$				0.90 $_{Organization}$				0.88 $_{Organization}$			
τ_{11}									0.07 $_{Organization.Health}$				0.05 $_{Organization.Health}$				0.07 $_{Organization.Health}$				0.11 $_{Organization.Health}$				0.06 $_{Organization.Health}$			
ρ_{01}									0.83 $_{Organization}$				0.55 $_{Organization}$				0.50 $_{Organization}$				0.62 $_{Organization}$				0.51 $_{Organization}$			
ICC	0.11				0.11				0.12				0.12				0.12				0.12				0.12			
N	100 $_{Organization}$				100 $_{Organization}$				100 $_{Organization}$				100 $_{Organization}$				100 $_{Organization}$				100 $_{Organization}$				100 $_{Organization}$			
Observations	545				545				545				545				545				545				545			
Marginal R^2 / Conditional R^2	0.000 / 0.106				0.077 / 0.178				0.053 / 0.162				0.117 / 0.224				0.110 / 0.221				0.100 / 0.216				0.107 / 0.214			

Note. Organizational level predictors were grand-mean centered; Individual level predictors were centered at the organization's mean; Unstandardized coefficients (*γ*) with standard errors (*SE*) are shown; σ^2 = within-person variance; τ_{00} = between-person intercept variance; τ_{01} = between-person slope variance; ρ_{01} = correlation between intercept and slope; . indicates p < .1; * indicates p < .05; ** indicates p < .01; *** indicates p < .001.

Table 4.3

Random intercept random slopes model. Dependent variable: late-career intentions

Predictor	Empty model	g	Controls only	g	Model without controls	g	Model with controls	g	ID interaction only	g	TR interaction only	g	CE interaction only
	γ *T* *SE* *CI*		*γ* *T* *SE* *CI*		*γ* *T* *SE* *CI*		*γ* *T* *SE* *CI*		*γ* *T* *SE* *CI*		*γ* *T* *SE* *CI*		*γ* *T* *SE* *CI*
Fixed effects													
(Intercept)	2.24*** 35.02 0.00 [2.12; 2.37]		2.31*** 24.74 0.00 [2.13; 2.50]		2.23*** 36.24 0.00 [2.11; 2.35]		2.28*** 23.97 0.00 [2.09; 2.46]		2.28*** 23.91 0.00 [2.09; 2.47]		2.28*** 23.97 0.00 [2.09; 2.47]		2.27*** 23.94 0.00 [2.08; 2.46]
Chronological age			0.02 1.51 0.13 [-0.01; 0.05]				0.02 1.48 0.14 [-0.01; 0.05]		0.02 1.56 0.12 [-0.01; 0.05]		0.02 1.46 0.15 [-0.01; 0.05]		0.02 1.57 0.12 [-0.01; 0.05]
Gender (male)			0.08 0.86 0.39 [-0.10; 0.26]				0.09 1.00 0.32 [-0.09; 0.27]		0.10 1.01 0.31 [-0.09; 0.28]		0.09 0.97 0.33 [-0.09; 0.27]		0.10 1.04 0.30 [-0.08; 0.28]
Education			0.05 1.64 0.10 [-0.01; 0.12]				0.04 1.17 0.24 [-0.03; 0.10]		0.04 1.38 0.17 [-0.02; 0.11]		0.04 1.40 0.16 [-0.02; 0.11]		0.04 1.12 0.26 [-0.03; 0.10]
Organization's number of employees			-0.20*** -4.65 0.00 [-0.28; -0.11]				-0.18*** -3.69 0.00 [-0.28; -0.09]		-0.18*** -3.67 0.00 [-0.28; -0.08]		-0.18*** -3.65 0.00 [-0.28; -0.08]		-0.18*** -3.72 0.00 [-0.28; -0.09]
Industry dummy: Public sector			-0.28* -2.41 0.02 [-0.50; -0.05]				-0.22. -1.84 0.07 [-0.45; 0.02]		-0.22. -1.86 0.07 [-0.46; 0.02]		-0.23. -1.90 0.06 [-0.47; 0.02]		-0.21. -1.81 0.08 [-0.45; 0.02]
Health					0.12* 2.69 0.01 [0.04; 0.21]		0.12* 2.50 0.02 [0.03; 0.20]		0.11* 2.07 0.04 [0.01; 0.21]		0.11* 2.19 0.03 [0.01; 0.20]		0.12* 2.47 0.02 [0.02; 0.21]
Individual development practices					0.06 0.29 0.77 [-0.32; 0.43]		0.17 0.94 0.35 [-0.17; 0.52]		0.16 0.91 0.36 [-0.18; 0.51]		0.17 0.92 0.36 [-0.18; 0.51]		0.17 0.93 0.36 [-0.18; 0.51]
Transition-to-retirement practices					0.00 0.03 0.98 [-0.31; 0.32]		-0.04 -0.28 0.78 [-0.33; 0.24]		-0.04 -0.26 0.80 [-0.33; 0.25]		-0.04 -0.29 0.78 [-0.33; 0.24]		-0.03 -0.21 0.83 [-0.32; 0.26]
Continued employment practices					0.41** 3.32 0.00 [0.17; 0.64]		0.13 0.94 0.35 [-0.13; 0.38]		0.13 0.98 0.33 [-0.13; 0.39]		0.13 0.96 0.34 [-0.13; 0.39]		0.12 0.93 0.36 [-0.14; 0.38]
Health x Individual development practices					-0.07 -0.43 0.67 [-0.36; 0.23]		-0.06 -0.40 0.69 [-0.33; 0.24]		0.09 0.64 0.53 [-0.19; 0.38]				
Health x Transition-to-retirement practices					0.19 1.43 0.16 [-0.07; 0.46]		0.18 1.31 0.20 [-0.08; 0.43]				0.24. 1.93 0.06 [-0.01; 0.48]		
Health x Continued employment practices					0.28** 2.72 0.01 [0.08; 0.47]		0.27* 2.64 0.01 [0.08; 0.46]						0.30** 3.04 0.00 [0.11; 0.49]
Random effects													
σ^2	1.05		1.03		0.98		0.98		0.97		0.98		0.97
τ_{00}	0.16 $_{Organization}$		0.08 $_{Organization}$		0.14 $_{Organization}$		0.09 $_{Organization}$		0.09 $_{Organization}$		0.09 $_{Organization}$		0.09 $_{Organization}$
τ_{11}					0.02 $_{Organization.Health}$		0.02 $_{Organization.Health}$		0.04 $_{Organization.Health}$		0.03 $_{Organization.Health}$		0.02 $_{Organization.Health}$
ρ_{01}					-0.03 $_{Organization}$		-0.18 $_{Organization}$		-0.08 $_{Organization}$		-0.03 $_{Organization}$		-0.25 $_{Organization}$
ICC	0.13		0.07		0.14		0.1		0.12		0.11		0.11
N	99 $_{Organization}$		99 $_{Organization}$		99 $_{Organization}$		99 $_{Organization}$		99 $_{Organization}$		99 $_{Organization}$		99 $_{Organization}$
Observations	533		533		533		533		533		533		533
Marginal R^2 / Conditional R^2	0.000 / 0.131		0.077 / 0.143		0.067 / 0.198		0.109 / 0.201		0.091 / 0.203		0.097 / 0.199		0.106 / 0.200

Note. Organizational level predictors were grand-mean centered; Individual level predictors were centered at the organization's mean; Unstandardized coefficients (*γ*) with standard errors (*SE*) are shown; σ^2 = within-person variance; τ_{00} = between-person intercept variance; τ_{01} = between-person slope variance; ρ_{01} = correlation between intercept and slope; . indicates p < .1; * indicates p < .05; ** indicates p < .01; *** indicates p < .001.

Hypotheses 2 to 4 stated that the positive relations between employees' health and their preferred retirement age, as well as their late-career intentions, are moderated by three groups of organizational practices, namely individual development practices, transition-to-retirement practices, and continued employment practices. For employees' preferred retirement age, results confirmed our hypotheses regarding the three cross-level interactions: The positive relation between employees' health and their preferred retirement age is stronger in organizations with individual development practices (Hypothesis 2a: $\gamma = 1.04$, $p < .01$), transition-to-retirement practices (Hypothesis 3a: $\gamma = .60$, $p < .05$), and continued employment practices (Hypothesis 4a: $\gamma = .56$, $p < .01$). The interactions are plotted in Figure 4.2. Because this study is limited in power and because the three groups of practices are correlated, the cross-level interactions were tested using three separate models. When tested with a single model, the cross-level interaction with individual development practices is the only significant interaction ($\gamma = .86$, $p < .05$). Nevertheless, the interaction plots (Figure 4.2) and according simple slopes tests suggest that the practices differ in that transition-to-retirement particularly lowers the intended retirement age of employees in poor health, while the other two groups of practices increase the intended retirement age of employees with good health.

For participants' late-career intentions, results partially supported our cross-level interaction hypotheses. The positive relation between employees' health and their late-career intentions was significantly stronger in organizations with continued employment practices (Hypothesis 4b: $\gamma = .30$, $p < .01$). This interaction is plotted in Figure 4.3. At the same time, individual development practices (Hypothesis 2b: $\gamma = .09$, n.s.) and transition-to-retirement practices (Hypothesis 3b: $\gamma = .24$, n.s.) did not show significant cross-level interaction effects.

Notably, none of the organizational practices showed a significant direct relation with employees' late-career intentions. To test whether our control variables confounded the results, we repeated the computations without control variables. All cross-level interactions persisted without control variables as well (Hypothesis 2a: $\gamma = 1.03$, $p < .01$; Hypothesis 3a: $\gamma = .67$, $p <$

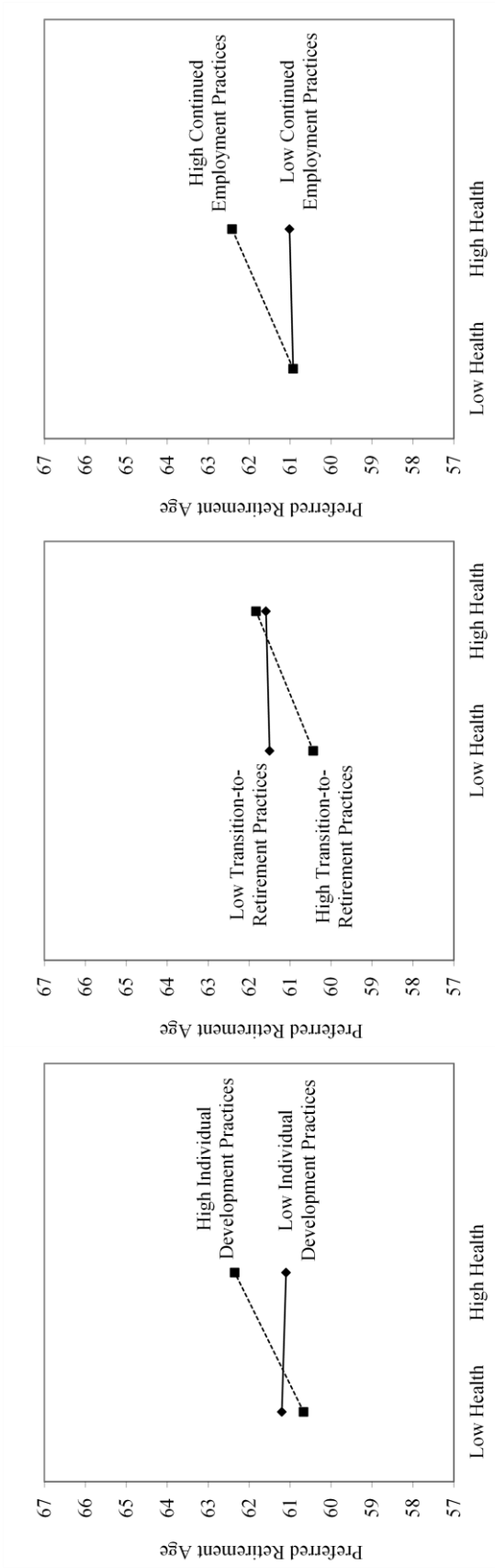
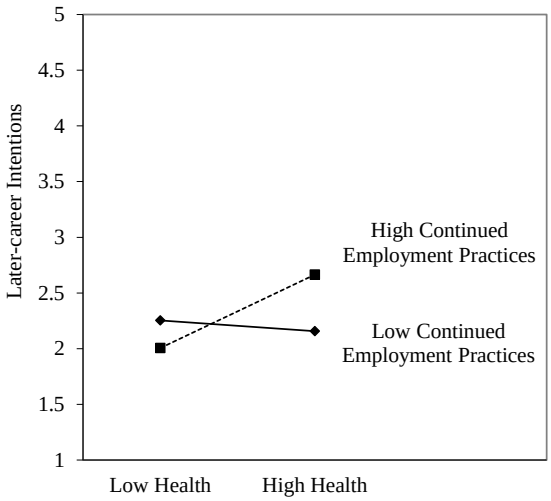


Figure 4.2

Significant cross-level interaction of organizational practices on the relation between employees' health and their preferred retirement age. Interactions were tested in separate models.

Figure 4.3

Significant cross-level interaction of organizational continued employment practices on the relation between employees' health and their intention to continue working for their current employer beyond the normal retirement age.



.05; Hypothesis 4a: $\gamma = .56$, $p < .01$; Hypothesis 4b: $\gamma = .31$, $p < .01$). In the model without control variables, however, a direct effect between continued employment practices and employees' late-career intentions age became evident ($\gamma = .40$, $p < .01$).

Discussion

This article addresses the conditional effects of organizational practices for older employees on the relationship between their health and their retirement intentions. We investigated three groups of practices that specifically address older employees' career, and in particular retirement options and decisions: first, organizational practices that foster older employees' individual development (e.g., life-long learning and career development); second, organizational practices that tailor the transition to retirement to older employees' individual needs (e.g., phased retirement); third, organizational practices that provide older employees with opportunities to continue working in later life (e.g., (re-)hiring of older employees and individualized forms of employment for employees even beyond the normal retirement age). Results from our multi-level and multi-organization study suggest that in organizations that

apply such practices, employees in good health intend to retire later and employees in poor health intend to retire earlier than in organizations not engaging in these practices.

With respect to employees' preferred retirement age, all three groups of organizational practices were found to strengthen the effect of employees' health. Individual development practices showed the strongest interaction effect, offsetting the effects of the other two practices if computed in the same model. This may be because in organizations that engage in practices tailoring the transition to retirement or that create opportunities for continued employment, strong individual development practices are likely present as well. Yet, results also suggest that individual development, transition-to-retirement, and continued employment practices differ in how they interact with employees' health. Transition-to-retirement practices were found to particularly increase intentions to retire early for employees with poor health, while these practices did not affect the retirement intentions of employees in good health. Contrarily, continued employment practices particularly increase intentions to postpone retirement for employees with good health, while those practices did not affect retirement intentions of employees with poor health (see Figure 4.3).

For employees' late-career intentions (i.e., the intention to continue working beyond the normal retirement age), continued employment practices that encompassed individualized employment deals and (re-)hiring of older employees by the organization were the only form of organizational practice that showed a stronger effect on the relation between health and employees' intention to continue working. This relationship was otherwise affected neither by individual development practices nor by transition-to-retirement practices. In line with prior research (e.g., Bal et al., 2012), this result indicates that the intention to engage in a late-career after retirement differs from employees' intended retirement age.

Consistent with previous literature, the results of this study confirm the importance not only of employees' health for their retirement intentions (e.g., Fisher et al., 2016; McGarry, 2004; Wang & Shultz, 2010) but also of organizational practices addressing older employees'

needs and abilities (e.g., Armstrong-Stassen, 2008; Bal et al., 2012). For instance, Bal et al. (2012) identified positive effects of idiosyncratic deals with older employees on their late-career intentions. This study shows similar results, insofar as that continued employment practices—including idiosyncratic deals for late careers—were directly related to older employees' preferred retirement age and late-career intentions. However, we only found this direct effect on late-career intentions, in case we did not control for organizational size or for the public sector. These results suggest that such flexible deals for prolonged careers exist particularly in small and nonpublic-sector organizations.

The main purpose of the present study, however, was to investigate not the direct effects but rather the moderating effects of organizational practices. Its results highlight that employees' health and organizational practices not only independently affect retirement intentions, but also interact. In line with the theory of work adjustment, the intention to continue working stems from the interaction between health as an individual characteristic and the work environment, characterized by practices that address older employees' individual needs. In the past, researchers have argued similarly that certain organizational climates can interact with personal resources. For example, Vignoli et al. (2019) found that in organizations characterized by age discrimination, the effect between personal resources and the desired retirement age is less strong than in organizations without age discrimination.

Furthermore, the present study extends existing research on the relation between employees' health and organizational practices. Previous research emphasized the importance of work design and occupational health practices to reduce the risk of work-related illnesses and thereby sustain employees' health. For example, Beier et al. (2020) found in longitudinal data that employees have fewer chronic health conditions if demands at work match their abilities. Yet it remains difficult for organizations to improve older employees' health directly (Bal et al., 2012), given that their health is affected by many other environmental and personal factors across their life-span (e.g., J. Ilmarinen, 2001). Thus, organizations have to cope with a

diverse health spectrum among older employees. The present study went a step further by investigating health as the independent variable. Health does not only result from a sufficient fit between personal factors and environmental factors at work, but changes in health may also impact the fit between employees and their work environment. For example, a decrease in health may impact the employee's abilities to the extent that employees are no longer able to meet demands at work. Retiring earlier or later can be a meaningful response to such differences in person-environment fit at work (Lahlouh et al., 2019). Thus, we did not focus on those practices generally benefitting older employees' health, their employability, and aging at work such as work design or health management practices, but on a set of organizational practices supporting an individualized adaptation of the environment to employees' health, by means of opportunities for individual development throughout the working life, a timely retirement planning, opportunities for an individualized transition to retirement, and opportunities to engage in a late-career. The present study reveals that these organizational practices can shape employees' preferences for an earlier or later retirement, depending on their individual abilities and needs. Actively tailoring the retirement transition to employees needs lead to intentions to retire earlier for employees with poor health, while options to continue working beyond normal retirement age resulted in intentions to retire later for employees in good health.

Theoretical and Practical Implications

Our findings have three important theoretical implications. First, they shed light on the relationship between older employees' health and their retirement intentions. Although this relationship has been widely supported in the literature, evidence on how this relationship can be influenced remains very limited. This study is among the first to investigate the role that organizational practices can play in shaping retirement intentions in alignment with employees' health.

Second, the present study contributes to our understanding of person-environment fit at work and its relevance for the retirement transition. We answered calls by Fischer et al. (2016) and Wang and Shultz (2010) for a more integrated perspective on individual and organizational factors affecting retirement decisions and timing. Previous research has focused primarily on either individual or organizational factors that influence retirement intentions. However, the theory of work adjustment, widely applied in research to explain career decisions in general, argues for an integrated perspective incorporating both personal and environmental factors. By focusing on individual health, we addressed a crucial individual antecedent of retirement intentions. In that respect, the present study might inspire researchers to further investigate the conditional effects of organizational practices by taking into account older employees' additional needs and capabilities.

Third, the present study provides additional empirical evidence for the structure of organizational practices that benefit an aging workforce. The study used three domains of the Later Life Workplace Index (Wilckens, Wöhrmann, Deller, et al., 2020), which were shown to be sufficiently distinct and to manifest on the organizational level. Although not hypothesized, this study identified that some practices are more relevant than others for shaping the effects on certain outcomes (e.g., continued employment practices showed cross-level interaction effects with health on the intention to continue working, whereas individual development and transition-to-retirement practices did not). Moreover, investigated practices showed contrary direct effects on older employees preferred retirement age—continued employment practices increased the preferred retirement age, while transition-to-retirement practices resulted in preferences for an earlier retirement (results showed that this was particularly the case for older employees with poor health). For these reasons, it is important to distinguish different sets of practices that are beneficial to older employees. Thus, our findings support recent endeavors to approach organizational practices for the aging workforce in a multifaceted manner, for example, the classification of HR practices along the selection, optimization, compensation

framework (cf. Baltes & Baltes, 1990) by Kooij et al. (2014), and, more recent the scale developments by Eppler-Hattab, Doron, and Meshoulam (2020), as well as Wilckens et al. (2020). These studies provide differentiated taxonomies to investigate the role that HR practices can play in motivating an aging workforce to work.

With regards to practice, the present paper suggests that late careers, if employees' health permits, are supported by organizational practices that foster employees' individual development and provide them with a perspective on their retirement transition and on continued employment at retirement age. Practices such as opportunities for training and development, regular meetings between managers and employees to discuss opportunities for individual development, and opportunities to engage in a late-career bolster older employees' willingness to extend their careers in the event of good health. Simultaneously, practices such as an early retirement planning between managers and employees, opportunities for individual retirement transition as phased retirement, and organizational support for the retirement transition also sensitize employees to the possibility of an early exit from work in the event of poor health. Employees may then benefit from the practices in terms of more differentiated and individualized consideration of their retirement transition and its fit with their individual health resources. For organizations, the practices may foster extended career intentions of employees in good health. Ultimately, these practices may contribute to solving labor force shortages and, consequently, benefit society.

Limitations and Future Directions

Despite the strengths of this multi-level, multi-organization study, several potential limitations are worth to be mentioned and may guide future research. Because this study is based on cross-sectional data, the inferred causal effects may be called into question. For the present study, however, we do not believe this to be a concern for conceptual and methodological reasons. Conceptually, the variables used in this study vary in kind. The

independent variable (health) is a long-term individual characteristic. The dependent variables (retirement intentions) are behavioral intention constructs. Being potentially subject to short-term change, these are therefore more likely to depend on long-term health than the other way around. Methodologically, we based our findings on interactions for which reverse causality is generally less a concern. Organizations may actively engage in particular organizational practices (e.g., occupational health programs) in response to changes in their employees' health or to ambitions to hasten or delay employees' retirement timing. Yet such reverse effects on organizational practices do not impact the investigated cross-level interactions.

Furthermore, findings from this study are limited to employees' retirement timing intentions, which do not necessarily manifest in their later behavior. Future research may require longitudinal data to investigate, whether the effects on retirement timing intentions identified in this study also manifest regarding older employees' behavior and their actual retirement decisions.

In addition, findings of this study may be limited to the German labor market context in which it was conducted. Compared with other labor markets, the labor market in Germany is characterized by a relatively fixed retirement age. However, a normal retirement age is deeply rooted in the regulatory frameworks and social security systems of many other countries, too (cf. Calvo et al., 2013; Radl, 2012). Consequently, the theoretical arguments of this study are not based on specifics of the German context. Nevertheless, future research needs to confirm the results in other national contexts.

Moreover, studies on older employees' health may be subject to the 'healthy worker effect': Employees with poor health are more likely to retire early, meaning that they are underrepresented in the sample of employees. This structural bias deflates the negative relation between age and health. For the present study, however, results are affected neither by a biased relationship between age and health nor by a slightly skewed distribution of health among the sample. Thus, the 'healthy worker effect' does not seem to have an impact on our findings.

Furthermore, the interaction perspective on organizational and individual factors predicting retirement intentions and decisions offers further research opportunities. For example, researchers should focus on other individual characteristics such as economic status, psychological factors (e.g., goals, needs), family factors, and domestic factors. Moreover, the health measure used in the present study is a general health construct, meaning that future research could investigate potential differences between physical, mental, and cognitive health. Furthermore, the interactions between organizational practices and health may be dependent on other factors within the organization, for example, attitudes and biases by the managers implementing the practices or a general age-friendly or age-discriminating organizational climate.

Additionally, the mechanisms proposed by this study to explain the identified relationships remain to be confirmed in future research. Mediators underlying the theoretical foundation of our hypotheses could not be included in this paper's multi-level study, because the questionnaire was limited in length. For example, further studies could include person-environment fit and future time perspective to test whether these constructs can fully explain the effects of organizational practices on the relationship between health and retirement intensions.

From the organizational standpoint, this study focused on three groups of practices that target sustained careers for those employees whose health allows them to continue working, and also appropriate transitions to retirement for the other employees. However, different sets of organizational practices may also specifically accommodate employees with poor health status, leading to different effects. For example, the Later Life Workplace Index outlines further organizational practices such as work design and health management practices. Furthermore, we did not include environmental factors outside the work environment. As Wang and Shultz (2010) pointed out, employees' late-career intentions may also be conceptualized as misfit between personal characteristics and characteristics of the retirement environment. Including

factors from the retirement environment would therefore open up additional research opportunities.

General Discussion

This dissertation describes the conceptualization, operationalization, validation, and initial application of the Later Life Workplace Index (LLWI), a multifaceted, comprehensive measure of organizational practices for the aging workforce. Beginning with its foundations in theory and international qualitative research (Article 1 and 2), the present dissertation provides evidence for the measurement quality of the new measure (Article 3), and its appropriateness to study organizational practices for the aging workforce on the organizational level (Article 4).

The set of articles addresses this dissertation's research objective to add to our understanding of organizational practices related to aging at work from various perspectives. Specifically, the dissertation aimed to answer three overarching research questions. These were (1) to holistically conceptualize organizational practices for the aging workforce, (2) to propose how organizational practices for the aging workforce can be measured validly, and (3) to add to our understanding, how individual organizational practices effect retirement intentions and, consequently, how organizations can facilitate aging at work. Thereby, the articles respond to calls from the literature for more organizational level research on employers' response to the aging workforce (Henkens et al., 2018), for a "multidimensional model […], which includes shared perceptions of more specific age-related organizational policies, norms, practices, and procedures" by Zacher and Yang (2016 p. 9), and for a more integrated person-environment fit approach to retirement decisions (Fisher et al., 2016; Wang & Shultz, 2010). Findings are rigorously obtained from empirical studies using qualitative data (Article 1), simulated quantitative data (Article 2), quantitative individual field data (Article 3), and quantitative multi-level field data (Article 4).

Theoretical Implications

Building upon previous research by Wöhrmann et al. (2018) and findings from the *Age Smart Employer Award* in New York City (Finkelstein et al., 2013) the first article holistically integrates organizational practices for the aging workforce. As a result, the proposed LLWI comprehensively summarizes the relevant organizational practices in the context of aging at work. By integrating German and U.S. data the taxonomy is not tied to a specific cultural, legal or societal context. Besides the empirical evidence from qualitative data underlying the taxonomy, we extensively reviewed the literature for further practices (see also Article 3), but did not find any which would not fit into the LLWI's taxonomy. We acknowledge that, by design, the practices are not uniform, but very different with regards to their effects within organizations and on older workers, and regarding the conditions, in which organizations meaningfully implement the practices. The practices are captured in nine distinct domains, namely an age-friendly organizational climate, leadership style, work design practices, individual development practices, health management practices, knowledge management practices, practices tailoring the retirement transition to the organization's and older workers' needs, practices facilitating continued employment beyond normal retirement age, and practices providing health and retirement coverage for employees. The methodological advancements presented in the second article provided evidence that the domains of practices are sufficiently distinct. While some of the LLWI's domains such as an age-friendly organizational climate or work design practices have intensively been researched already (e.g., Armstrong-Stassen, 2008), the index also incorporates certain practices limitedly researched so far. For example, practices tailoring the transition to retirement are fairly new concepts and we do not yet sufficiently know how they affect older workers and their employment. Consequently, the LLWI not only provides a multifaceted framework for those practices well recognized in the literature, but also adds new practices to form a comprehensive taxonomy.

This dissertation's second research objective to develop a psychometrically sound, multifaceted measure of organizational practices for the aging workforce is addressed in the third article. The three studies comprising the article describe the development of the new measure. They provide quantitative evidence for the structure of the LLWI's nine domains and the measurement quality of the 80-item measure, both, in terms of reliability and validity. The measured nine domains of practices were positively related to person-job fit, person-organization fit, employees' engagement, well-being, and post-retirement work intentions. At the same time, the practices were negatively related to stress and turnover intentions. By providing the LLWI as a comprehensive, multifaceted, and psychometrically sound measure, results from this dissertation have the potential to shape our view on organizational practices in future. In line with previous research (e.g., Armstrong-Stassen & Templer, 2006; Kooij et al., 2014) the LLWI's multifaceted structure supports the distinctiveness of practices. Thus, the findings question to what extent *age-friendly organizational practices* can meaningfully be investigated by a lump-sum assessment. Certainly, the age-related organizational practices are correlated. Nevertheless, results from the fourth article suggest that different organizational practices have different effects. Thus, single-dimensional measures of organizational practices may not be sufficient to fully understand how organizations can improve employment of older employees. Previous measures of organizational practices related to aging at work are single-dimensional (e.g., Taneva & Arnold, 2018) or lack either a sound conceptualization or psychometric validation (e.g., lists of organizational practices). Thus, the LLWI has the potential to substantially improve how researchers investigate practices, by which organizations can facilitate and successfully deal with an aging workforce. Results of this dissertation suggest that organizational practices for the aging workforce are not uniform, but should be disentangled to understand their individual effects. The LLWI allows future studies to research organizational practices on a granular level, to compare the practices among each other, and to derive more tailored recommendations for practice.

Beyond the scope of measuring organizational practices for the aging workforce, this dissertation may also contribute to research on organizational practices in general. In line with previous research by for example, Bal et al. (2016) and Taneva and Arnold (2018) the qualitative index development revealed that organizational practices relevant in the context of aging at work are only partially formally enacted by the management (cf. Article 1 and 3). Instead many practices materialize for the workers as informal habits. To understand how organizational practices influence individuals it thus is important to not only focus on practices the management is engaged in, but also those informal practices present for the employees independently of managerial decisions. Further, the LLWI is constructed as a formative index, because the nine index' domains (i.e. sets of practices with a common intentions as individual development or heath management practices) are not indicators for an underlying causing factor such as overall age-friendliness but are instead independent fields of managerial decisions. This does not only apply to age-related organizational practices but also to practices in general. Thus, research should question how organizational practices relevant in other fields are measured and whether the measurement model properly reflects assumed causality.

Finally, the fourth article uses the LLWI in a multi-level study among organizations in Germany to provide initial answers to this dissertation's third research objective to investigate the role of organizational practices for older employees' retirement intentions in more detail. Specifically, the study investigated the interaction between older employees' health and three organizational practices, namely individual development for older employees, transition-to-retirement practices, and practices allowing to continue working beyond normal retirement age. The study enhances our understanding of aging at work by taking a person-environment fit perspective on the role of organizational practices in the retirement process. In line with previous research, results support the importance of both, organizational practices and individual characteristics (Fisher et al., 2016). Results suggest that organizational practices interact with older workers' individual characteristics. Transition-to-retirement practices

stronger reduced the intended retirement age of older workers with poor health, while individual development and continued employment practices particularly increased the intended retirement age of healthy older workers. Research may thus stronger focus on the different effects between organizational practices related to aging at work in order to derive meaningful recommendations for organizations which practices to implement under which conditions and for what purposes. Thus, this dissertation's contribution to our understanding of the differences between relevant organizational practices for older workers likely interact with research on the heterogeneity among older workers (e.g., different health, socio-economic status, gender, family status, needs) in future studies (see also Taylor et al., 2016). The organizational practices incorporated in the LLWI share the notion that organizations should engage in individualized solutions that fit the various needs and capabilities of older and aging workers to deal with the aging workforce (cf. Oostrom et al., 2016). Not only quantitative evidence from the fourth article, but also qualitative results from the first article emphasize the importance and benefit of practices that take individual differences of the workers into account. Taking individual differences into account likely is not only relevant for organizational practices for an aging workforce, but may also guide organizational practices in other fields as for example, disability and inclusion.

Moreover, this dissertation adds to scale development, and more generally, qualitative content analysis from a methodological point of view in the second article. Development of a new measure requires a thorough conceptual framework which is most often based on qualitative evidence. Results of the methodological article on reliability assessment for multi-valued qualitative data enhance researchers' ability to test their conclusion from qualitative sources, such as interviews or observations. Results generally confirm the approach to compute reliability coefficients for multi-valued data proposed by Krippendorff (2019). However, the article's methodological advancements including an alternative difference metric have the potential to improve how researchers deduct meaning from complex qualitative data, in which

a single unit of analysis, such as a thought in an interview or a scene during an observation may involve more than one concept. The methodological advancements improved how the LLWI was derived from qualitative interviews on organizations' response to the aging workforce. The methodology to derive Krippendorff's alpha for multi-valued data as a reliability measure may support qualitative researchers in similar challenges in future.

Practical Implications

Development of the LLWI took place with the objective to create a measure valuable in practice. Decision-makers in organizations facing an aging workforce need to know how to specifically improve their organizations response. Disentangling the various practices and allowing organizations to assess themselves regarding these practices is a starting point to derive areas for improvement. Even though research has just started investigating how these organizational practices benefit older employees and their organizations in detail, this dissertation provides organizations with the possibility to evaluate the status quo. Benchmark studies may support identifying gaps between an organization and its peers in the industry. The multi-organizational study presented in the fourth article of this dissertation was used to provide organizations with an analysis of their practices compared to other participating organizations comparable in size and industry sector. In many organizations results from the LLWI study were used as a starting point to improve employment conditions for the organizations' older employees. Thus, this dissertation also marks the LLWI's first application in practice. The interest of the various organizations to take part in the study, as well as the fact that many organizations used their results from the study as a starting point for improvement projects, showed the benefit the LLWI may play for practice in future. The LLWI has the potential to contribute to more evidence-based management of the aging workforce in practice (see also Henkens et al., 2018).

Moreover, the LLWI's framework with its nine domains may also help to structure ongoing political and regulatory efforts to improve practices for older employees in organizations. For example, the International Organization for Standardization (2020) took up the LLWI as a basis for a new guideline on age-inclusive workforces through the course of the LLWI project. Thereby, the evidence incorporated in the LLWI likely reaches a broader audience beyond the scientific community.

However, results of this dissertation also show that different domains of practices covered by the LLWI may have different effects depending on the employees' characteristics (see Article 4). The organizational level study underlying the fourth article revealed that the presence of practices varies significantly between organizations of different kind (e.g., public vs. private organizations, large vs. small organizations; see also van Dalen et al., 2015). For example, small organizations may lack sufficient resources to implement practices (see Article 3). But, organizational practices covered by the LLWI may also have very different relevance depending on the organizational context. Further research is required to examine contextual factors and characteristics of the employee that effect the benefit of the practices (see Article 4). Consequently, application of the LLWI in practice provides management with their organization's status quo regarding practices relevant for the aging workforce. Yet, it is likely misleading to conclude that each and every organization and their employees would benefit similarly from fully implementing all practices.

Limitations and Future Directions

Despite the conceptual and empirical strength of this dissertation's articles, there are several limitations and starting points for future research. Most prominent, this dissertation provides a thoroughly developed measure for organizational practices regarding aging at work, but does not yet provide many answers regarding how these individual practices effect employment of older employees on the individual and the organizational level. The third article

provides some cross-sectional evidence that measured practices are positively related to several behavioral and motivational variables of older employees, such as work engagement and person-organization fit. Moreover, results of the multi-level study presented in the fourth article suggest that individual development practices and continued employment practices motivate older employees in good health to postpone retirement, while transition-to-retirement practices result in an earlier retirement of employees with poor health. Yet, further relevant questions regarding the effects of the practices have not yet been answered: For example, future research might investigate to what extend organizations benefit from implementing certain practices in terms of their older employees' performance. Previous research provided evidence that some of the LLWI's domains benefit performance of older employees (e.g., individual development practices researched by Boehm et al., 2014). Yet, the overall picture on the diverse set of organizational practices is still missing. Implementation of practices in organizations most often requires a positive business case, which does not yet exist for most practices. To promote and improve aging at work, future research needs to prove the (long-term) benefit of practices not only for the workers but also for organizations, and to take into account related cost to provide a holistic cost-benefit analysis from the workers', the organizations', and society's perspective. Furthermore, the third article outlined substantial differences between organizations in terms of which practices they implement. Understanding the reasons for these differences likely provides further insights in how organizations can support an aging workforce. Researchers investigated the effect of employers' age discrimination on their response to the aging workforce (Henkens, 2005). However, other contextual factors, such as the severity of labor force shortages and employers' knowledge about the practices may also play a significant role.

The present dissertation provides the LLWI as a validated measure, which was developed based on qualitative evidence from Germany and the U.S. Yet, the qualitative data for the U.S. was collected in New York City only. To what extend this New York City data is sufficiently informative for the entire U.S. remains to be validated. In other cultures and other

regulatory schemes, different organizational practices that are not yet part of the LLWI may be relevant. Moreover, the index' operationalization has so far only been validated among organizations in Germany. Retirement timing and pension schemes in Germany are characterized by a statutory pension, for which employees become eligible at a comparably fixed age (OECD, 2017). Future studies could operationalize the LLWI in other languages and validate the index' structure among organizations from other regulatory schemes. This would allow cross-cultural research to examine cultural differences in the requirements, antecedents, and effects of organizational practices for older workers, the organizations, and society.

Furthermore, the multi-level study among numerous organizations presented in the fourth article provided evidence for the organizational nature of the LLWI domains researched. Data from the study not only supports a significant portion of shared variance between respondents of the same organizations for the three domains addressed in the article, but also for the other six LLWI domains. Yet, the study also shows that respondents from the same organization vary in their assessments. Particularly, managers and older workers rate practices within their organization differently, which might result from differences in knowledge about offered practices and utilization of the practices between the groups of respondents (see Nishii et al., 2018). To further enhance our understanding of the practices, future research might investigate how the practices become effective for older employees, to what extend and under what conditions they are utilized, and why ratings differ among respondents of the same organization.

More generally, development of the LLWI provide new opportunities for researchers in the fields of aging at work, because the LLWI allows them to validly assess and disentangle different organizational practices based on a thorough conceptual framework. The multifacetedness allows researchers to tap into the differences between individual practices. However, individual domains of the LLWI may also be used as single-dimensional scales of the respective practices.

Summary and Outlook

To summarize, this dissertation advances our understanding of age-related organizational practices. By providing the LLWI as a multifaceted measure of practices this dissertation contributes to future research in the field of aging at work and allows for assessing practices within organizations. Findings on the LLWI's practices emphasize the importance to disentangle organizational practices for the aging workforce and to further research their effects. The new measure eases and improves the assessment of the practices in research and practice. Thus, it may help us to better understand how "employers [can] make an aging staff 'work'" (Henkens et al., 2018, p. 809). In combination with the visibility on the situation within organizations provided by the LLWI, this may induce better decisions on how to improve employment of older employees and result in extended, yet healthy, productive, and motivated working lives.

References

Adams, G. A., & Beehr, T. A. (1998). Turnover and retirement: A comparison of their similarities and differences. *Personnel Psychology*, *51*(3), 643–665. https://doi.org/10.1111/j.1744-6570.1998.tb00255.x

Akkermans, J., de Lange, A. H., van der Heijden, B. I. J. M., Kooij, D. T. A. M., Jansen, P. G. W., & Dikkers, J. S. E. (2016). What about time? Examining chronological and subjective age and their relation to work motivation. *Career Development International*, *21*(4), 419–439. https://doi.org/10.1108/CDI-04-2016-0063

Armstrong-Stassen, M. (2008). Organisational practices and the post-retirement employment experience of older workers. *Human Resource Management Journal*, *18*(1), 36–53. https://doi.org/10.1111/j.1748-8583.2007.00057.x

Armstrong-Stassen, M., & Lee, S. H. (2009). The effect of relational age on older Canadian employees' perceptions of human resource practices and sense of worth to their organization. *The International Journal of Human Resource Management*, *20*(8), 1753–1769. https://doi.org/10.1080/09585190903087156

Armstrong-Stassen, M., & Schlosser, F. (2008). Benefits of a supportive development climate for older workers. *Journal of Managerial Psychology*, *23*(4), 419–437. https://doi.org/10.1108/02683940810869033

Armstrong-Stassen, M., & Schlosser, F. (2011). Perceived organizational membership and the retention of older workers. *Journal of Organizational Behavior*, *32*(2), 319–344. https://doi.org/10.1002/job.647

Armstrong-Stassen, M., & Templer, A. J. (2006). The response of Canadian public and private sector human resource professionals to the challenge of the aging workforce. *Public Personnel Management*, *35*(3), 247–260.

https://doi.org/10.1177/009102600603500307

Bal, P. M., De Jong, S. B., Jansen, P. G. W., & Bakker, A. B. (2012). Motivating employees to work beyond retirement: A multi-level study of the role of i-deals and unit climate. *Journal of Management Studies, 49*(2), 306–331. https://doi.org/10.1111/j.1467-6486.2011.01026.x

Bal, P. M., Jansen, P. G. W., van der Velde, M. E. G., de Lange, A. H., & Rousseau, D. M. (2010). The role of future time perspective in psychological contracts: A study among older workers. *Journal of Vocational Behavior, 76*(3), 474–486. https://doi.org/10.1016/j.jvb.2010.01.002

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). Cambridge University Press.

Barnes-Farrell, J. L. (2003). Beyond health and wealth: Attitudinal and other influences on retirement decision-making. In G. A. Adams & T. A. Beehr (Eds.), *Retirement: Reasons, processes, and results* (pp. 159–187). Springer.

Barnes-Farrell, J. L. (2005). Older workers. In J. Barling, E. K. Kelloway, & M. R. Frone (Eds.), *Handbook of work stress* (pp. 431–454). Sage.

Barnes, H., Smeaton, D., & Taylor, R. (2009). *An ageing workforce: The employer's perspective*. Institute for Employment Studies. https://www.employment-studies.co.uk/resource/ageing-workforce-employers-perspective

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beier, M. E., Torres, W. J., Fisher, G. G., & Wallace, L. E. (2020). Age and job fit: The relationship between demands–ability fit and retirement and health. *Journal of*

Occupational Health Psychology, *25*(4), 227–243. https://doi.org/10.1037/ocp0000164

Bliese, P. D. (1998). Group size, ICC values, and group-level correlations: A simulation. *Organizational Research Methods*, *1*(4), 355–373. https://doi.org/10.1177/109442819814001

Boehm, S. A., & Dwertmann, D. J. G. (2015). Forging a single-edged sword: Facilitating positive age and disability diversity effects in the workplace through leadership, positive climates, and HR practices. *Work, Aging and Retirement*, *1*(1), 41–63. https://doi.org/10.1093/workar/wau008

Boehm, S. A., Kunze, F., & Bruch, H. (2014). Spotlight on age-diversity climate: The impact of age-inclusive HR practices on firm-level outcomes. *Personnel Psychology*, *67*(3), 667–704. https://doi.org/10.1111/peps.12047

Börsch-Supan, A., & Weiss, M. (2016). Productivity and age: Evidence from work teams at the assembly line. *The Journal of the Economics of Ageing*, *7*(C), 30–42. https://doi.org/10.1016/j.jeoa.2015.12.001

Börsch-Supan, A., & Wilke, C. B. (2004). *The German public pension system: How it was, how it will be* (No. 10525; NBER Working Paper Series). National Bureau of Economic Research.

Boselie, P., Dietz, G., & Boon, C. (2005). Commonalities and contradictions in HRM and performance research. *Human Resource Management Journal*, *15*(3), 67–94. https://doi.org/10.1111/j.1748-8583.2005.tb00154.x

Breyer, B., & Bluemke, M. (2016). Deutsche Version der Positive and Negative Affect Schedule PANAS (GESIS Panel) [German version of the positive and negative affect schedule]. *Zusammenstellung Sozialwissenschaftlicher Items Und Skalen*. https://doi.org/10.6102/zis242

Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*,

38(1), 25–56. https://doi.org/10.1207/S15327906MBR3801_2

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Bureau of Labor Statistics. (2016). *Employee tenure in 2016*. https://www.bls.gov/news.release/pdf/tenure.pdf

Bureau of Labor Statistics. (2017). *Labor force statistics*. https://www.bls.gov/cps/cpsaat11b.htm

Cable, D. M., & DeRue, D. S. (2002). The convergent and discriminant validity of subjective fit perceptions. *Journal of Applied Psychology*, *87*(5), 875–884. https://doi.org/10.1037/0021-9010.87.5.875

Cadiz, D. M., Brady, G., Rineer, J. R., & Truxillo, D. M. (2019). A review and synthesis of the work ability literature. *Work, Aging and Retirement, 5*(1), 114–138. https://doi.org/10.1093/workar/way010

Cahill, K. E., Giandrea, M. D., & Quinn, J. F. (2015). Retirement patterns and the macroeconomy, 1992–2010: The prevalence and determinants of bridge jobs, phased retirement, and reentry among three recent cohorts of older Americans. *The Gerontologist*, *55*(3), 384–403. https://doi.org/10.1093/geront/gnt146

Calvo, E., Sarkisian, N., & Tamborini, C. R. (2013). Causal effects of retirement timing on subjective physical and emotional health. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(1), 73–84. https://doi.org/10.1093/geronb/gbs097

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, *56*(2), 81–105.

Carstensen, L. L. (1987). Age-related changes in social activity. In L. L. Carstensen & B. A. Edelstein (Eds.), *Pergamon General Psychology Series, Vol. 146. Handbook of clinical gerontology* (pp. 222–237). Pergamon Press.

Carstensen, L. L. (1991). Socioemotional selectivity theory: Social activity in life-span context. *Annual Review of Gerontology and Geriatrics*, *11*(1), 195–217.

Carstensen, L. L. (1992). Social and emotional patterns in adulthood: Support for socioemotional selectivity theory. *Psychology and Aging*, *7*(3), 331–338. https://doi.org/10.1037/0882-7974.7.3.331

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234–246. https://doi.org/10.1037/0021-9010.83.2.234

Cheung, F., & Wu, A. M. S. (2013). Older workers' successful aging and intention to stay. *Journal of Managerial Psychology*, *28*(6), 645–660. https://doi.org/10.1108/JPM-09-2011-0062

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Conen, W. S., Henkens, K., & Schippers, J. (2012). Employers' attitudes and actions towards the extension of working lives in Europe. *International Journal of Manpower*, *33*(6), 648–665. https://doi.org/10.1108/01437721211261804

Conway, E. (2004). Relating career stage to attitudes towards HR practices and commitment: Evidence of interaction effects? *European Journal of Work and Organizational Psychology*, *13*(4), 417–446. https://doi.org/10.1080/13594320444000155

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*,

16(3), 297–334. https://doi.org/10.1007/BF02310555

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

Davis, M. A. (2003). Factors related to bridge employment participation among private sector early retirees. *Journal of Vocational Behavior*, *63*(1), 55–71. https://doi.org/10.1016/S0001-8791(02)00016-7

Dawis, R. V., England, G. W., & Lofquist, L. H. (1964). *A theory of work adjustment*. University of Minnesota Press.

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment*. University of Minnesota Press.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.

Dingemans, E., & Henkens, K. (2015). How do retirement dynamics influence mental well-being in later life? A 10-year panel study. *Scandinavian Journal of Work, Environment & Health, 41*(1), 16–23. https://doi.org/10.5271/sjweh.3464

Dingemans, E., Henkens, K., & van Solinge, H. (2016). Access to bridge employment: Who finds and who does not find work after retirement? *The Gerontologist*, *56*(4), 630–640. https://doi.org/10.1093/geront/gnu182

Edwards, J. R. (1991). Person-job fit: A conceptual integration, literature review, and methodological critique. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (Vol. 6, pp. 283–357). John Wiley & Sons.

Edwards, J. R., & Shipp, A. J. (2007). The relationship between person-environment fit and outcomes: An integrative theoretical framework. In C. Ostroff & T. A. Judge (Eds.), *Perspectives on organizational fit* (pp. 209–258). Jossey-Bass. https://doi.org/10.4324/9780203810026

Eppler-Hattab, R., Meshoulam, I., & Doron, I. (2020). Conceptualizing age-friendliness in

workplaces: Proposing a new multidimensional model. *The Gerontologist*, *60*, 12–21. https://doi.org/10.1093/geront/gny184

Eurofound. (2012). *Income from work after retirement in the EU*. Publications Office of the European Union. https://doi.org/10.2806/42416

Eurofound. (2015). *Job tenure in turbulent times*. Publications Office of the European Union.

European Commission, & Economic Policy Committee. (2017). *The 2018 ageing report: Underlying assumptions and projection methodologies*. Publications Office of the European Union. https://doi.org/10.2765/15373

Fasbender, U., Deller, J., Wang, M., & Wiernik, B. M. (2014). Deciding whether to work after retirement: The role of the psychological experience of aging. *Journal of Vocational Behavior*, *84*(3), 215–224. https://doi.org/10.1016/j.jvb.2014.01.006

Fasbender, U., Wöhrmann, A. M., Wang, M., & Klehe, U.-C. (2019). Is the future still open? The mediating role of occupational future time perspective in the effects of career adaptability and aging experience on late career planning. *Journal of Vocational Behavior*, *111*(1), 24–38. https://doi.org/10.1016/j.jvb.2018.10.006

Feldman, D. C. (2012). Feeling like it's time to retire: A fit perspective on early retirement decisions. In M. Wang (Ed.), *The Oxford handbook of retirement* (pp. 280–292). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199746521.013.0109

Feng, G. C. (2013). Factors affecting intercoder reliability: A Monte Carlo experiment. *Quality and Quantity*, *47*(5), 2959–2982. https://doi.org/10.1007/s11135-012-9745-9

Finkelstein, R., Roher, S., & Owusu, S. (2013). *Compendium of strategies and practices*. The New York Academy of Medicine. http://aging.columbia.edu/sites/default/files/AgeSmartEmployer_Compendium.pdf

Fisher, G. G., Chaffee, D. S., & Sonnega, A. (2016). Retirement timing: A review and recommendations for future research. *Work, Aging and Retirement*, *2*(2), 230–261. https://doi.org/10.1093/workar/waw001

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378–382. https://doi.org/10.1037/h0031619

Furunes, T., & Mykletun, R. J. (2010). Age discrimination in the workplace: Validation of the Nordic Age Discrimination Scale (NADS). *Scandinavian Journal of Psychology, 51*(1), 23–30. https://doi.org/10.1111/j.1467-9450.2009.00738.x

Gałecki, A., & Burzykowski, T. (2013). *Linear mixed-effects models using R*. Springer. https://doi.org/10.1007/978-1-4614-3900-4

Garner, W. R. (1960). Rating scales, discriminability, and information transmission. *Psychological Review*, *67*(6), 343–352. https://doi.org/10.1037/h0043047

Gehman, J., Glaser, V. L., Eisenhardt, K. M., Gioia, D., Langley, A., & Corley, K. G. (2018). Finding theory–method fit: A comparison of three qualitative approaches to theory building. *Journal of Management Inquiry*, *27*(3), 284–300. https://doi.org/10.1177/1056492617706029

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research, 25*(2), 186. https://doi.org/10.2307/3172650

Giarini, O. (2012). The four pillars project: 25 years on. *Life and Pensions Newsletter, 50*, 1–2.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Göbel, C., & Zwick, T. (2013). Are personnel measures effective in increasing productivity of old workers? *Labour Economics, 22*(1), 80–93. https://doi.org/10.1016/j.labeco.2012.11.005

Gobeski, K. T., & Beehr, T. A. (2009). How retirees work: Predictors of different types of bridge employment. *Journal of Organizational Behavior, 30*(3), 401–425.

https://doi.org/10.1002/job.547

Green, P. E., & Rao, V. R. (1970). Rating scales and information recovery. How many scales and response categories to use? *Journal of Marketing*, *34*(3), 33–39. https://doi.org/10.2307/1249817

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, *61*(1), 29–48. https://doi.org/10.1348/000711006X126600

Harper, M. C., & Shoffner, M. F. (2004). Counseling for continued career development after retirement: An application of the theory of work adjustment. *Career Development Quarterly*, *52*(3), 272–284. https://doi.org/10.1002/j.2161-0045.2004.tb00648.x

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. https://doi.org/10.1080/19312450709336664

Henkens, K. (2005). Stereotyping older workers and retirement: The managers' point of view. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, *24*(4), 353–366. https://doi.org/10.1353/cja.2006.0011

Henkens, K., Van Dalen, H. P., Ekerdt, D. J., Hershey, D. A., Hyde, M., Radl, J., Van Solinge, H., Wang, M., & Zacher, H. (2018). What we need to know about retirement: Pressing issues for the coming decade. *Gerontologist*, *58*(5), 805–812. https://doi.org/10.1093/geront/gnx095

Hennekam, S., & Herrbach, O. (2013). HRM practices and low occupational status older workers. *Employee Relations*, *35*(3), 339–355. https://doi.org/10.1108/01425451311320512

Hennekam, S., & Herrbach, O. (2015). The influence of age-awareness versus general HRM practices on the retirement decision of older workers. *Personnel Review*, *44*(1), 3–21. https://doi.org/10.1108/PR-01-2014-0031

Henry, H., Zacher, H., & Desmette, D. (2017). Future time perspective in the work context: A systematic review of quantitative studies. *Frontiers in Psychology*, *8*(413), 1–22. https://doi.org/10.3389/fpsyg.2017.00413

Hershey, D. A., & Henkens, K. (2014). Impact of different types of retirement transitions on perceived satisfaction with life. *The Gerontologist*, *54*(2), 232–244. https://doi.org/10.1093/geront/gnt006

Hertweck, M. S., & Sigrist, O. (2015). The ins and outs of German unemployment: A transatlantic perspective. *Oxford Economic Papers*, *67*(4), 1078–1095. https://doi.org/10.1093/oep/gpv031

Hesketh, B., Griffin, B., Dawis, R., & Bayl-Smith, P. (2015). Extensions to the dynamic aspects of the Retirement Transition and Adjustment Framework (RTAF): Adjustment behaviors, work styles, and identity. *Work, Aging and Retirement*, *1*(1), 79–91. https://doi.org/10.1093/workar/wau004

Hesketh, B., Griffin, B., & Loh, V. (2011). A future-oriented retirement transition adjustment framework. *Journal of Vocational Behavior*, *79*(2), 303–314. https://doi.org/10.1016/j.jvb.2011.03.008

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104–121. https://doi.org/10.1177/109442819800100106

Hsu, L. M., & Field, R. (2003). Interrater agreement measures: Comments on Kappan, Cohen's Kappa, Scott's π, and Aickin's α. *Understanding Statistics*, *2*(3), 205–219. https://doi.org/10.1207/S15328031US0203_03

Humphreys, L. G., & Montanelli Jr., R. G. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*, *10*(2), 193–205. https://doi.org/10.1207/s15327906mbr1002_5

Huselid, M. A. (1995). The impact of human resource management practices on turnover,

productivity, and corporate financial performance. *Academy of Management Journal*, *38*(3), 635–672. https://doi.org/10.2307/256741

Ilmarinen, J. (2001). Aging workers. *Occupational and Environmental Medicine*, *58*(8), 546–546. https://doi.org/10.1136/oem.58.8.546

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*(8), 1445–1460. https://doi.org/10.1080/00140139.2015.1005167

International Organization for Standardization. (2020). *ISO/TC314 Aging Societies WG1 Aging workforce*. https://www.iso.org/standard/76420.html

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199–218. https://doi.org/10.1086/376806

Jiang, K., Lepak, D. P., Hu, J., & Baer, J. C. (2012). How does human resource management influence organizational outcomes? A meta-analytic investigation of mediating mechanisms. *Academy of Management Journal*, *55*(6), 1264–1294. https://doi.org/10.5465/amj.2011.0088

Johnson, R., Kawachi, J., & Lewis, E. (2009). *Older workers on the move: Recareering in later life*. AARP Public Policy Institute. https://www.urban.org/research/publication/older-workers-move-recareering-later-life

Jung, P., & Kuhn, M. (2014). Labour market institutions and worker flows: Comparing Germany and the US. *The Economic Journal*, *124*(581), 1317–1342. https://doi.org/10.1111/ecoj.12118

Karpinska, K., Henkens, K., & Schippers, J. (2013). Hiring retirees: Impact of age norms and stereotypes. *Journal of Managerial Psychology*, *28*(7/8), 886–906.

https://doi.org/10.1108/JMP-07-2013-0223

Khilji, S. E., & Wang, X. (2006). 'Intended' and 'implemented' HRM: The missing linchpin in strategic human resource management research. *The International Journal of Human Resource Management, 17*(7), 1171–1189. https://doi.org/10.1080/09585190600756384

Kim, H., & Stoner, M. (2008). Burnout and turnover intention among social workers: Effects of role stress, job autonomy and social support. *Administration in Social Work, 32*(3), 5–25. https://doi.org/10.1080/03643100801922357

Kim, S., & Feldman, D. C. (2000). Working in retirement: The antecedents of bridge employment and its consequences for quality of life in retirement. *The Academy of Management Journal, 43*(6), 1195–1210. https://doi.org/10.5465/1556345

Klaffke, M. (Ed.). (2014). *Generationen-Management: Konzepte, Instrumente, Good-Practice-Ansätze* [Management of generations: Concepts, instruments, and good-practices]. Springer Gabler.

Kleinke, K. (2017). Multiple imputation under violated distributional assumptions: A systematic evaluation of the assumed robustness of predictive mean matching. *Journal of Educational and Behavioral Statistics, 42*(4), 371–404. https://doi.org/10.3102/1076998616687084

Kooij, D. T. A. M. (2015). Successful aging at work: The active role of employees. *Work, Aging and Retirement, 1*(4), 309–319. https://doi.org/10.1093/workar/wav018

Kooij, D. T. A. M., Guest, D. E., Clinton, M., Knight, T., Jansen, P. G. W., & Dikkers, J. S. E. (2013). How the impact of HR practices on employee well-being and performance changes with age. *Human Resource Management Journal, 23*(1), 18–35. https://doi.org/10.1111/1748-8583.12000

Kooij, D. T. A. M., Jansen, P. G. W., Dikkers, J. S. E., & de Lange, A. H. (2010). The influence of age on the associations between HR practices and both affective commitment and job satisfaction: A meta-analysis. *Journal of Organizational Behavior,*

31(8), 1111–1136. https://doi.org/10.1002/job.666

Kooij, D. T. A. M., Jansen, P. G. W., Dikkers, J. S. E., & de Lange, A. H. (2014). Managing aging workers: A mixed methods study on bundles of HR practices for aging workers. *The International Journal of Human Resource Management*, *25*(15), 2192–2212. https://doi.org/10.1080/09585192.2013.872169

Kooij, D. T. A. M., Kanfer, R., Betts, M., & Rudolph, C. W. (2018). Future time perspective: A systematic review and meta-analysis. *Journal of Applied Psychology*, *103*(8), 867–893. https://doi.org/10.1037/apl0000306

Kooij, D. T. A. M., & Van de Voorde, K. (2011). How changes in subjective general health predict future time perspective, and development and generativity motives over the lifespan. *Journal of Occupational and Organizational Psychology*, *84*(2), 228–247. https://doi.org/10.1111/j.2044-8325.2010.02012.x

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). Jossey-Bass.

Krippendorff, K. (1970). Bivariate agreement coefficients for reliability of data. *Sociological Methodology*, *2*, 139–150.

Krippendorff, K. (1992). Recent developments in reliability analysis. *42nd Annual Meeting of the International Communication Association*. http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED352390

Krippendorff, K. (2004). Measuring the reliability of qualitative text analysis data. *Quality and Quantity*, *38*(6), 787–800. https://doi.org/10.1007/s11135-004-8107-7

Krippendorff, K. (2011). Agreement and information in the reliability of coding. *Communication Methods and Measures*, *5*(2), 93–112. https://doi.org/10.1080/19312458.2011.568376

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Sage.

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). Sage.

Krippendorff, K., & Craggs, R. (2016). The reliability of multi-valued coding of data. *Communication Methods and Measures, 10*(4), 181–198. https://doi.org/10.1080/19312458.2016.1228863

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, *49*(1), 1–49. https://doi.org/10.1111/j.1744-6570.1996.tb01790.x

Kunze, F., Boehm, S., & Bruch, H. (2013). Organizational performance consequences of age diversity: Inspecting the role of diversity-friendly HR policies and top managers' negative age stereotypes. *Journal of Management Studies*, *50*(3), 413–442. https://doi.org/10.1111/joms.12016

Lacy, S., Watson, B. R., Riffe, D., & Lovejoy, J. (2015). Issues and best practices in content analysis. *Journalism and Mass Communication Quarterly*, *92*(4), 791–811. https://doi.org/10.1177/1077699015607338

Lahlouh, K., Lacaze, D., & Huaman-Ramirez, R. (2019). Bridge employment and full retirement intentions: The role of person-environment fit. *Personnel Review*, *48*(6), 1469–1490. https://doi.org/10.1108/PR-02-2018-0067

Lawler III, E. E. (1992). *The ultimate advantage: Creating the high involvement organization*. Jossey-Bass.

Li, Y., Gong, Y., Burmeister, A., Wang, M., Alterman, V., Alonso, A., & Robinson, S. (2021). Leveraging age diversity for organizational performance: An intellectual capital perspective. *Journal of Applied Psychology*, *106*(1), 71–91. https://doi.org/10.1037/apl0000497

Loch, C. H., Sting, F. J., Bauer, N., & Mauermann, H. (2010). How BMW is defusing the demographic timebomb. *Harvard Business Review*, *88*(3), 99–102.

Longman, R. S., Cota, A. A., Holden, R. R., & Fekken, G. C. (1989). A regression equation for the parallel analysis criterion in principal components analysis: Mean and 95th percentile eigenvalues. *Multivariate Behavioral Research, 24*(1), 59–69. https://doi.org/10.1207/s15327906mbr2401_4

Mahon, J. F., & Millar, C. C. J. M. (2014). ManAGEment: The challenges of global age diversity for corporations and governments. *Journal of Organizational Change Management, 27*(4), 553–568. https://doi.org/10.1108/JOCM-05-2014-0100

Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (11th ed.) [Qualitaitve content analysis: Foundations and methodology]. Beltz.

McGarry, K. (2004). Health and retirement: Do changes in health affect retirement expectations? *The Journal of Human Resources, 39*(3), 624–648. https://doi.org/10.2307/3558990

McGonagle, A. K., Fisher, G. G., Barnes-Farrell, J. L., & Grosch, J. W. (2015). Individual and work factors related to perceived work ability and labor force outcomes. *Journal of Applied Psychology, 100*(2), 376–398. https://doi.org/10.1037/a0037974

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085

Mermin, G. B. T., Johnson, R. W., & Murphy, D. P. (2007). Why do boomers plan to work longer? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 62*(5), S286–S294. https://doi.org/10.1093/geronb/62.5.S286

Moen, P. (2016). *Encore adulthood. Boomers on the edge of risk, renewal, and purpose*. Oxford University Press.

Morrow-Howell, N., Gonzales, E., James, J. B., Matz-Costa, C., & Putnam, M. (2018). Advance long and productive lives. In R. Fong, J. Lubben, & R. P. Barth (Eds.), *Grand challenges for social work and society* (pp. 81–102). Oxford University Press.

Naegele, G., & Walker, A. (2006). *Guide to good practice in age management*. Office for

Official Publications of the European Communities.

Nishii, L. H., Khattab, J., Shemla, M., & Paluch, R. M. (2018). A multi-level process model for understanding diversity practice effectiveness. *Academy of Management Annals, 12*(1), 37–82. https://doi.org/10.5465/annals.2016.0044

Nissly, J. A., Barak, M. E. M., & Levin, A. (2005). Stress, social support, and workers' intentions to leave their jobs in public child welfare. *Administration in Social Work, 29*(1), 79–100. https://doi.org/10.1300/J147v29n01_06

Oakman, J., & Wells, Y. (2016). Working longer: What is the relationship between person-environment fit and retirement intentions? *Asia Pacific Journal of Human Resources, 54*(2), 207–229. https://doi.org/10.1111/1744-7941.12075

OECD. (2017). *Pensions at a glance 2017: OECD and G20 indicators*. OECD Publishing. https://doi.org/10.1787/pension_glance-2017-en

Oostrom, J. K., Pennings, M., & Bal, P. M. (2016). How do idiosyncratic deals contribute to the employability of older workers? *Career Development International, 21*(2), 176–192. https://doi.org/10.1108/CDI-08-2015-0112

Osborne, J. W. (2012). Psychological effects of the transition to retirement. *Canadian Journal of Counselling and Psychotherapy, 46*(1), 45–58.

Pak, K., Kooij, D. T. A. M., De Lange, A. H., Heuvel, S., & Van Veldhoven, M. J. P. M. (2021). The influence of human resource practices on perceived work ability and the preferred retirement age: A latent growth modelling approach. *Human Resource Management Journal, 31*(1), 311–325. https://doi.org/10.1111/1748-8583.12304

Pak, K., Kooij, D. T. A. M., de Lange, A. H., & Van Veldhoven, M. J. P. M. (2019). Human resource management and the ability, motivation and opportunity to continue working: A review of quantitative studies. *Human Resource Management Review, 29*(3), 336–352. https://doi.org/10.1016/j.hrmr.2018.07.002

Parker, S. K., & Andrei, D. M. (2020). Include, individualize, and integrate: Organizational

meta-strategies for mature workers. *Work, Aging and Retirement*, *6*(1), 1–7. https://doi.org/10.1093/workar/waz009

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(2), 287–312. https://doi.org/10.1207/S15328007SEM0802_7

Podsakoff, P. M., Ahearne, M., & MacKenzie, S. B. (1997). Organizational citizenship behavior and the quantity and quality of work group performance. *Journal of Applied Psychology*, *82*(2), 262–270. https://doi.org/10.1037/0021-9010.82.2.262

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Polat, T., Bal, P. M., & Jansen, P. G. W. (2017). How do development HR practices contribute to employees' motivation to continue working beyond retirement age? *Work, Aging and Retirement*, *3*(4), 366–378. https://doi.org/10.1093/workar/wax007

Posthuma, R. A., & Campion, M. A. (2009). Age stereotypes in the workplace: Common stereotypes, moderators, and future research directions. *Journal of Management*, *35*(1), 158–188. https://doi.org/10.1177/0149206308318617

Radl, J. (2012). Too old to work, or too young to retire? The pervasiveness of age norms in Western Europe. *Work, Employment and Society*, *26*(5), 755–771. https://doi.org/10.1177/0950017012451644

Rauvola, R. S., Rudolph, C. W., Ebbert, L. K., & Zacher, H. (2019). Person–environment fit and work satisfaction: Exploring the conditional effects of age. *Work, Aging and Retirement*, *6*(2), 101–117. https://doi.org/10.1093/workar/waz011

Riff, D. (2014). *Analyzing media messages*. Routledge.

https://doi.org/10.4324/9780203551691

Rodríguez-López, J., & Romaguera, S. (2004). The Hausdorff fuzzy metric on compact sets. *Fuzzy Sets and Systems, 147*(2), 273–283. https://doi.org/10.1016/j.fss.2003.09.007

Rudolph, C. W. (2016). Lifespan developmental perspectives on working: A literature review of motivational theories. *Work, Aging and Retirement*, *2*(2), 130–158. https://doi.org/10.1093/workar/waw012

Rudolph, C. W., de Lange, A. H., & van der Heijden, B. I. J. M. (2015). Adjustment processes in bridge employment: Where we are and where we need to go. In P. M. Bal, D. T. A. M. Kooij, & D. M. Rousseau (Eds.), *Aging workers and the employee-employer relationship* (pp. 221–242). Springer. https://doi.org/10.1007/978-3-319-08007-9_13

Sackreuther, I., Mergenthaler, A., Cihlar, V., Micheel, F., Lessenich, S., Lippke, S., Schneider, N. F., & Staudinger, U. M. (2017). *(Un-)ruhestände in Deutschland: Übergänge, Potenziale und Lebenspläne älterer Menschen im Wandel* [Late careers and retirement in Germany: Changing transitions, potentials, and plans for life of older people]. Bundesinstitut für Bevölkerungsforschung.

Salminen, H., von Bonsdorff, M., & von Bonsdorff, M. (2019). Investigating the links between resilience, perceived HRM practices, and retirement intentions. *Evidence-Based HRM: A Global Forum for Empirical Scholarship*, *7*(1), 75–92. https://doi.org/10.1108/EBHRM-02-2018-0011

Sautier, L., Scherwath, A., Weis, J., Sarkar, S., Bosbach, M., Schendel, M., Ladehoff, N., Koch, U., & Mehnert, A. (2015). Assessment of work engagement in patients with hematological malignancies: Psychometric properties of the German version of the Utrecht Work Engagement Scale 9 (UWES-9). *Die Rehabilitation*, *54*(5), 297–303. https://doi.org/10.1055/s-0035-1555912

Schaufeli, W. B., Shimazu, A., Hakanen, J., Salanova, M., & De Witte, H. (2019). An ultra-short measure for work engagement. *European Journal of Psychological Assessment*,

35(4), 577–591. https://doi.org/10.1027/1015-5759/a000430

Schuett, S. (2014). *Demografie-Management in der Praxis - Mit der Psychologie des Alterns wettbewerbsfähig bleiben* [Demography management in practice - staying competitive through psychology of aging]. Springer.

Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, *19*(3), 321. https://doi.org/10.1086/266577

Shaw, J. D., Delery, J. E., Jenkins, G. D. J., & Gupta, N. (1998). An organization-level analysis of voluntary and involuntary turnover. *The Academy of Management Journal*, *41*(5), 511–525. https://doi.org/https://doi.org/10.5465/256939

Silver, M. P., Settels, J., Schafer, M. H., & Schieman, S. (2019). Getting the hours you want in the preretirement years: Work hour preferences and mismatch among older Canadian workers. *Work, Aging and Retirement*, *5*(2), 175–188. https://doi.org/10.1093/workar/way015

Social Security Administration. (2018). *Retirement benefits* (05-10035, ICN 457500). https://www.ssa.gov/pubs/EN-05-10035.pdf

Stanton, J. M., Balzer, W. K., Smith, P. C., Parra, L. F., & Ironson, G. (2001). A general measure of work stress: The Stress in General scale. *Educational and Psychological Measurement*, *61*(5), 866–888. https://doi.org/10.1177/00131640121971455

Statistisches Bundesamt. (2015). *Bevölkerung Deutschlands bis 2060: 13. koordinierte Bevölkerungsvorausberechnung* [Population in Germany until 2060: 13th coordinated projection]. Statistisches Bundesamt. https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/VorausberechnungBevoelkerung/BevoelkerungDeutschland2060Presse5124204099004.pdf?__blob=publicationFile

Statistisches Bundesamt. (2019). *Erwerbsbeteiligung der Bevölkerung* [Labor market participation in Germany]. Statistisches Bundesamt.

https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Erwerbstaetigkeit/Publikationen/Downloads-Erwerbstaetigkeit/erwerbsbeteiligung-bevoelkung-2010410187004.pdf?__blob=publicationFile

Staudinger, U. M., Finkelstein, R., Calvo, E., & Sivaramakrishnan, K. (2016). A global view on the effects of work on health in later life. *The Gerontologist*, *56*(S2), S281–S292. https://doi.org/10.1093/geront/gnw032

Sullivan, S. E., & Al Ariss, A. (2019). Employment after retirement: A review and framework for future research. *Journal of Management*, *45*(1), 262–284. https://doi.org/10.1177/0149206318810411

Tan, P. N., Kumar, V., & Srivastava, J. (2004). Selecting the right objective measure for association analysis. *Information Systems*, *29*(4), 293–313. https://doi.org/10.1016/S0306-4379(03)00072-3

Taneva, S. K., & Arnold, J. (2018). Thriving, surviving and performing in late career: A mixed-method study of pathways to successful aging in organizations. *Work, Aging and Retirement*, *4*(2), 189–212. https://doi.org/10.1093/workar/wax027

Taylor, P., Loretto, W., Marshall, V., Earl, C., & Phillipson, C. (2016). The older worker: Identifying a critical research agenda. *Social Policy and Society*, *15*(4), 675–689. https://doi.org/10.1017/S1474746416000221

Taylor, P., & Walker, A. (1998). Policies and practices towards older workers: A framework for comparative research. *Human Resource Management Journal*, *8*(3), 61–76. https://doi.org/10.1111/j.1748-8583.1998.tb00174.x

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. https://doi.org/10.1177/0022022106297301

Tinsley, H. E. A., & Weiss, D. J. (2000). Interrater reliability and agreement. In E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical*

modeling (pp. 95–124). Academic Press. https://doi.org/10.1016/B978-012691360-6/50005-7

Tisch, A. (2015). Firms' contribution to the internal and external employability of older employees: Evidence from Germany. *European Journal of Ageing*, *12*(1), 29–38. https://doi.org/10.1007/s10433-014-0323-y

Topa, G., & Alcover, C.-M. (2015). Psychosocial factors in retirement intentions and adjustment: A multi-sample study. *Career Development International*, *20*(4), 384–408. https://doi.org/10.1108/CDI-09-2014-0129

Topa, G., Moriano, J. A., Depolo, M., Alcover, C.-M., & Morales, J. F. (2009). Antecedents and consequences of retirement planning and decision-making: A meta-analysis and model. *Journal of Vocational Behavior*, *75*(1), 38–55. https://doi.org/10.1016/j.jvb.2009.03.002

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, *84*(3), 167–176. https://doi.org/10.1159/000376585

United Nations. (2019). *World population prospects 2019*. United Nations.

van Dalen, H. P., & Henkens, K. (2018). Why demotion of older workers is a nogGo area for managers. *The International Journal of Human Resource Management*, *29*(15), 2303–2329. https://doi.org/10.1080/09585192.2016.1239214

van Dalen, H. P., Henkens, K., & Wang, M. (2015). Recharging or retiring older workers? Uncovering the age-based strategies of European employers. *The Gerontologist*, *55*(5), 814–824. https://doi.org/10.1093/geront/gnu048

van den Berg, T. I. J., Elders, L. A. M., & Burdorf, A. (2010). Influence of health and work on early retirement. *Journal of Occupational and Environmental Medicine*, *52*(6), 576–583. https://doi.org/10.1097/JOM.0b013e3181de8133

van Solinge, H., & Henkens, K. (2010). Living longer, working longer? The impact of

subjective life expectancy on retirement intentions and behaviour. *The European Journal of Public Health*, *20*(1), 47–51. https://doi.org/10.1093/eurpub/ckp118

van Solinge, H., & Henkens, K. (2014). Work-related factors as predictors in the retirement decision-making process of older workers in the Netherlands. *Ageing and Society*, *34*(9), 1551–1574. https://doi.org/10.1017/S0144686X13000330

Vandenberg, R. J., Richardson, H. A., & Eastman, L. J. (1999). The impact of high involvement work processes on organizational effectiveness. *Group & Organization Management*, *24*(3), 300–339. https://doi.org/10.1177/1059601199243004

Verquer, M. L., Beehr, T. A., & Wagner, S. H. (2003). A meta-analysis of relations between person–organization fit and work attitudes. *Journal of Vocational Behavior*, *63*(3), 473–489. https://doi.org/10.1016/S0001-8791(02)00036-2

Vignoli, M., Zaniboni, S., Chiesa, R., Alcover, C., Guglielmi, D., & Topa, G. (2019). Maintaining and engaging older workers at work: The trigger role of personal and psychosocial resources. *The International Journal of Human Resource Management*, Advance online publication, 1–23. https://doi.org/10.1080/09585192.2019.1579252

Visser, M., Lössbroek, J., & van der Lippe, T. (2020). The use of HR policies and job satisfaction of older workers. *Work, Aging and Retirement*, Advance online publication, 1–19. https://doi.org/10.1093/workar/waaa023

von Bonsdorff, M. E., Zhou, L., Wang, M., Vanhala, S., von Bonsdorff, M. B., & Rantanen, T. (2018). Employee age and company performance: An integrated model of aging and human resource management practices. *Journal of Management*, *44*(8), 3124–3150. https://doi.org/10.1177/0149206316662314

WAI-Netzwerk. (2015). *WAI Manual: Anwendung des Work-Ability Index* [WAI manual: Application of the work-ability index]. http://wai-netzwerk.uni-wuppertal.de/picture/upload/file/WAI-Manual.pdf

Wang, M., & Shultz, K. S. (2010). Employee retirement: A review and recommendations for

future investigation. *Journal of Management, 36*(1), 172–206. https://doi.org/10.1177/0149206309347957

Wang, M., Zhan, Y., Liu, S., & Shultz, K. S. (2008). Antecedents of bridge employment: A longitudinal investigation. *Journal of Applied Psychology, 93*(4), 818–830. https://doi.org/10.1037/0021-9010.93.4.818

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing, 27*(3), 236–247. https://doi.org/10.1016/j.ijresmar.2010.02.004

Wilckens, M. R., Wöhrmann, A. M., Adams, C., Deller, J., & Finkelstein, R. (2020). Integrating the German and US perspective on organizational practices for later life work: The Later Life Work Index. In S. Czaja, J. Sharit, & J. James (Eds.), *Current and emerging trends in aging and work* (pp. 59–79). Springer. https://doi.org/10.1007/978-3-030-24135-3_4

Wilckens, M. R., Wöhrmann, A. M., Deller, J., & Wang, M. (2020). Organizational practices for the aging workforce: Development and validation of the Later Life Workplace Index. *Work, Aging and Retirement*, Advance online publication. https://doi.org/10.1093/workar/waaa012

Wöhrmann, A. M., Deller, J., & Pundt, L. (2018). Complementing AAI at the meso level: The Silver Work Index. In A. Zaidi, S. Harper, K. Howse, G. Lamura, & J. Perek-Bialas (Eds.), *Building evidence for active ageing policies* (pp. 75–94). Palgrave Macmillan. https://doi.org/10.1007/978-981-10-6017-5_5

Wöhrmann, A. M., Deller, J., & Wang, M. (2013). Outcome expectations and work design

characteristics in post-retirement work planning. *Journal of Vocational Behavior*, *83*(3), 219–228. https://doi.org/10.1016/j.jvb.2013.05.003

Wöhrmann, A. M., Deller, J., & Wang, M. (2014). A mixed-method approach to post-retirement career planning. *Journal of Vocational Behavior*, *84*(3), 307–317. https://doi.org/10.1016/j.jvb.2014.02.003

Wöhrmann, A. M., Fasbender, U., & Deller, J. (2017). Does more respect from leaders postpone the desire to retire? Understanding the mechanisms of retirement decision-making. *Frontiers in Psychology*, *8*(1400), 1–11. https://doi.org/10.3389/fpsyg.2017.01400

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Wright, P. (2002). Desegregating HRM: A review and synthesis of micro and macro human resource management research. *Journal of Management*, *28*(3), 247–276. https://doi.org/10.1016/S0149-2063(02)00128-9

Zacher, H. (2015a). Successful aging at work. *Work, Aging and Retirement*, *1*(1), 4–25. https://doi.org/10.1093/workar/wau006

Zacher, H. (2015b). The importance of a precise definition, comprehensive model, and critical discussion of successful aging at work. *Work, Aging and Retirement*, *1*(4), 320–333. https://doi.org/10.1093/workar/wav020

Zacher, H., & Frese, M. (2009). Remaining time and opportunities at work: Relationships between age, work characteristics, and occupational future time perspective. *Psychology and Aging*, *24*(2), 487–493. https://doi.org/10.1037/a0015425

Zacher, H., Rauvola, R. S., & Rudolph, C. W. (2019). Innovating career counselling to manage the transitions to bridge employment and retirement. In J. G. Maree (Ed.), *Handbook of innovative career counselling* (pp. 173–192). Springer.

https://doi.org/10.1007/978-3-030-22799-9_11

Zacher, H., & Yang, J. (2016). Organizational climate for successful aging. *Frontiers in Psychology, 7*(1007), 1–12. https://doi.org/10.3389/fpsyg.2016.01007

Zaniboni, S. (2015). The interaction between older workers' personal resources and perceived age discrimination affects the desired retirement age and the expected adjustment. *Work, Aging and Retirement*, *1*(3), 266–273. https://doi.org/10.1093/workar/wav010

Zygmont, C., & Smith, M. R. (2014). Robust factor analysis in the presence of normality violations, missing data, and outliers: Empirical questions and possible solutions. *The Quantitative Methods for Psychology*, *10*(1), 40–55. https://doi.org/10.20982/tqmp.10.1.p040

Appendix

Appendix A. Dimensions and Indicators of the LLWI

Table A.1

Dimensions and indicators of the Later Life Work Index

Dimensions
Indicators

OC Organizational Culture

The organizational culture dimension includes the set standards and actions of an employer shaped by the mission and values of the organization. An organizational culture that fosters good management of employees just before and in retirement age especially promotes equal opportunities and a positive image for all age groups. Indicators are:

OC1 Equality of opportunity: Initial conditions should be the same for every employee regardless of age. Further, no discrimination or stigmatization due to age should occur. Each employee therefore has the same opportunities, for example, participation in training and professional qualification or in the need of downsizing.

OC2 Positive image of age: Prevailing beliefs and attitudes regarding older employees are shaped by a positive attitude within the organization. Aging should be understood as an individual change process of competencies, motivation, values, and behavior. Opportunities should be recognized, valued and realized. For example, by identifying and assigning tasks which correspond to the specific competencies of older individuals.

OC3 Open and target group-oriented communication: The organization is characterized by a differentiated image of age that is communicated through external and internal representation of the organization. This explicitly includes open and transparent exchange between employees and their managers regarding retirement and/or continued opportunities for work. Positive images representing all age groups within the employee magazine, on the intranet or website are another example.

LE Leadership

The leadership dimension includes the responsibility of organizational executives to harness the potential of employees at all ages and particularly just before and in retirement age. This is achieved through the consideration of each individual employee's strengths and by showing appreciation for their talents and contributions. Indicators are:

LE1 Appreciation: Managers of an organization should have an appreciative attitude towards their employees of all ages, manifested through a consistent demonstration of respect and kindness. Managers should reward the experience and achievements of their employees by offering higher levels of job autonomy and responsibility. Celebrating milestones and farewells are another way to convey gratitude, particularly when an employee is going into retirement.

LE2 Responsiveness to individuality: Managers of an organization should be sensitive to individual needs and events that occur at different life stages. They should also take into account each individual's personality and performance capability. Managers are responsible for recognizing and harnessing individual potential regardless of age and for creating performance-enhancing conditions. Among other factors, this includes the consideration of employees' wishes and suggestions regarding the design of their work space as well as the consideration of individual life circumstances, such as the need to care for family.

WD Work Design

The work design dimension includes the adaptation of work location, time and physical space to fit the individual needs and abilities of employees, relieve strain, and increase job satisfaction and efficiency. Indicators are:

WD1 Flexible work time arrangements: The organization should allow employees to change their work time depending on individual needs. Specific solutions will depend on the nature of an employee's work. Options for flexibility could include a long- or short-term switch to part time, offering flextime, job sharing, the possibility of swapping shifts, and unpaid leaves.

WD2 Flexible workplaces: When possible, employees should be able to choose their work location based on their individual needs and what is most efficient. Examples include the facilitation and technical support of home-office-solutions or the installation of silent work places within the office.

WD3 Work according to capabilities: Employees should have adequate jobs corresponding to their individual physical and mental performance capability and resilience. If not the case, this could be realized through a temporary or permanent change to another role that is less straining. Swapping jobs or reconsidering and adapting work flows should also be taken into consideration.

WD4 Ergonomic working conditions: The work place should be designed according to ergonomic requirements and should also take into account the individual circumstances of the employee. For example, occupational safety measures should be taken and supportive equipment and/or tools should be provided.

HM Health Management

The health management dimension includes all organizational activities that aim to maintain and promote employees' health and work ability. Health management should be characterized by a holistic approach addressing not only specific interventions but also health-promoting work design and leadership. Indicators are:

HM1 Availability of physical exercise and nutrition opportunities: Initiatives to strengthen health and work ability should be offered, such as company sports activities, active breaks, and nutritional guidance.

HM2 Workplace Medical treatment: Measures should be taken to help employees avoid medical conditions and assistance to aid in the recovery of sick employees should be offered. Examples include company doctors, on-site medical check-ups and physical therapy, along with wellness programs.

Dimensions

Indicators

HM3 Health promotion: Measures should be taken to disseminate knowledge about healthy behaviors to help employees make responsible and healthy decisions. This could be done by providing information on healthy living. Moreover, managers should act as role models for healthy behaviors and promote a healthy work environment. This includes taking part in physical exercise, nutrition opportunities, and related programs themselves, as well as encouraging a sustainable work-life balance.

ID Individual Development

Employees should be supported in their professional and personal development during their entire work life. A special emphasis is put on the importance of lifelong learning through continued education and training. There should also be opportunities for career development through internal advancement and promotions. Indicators are:

ID1 Continuous development planning: Planning for each individual employee's future should be done on an ongoing basis at all ages and stages of the work life. This could be done through individual meetings between managers and employees and by providing professional workshops that allow for self-reflection on abilities, competencies, and goals.

ID2 Appropriate solutions for training and development: The organization should provide further training and education aligned with the individual employee's professional, educational, and life experience as well as with organizational goals. Further, training content and methods should be targeted towards specific groups. Examples of appropriate training and development solutions are workshops, seminars and industry conferences, training for new technologies or equipment, cross-training, and internships for people of all ages. These training and development opportunities can be facilitated on-site or through reimbursement of tuition or fees.

ID3 Enabling development steps and job changes: Modifications to an employee's current position, function or job should be made possible to reflect the specific competencies and development interests of an individual. For example, this could be achieved by increasing job responsibilities, inclusion into other projects, or a horizontal or vertical change of position, which could also mean an additional apprenticeship or a new job within a different department.

KM Knowledge Management

The knowledge management dimension includes procedures for the transfer, exchange, and conservation of knowledge between different generations of employees. Indicators are:

KM1 Institutionalized knowledge transfer: Institutionalized structures that transfer knowledge from experienced employees to their successors should be in place. This can be achieved through mentoring and "buddy" programs or through a systematic knowledge transfer process before employees leave the organization for retirement.

KM2 Inter-generative collaboration: The organization should allow for mutual transfer of knowledge and experience between generations. This transfer goes in both directions, young to old, and old to young. Its structure is not necessarily determined by the organization. For example, collaboration can happen within intergenerational pairs or age-mixed teams.

TR Transition to Retirement

The transition into retirement dimension includes the necessary conversations, planning, and workplace solutions for any employee who is on the verge of retiring. Information and counseling should be provided to help the employee transition. Indicators are:

TR1 Timely transition planning: Managers should talk with employees about their personal plans for entering the retirement stage, including a succession plan. Potential transition scenarios should be actively discussed to find individual solutions, for example, through annual employee interviews.

TR2 Phased retirement and individualized transition solutions: Generic solutions for the transition into retirement should be tailored according to employees' individual needs. Flexibility and imagination should be present when designing the employee's individual transition into retirement. Phased retirement through a gradual reduction of working time should be offered companywide. Phased retirement can take place over a shorter or longer period of time, depending on needs.

TR3 Counselling for retirement life preparation: Organizations should support their employees in preparing mentally for the life change of retirement by providing advising and counseling. Employees should be motivated to actively design their retirement life prior to transition. For example, individual preparation can be fostered through a structured approach that reflects individual expectations and plans. There may also be opportunities to establish alternative activities beyond employment.

TR4 Continuous inclusion and maintaining contact: Tools should be in place to maintain contact with employees even after their retirement and to help them stay engaged as part of the organization. This could be facilitated through an active management of relationships by means of an alumni network, invitations to organizational events or by allowing for voluntary work.

CE Continued Employment

The continued employment dimension includes the organizational design and employment options for employees at retirement age. This includes former employees of the organization as well as external employees looking for continued employment.

CE1 Individualized employment options: Employment options for individuals, who would otherwise be fully retired, should be offered systematically. To ensure employment options are meaningful for both the organization and the employee, integration of those employees into the organization should be strategically planned. For example, they might be brought in on a temporary basis at peak production times. Tasks and work time should be adaptable to the individual employee. This can be achieved through alternative contract forms such as consulting and mentoring work or flexible work time arrangements with generally fewer hours than a full-time position.

CE2 (Re-) hiring of older employees: Older individuals, particularly including already and almost retired employees should be specifically addressed by job marketing, hiring, and re-employment processes. This is achieved through age-friendly communication of job offers and the use of alternative marketing paths to address external as well as internal individuals. This explicitly includes employees with long careers in other industries or companies.

Dimensions

Indicators

CC Health & Retirement Coverage

Organizations should support their employees with retirement savings and insurance coverage, if not sufficiently provided by public systems. Requirements vary due to different regulations and social systems. The support may be a direct financial benefit or put into practice as individual planning and assistance. Indicators are:

CC1 Retirement savings and pensions: Employees should be offered options for retirement savings, if not sufficiently covered by public systems. Organizations may include pensions and retirement saving accounts into their full compensation packages, offer optional saving possibilities to be opened by the employees individually, and support their employees in timely planning and organization of their retirement savings.

CC2 Insurances and financial emergency support: Organizations should offer health-related insurance coverage, if not sufficiently covered by public systems. This includes (additional) health-, disability-, care- or life insurances, which particularly cover risks that increase with age. Additional financial support may be offered in case of family emergencies, as for example, in a case of nursing care or child sickness.

Appendix B. Practices Identified by the Age Smart Employer Award

1. Flexibility.
- (a) Employees can choose/ swap shifts,
- (b) Work hours can change as needed,
- (c) Employees can work from home/ telecommute,
- (d) Employees can take unpaid leave

2. Benefits.
- (a) Profit sharing/ bonuses,
- (b) Paid time off,
- (c) Health insurance,
- (d) Retirement savings account,
- (e) Pension,
- (f) Tuition Reimbursement,
- (g) Wellness program,
- (h) Financial planning assistance,
- (i) Caregiver support program,
- (j) Paid family leave

3. Environment.
- (a) Ergonomic working conditions,
- (b) Employees can make adjustments/ suggestions

4. Work Atmosphere.
- (a) Supportive/ team mentality,
- (b) Celebrates mile-stones,
- (c) English as a second language classes offered,
- (d) "Family-like" environment,
- (e) Fund/ financial held for employee emergencies

5. Training.
- (a) Apprenticeships/ Internships,
- (b) Cross-training,
- (c) Mentorship/ paired learning,
- (d) Classes/ licensure are paid for,
- (e) Work-shops/ seminars during the workday,
- (f) Employees can attend industry conferences/ events,
- (g) Training for new technology or equipment

6. Job Restructuring.
- (a) Job sharing,
- (b) Roles change based on ability of employee,
- (c) New roles created based on ability of employee

7. Retirement.
- (a) Has a succession plan,
- (b) Can dial-down/ up work as an employee needs,
- (c) Retirees can do part-time/ consulting work,
- (d) Retirees can volunteer

8. Hiring.
- (a) Hires people who have retired from other company,
- (b) Actively recruits older workers 50+,
- (a) Hires people with long careers in other industries,
- (b) Promotes from within/ develops staff

Appendix C. R Program to Compute Krippendorff's Alpha for Multi-Valued Data

```
# This macro computes Krippendorff's alpha reliability estimate for judgments of multi value
# nominal judgements, for any number of judges, with or without missing data.
# The macro assumes the data is set up as a file with judges as the variables and the units
# being judged in the rows. The entries in the data matrix should be the coding
# (comma separated list of quantified or numerically coded nominal judgments) applied
# to the unit in that row by the judge in that column. Once the macro is
# activated (by running the command set below), the syntax is
#
# KALPHA_multi (data = a, p_value = b, alpha_min = c, include_empty_set=TRUE,
#     detail=1, sep=",", diff_fct="Krippendorff_default", emptySetLabel="none")
#
# where 'a' is the data matrix (if data is NULL or omitted, the macro will prompt for a table
# import), 'b' is the level of significance used to calculate confidence intervals (default is
# 0.05), 'c' is the minimum C-Alpha used to calculate the Type I error. 'include_empty_set'
# can be set to FALSE to exclude codings, for with no value has been assigned. 'detail' allows
# to adjust the amount of information printed during execution. Set detail to 1 if you want to
# print the coincidence and difference matrices. 'sep' can be set to define the character
# delimiting multiple values in the input file. 'diff_fct' defines the difference function to
# use. Currently, either 'Krippendorff_default' or 'Krippendorff_set_theoretic_difference' can
# be used. 'emptySetLabel' defines the label, which is applied to the empty value. All
# parameters are optional.
#
# Missing data should be represented with a period character ('.').
# Units that are not coded by at least two judges are excluded from the analysis.
#
# This macro is version 1.0, updated on January 21, 2020
#
# Written by Max Wilckens
# Institute of Management & Organization
# Leuphana University Lüneburg
#
# This program is free software: you can redistribute it and/or modify it under the terms of
# the GNU General Public License as published by the Free Software Foundation, either version
# 3 of the License, or (at your option) any later version.
# This program is distributed in the hope that it will be useful,
# but WITHOUT ANY WARRANTY; without even the implied warranty of
# MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
# GNU General Public License for more details.
# You should have received a copy of the GNU General Public License
# along with this program.  If not, see <http://www.gnu.org/licenses/>.


KALPHA_multi <- function(data=NULL, standard=NULL, p_value=0.05, alpha_min=0.7,
         include_empty_set=TRUE, detail=1, sep=",", diff_fct="Krippendorff_default",
         emptySetLabel="none"){
  library(dplyr)
  library(tidyr)
  library(reshape2)
  library(abind)

  if(is.null(data))
    data <- read.table(file.choose(), header=TRUE, sep=",", na.strings=c(".","NA"),
         stringsAsFactors = FALSE)

  if(length(dim(data))==2){
    # data is given in a n x r matrix with n units and r raters. Each cell contains a comma
         seperated string of values.

    units  <- nrow(data)
    raters <- colnames(data)

    # Treat empty set as seperate value
    data[!is.na(data) & data == ""] <- emptySetLabel

    # split strings
    temp2 <- lapply(data,strsplit,split=sep)

    # values
    values <- sort(unique(unlist(temp2)))

    temp3 <- lapply(temp2,function(l1){
      lapply(l1, function(l2){
        if(length(l2) == 0)
          rep(FALSE,length(values))
        else if(is.na(l2[[1]]))
```

```
        rep(NA,length(values))
      else
        values %in% l2
    })
  })

  data_array <- array(unlist(temp3),dim=list(length(values),nrow(data),ncol(data)),
       dimnames=list(values,1:nrow(data),raters))
  data_array <- aperm(data_array,c(2,3,1))

} else if(length(dim(data))==3){
  # data is given in a n x r x c array with n units, r raters, and c categories.

  temp <- colSums(aperm(data,c(3,1,2)),na.rm=TRUE)
  if((any(temp2 <- temp == 0) | any(temp == dim(data)[3])) & include_empty_set){
    data_array <- abind(data, temp2, along=3)
    dimnames(data_array)[[3]][dim(data_array)[3]] <- emptySetLabel
  } else {
    data_array <- data
  }
  data_array <- data_array[,,apply(!is.na(data_array),3,sum)>0]

  raters <- dimnames(data_array)[[2]]
  values <- dimnames(data_array)[[3]]

} else {
  stop("ERROR: data has invalid dimensions.","\n")
}

sets <- t(unique(apply(data_array,c(3),rbind)))
sets <- sets[,!apply(sets,2,anyNA)]
colnames(sets) <- apply(sets,2,function(x){
  paste(as.integer(x),collapse="")
})

sets_alpha <- sort(apply(sets,2,function(x){
  paste(values[x==TRUE],collapse=",")
}))
sets <- sets[, names(sets_alpha)]

data_array_sets <- factor(apply(data_array,c(1,2),function(x){
  ifelse(anyNA(x),NA,paste(as.integer(x),collapse=""))
}), levels=c(names(sets_alpha),NA))
dim(data_array_sets)   <- dim(data_array)[c(1,2)]
dimnames(data_array_sets) <- list(1:nrow(data),raters)
sets_keys <- levels(data_array_sets)[!is.na(levels(data_array_sets))]
n_sets <- length(sets_keys)

raters_by_unit <- rowSums(!is.na(data_array_sets))

# display statistics
if(detail>=1){
  cat("Number of raters:          ",dim(data_array)[2],"\n")
  cat("Number of units:           ",dim(data_array)[1],"\n")
  cat("Number of unique values:   ",dim(data_array)[3],"\n")
  cat("Number of missing ratings: ",sum(is.na(data_array_sets)),"\n")
}

# create sets coincidence matrix
temp <- 1 / (raters_by_unit-1)
excluded <- sum(temp == Inf)
temp[temp == Inf] <- 0
rater_combn <- if(is.null(standard))  as.data.frame(combn(1:length(raters),2))   else
       as.data.frame(rbind(rep(standard,length(raters)-1),(1:length(raters))[-standard]))
temp <- lapply(rater_combn,function(l){
  data.frame(
    "Rating1" = as.numeric(data_array_sets[,l[1]]),
    "Rating2" = as.numeric(data_array_sets[,l[2]]),
    "value"   = temp
  )
})
temp <- bind_rows(temp)
temp <- acast(temp[!is.na(temp$Rating1) & !is.na(temp$Rating2),], Rating1~Rating2, sum)
coincidence <- matrix(0,n_sets,n_sets)
coincidence[match(rownames(temp),1:n_sets),match(colnames(temp),1:n_sets)] <- temp
coincidence <- coincidence + t(coincidence)

if(detail>=1){
```

```
  cat("Number of unique sets:     ",n_sets,"\n\n")
  cat("Number of excluded units due to missing ratings: ",excluded,"\n\n")
}
if(detail>=2){
  cat("Coincidence matrix by sets:\n")
  dimnames(coincidence) <- list(sets_alpha,sets_alpha)
  print(coincidence[order(colSums(sets)),order(colSums(sets))],digits=2)
  cat("\n")
}

n <- rowSums(coincidence)
n_values <- length(values)
length_sets <- setNames(colSums(sets),NULL)

sets_molten <- cbind(
  1:(n_sets*n_values),          # setvalue
  rep(1:n_sets,each=n_values),  # set
  rep(1:n_values,times=n_sets), # value
  as.vector(sets)               # value present in set?
)
# matrix of all set/value combinations
master <- matrix(NA,nrow=(n_sets*n_values)^2,ncol=13)
master[,1:4] <- sets_molten[rep(1:nrow(sets_molten),times=nrow(sets_molten)),]
master[,5:8] <- sets_molten[rep(1:nrow(sets_molten), each=nrow(sets_molten)),]
master <- master[(master[,1] <= master[,5]),]

# Columns
# 1: setvalue id a
# 2: set id a
# 3: value id a
# 4: value member of set a
# 5: setvalue id b
# 6: set id b
# 7: value id b
# 8: value member of set b
if(diff_fct=="Krippendorff_default_difference"){
  filter <- master[,4] & master[,8]

  # 9: value_difference
  master[,9]  <- ifelse(filter, !sets[cbind(master[,3],master[,6])] |
       !sets[cbind(master[,7],master[,2])], 0)
  # 10: value_coincidence
  master[,10] <- ifelse(filter, coincidence[cbind(master[,2], master[,6])], 0)
  # 11: value_products
  master[,11] <- ifelse(filter, n[master[,2]] * (n[master[,6]] - (master[,1]==master[,5])),
       0)

} else if(diff_fct=="Krippendorff_set_theoretic_difference"){
  filter <- master[,4] & master[,8]
  set_length_2 <- length_sets[master[,2]]
  set_length_6 <- length_sets[master[,6]]

  # 9: value_difference
  master[,9]  <- ifelse(
    filter,
    (!sets[cbind(master[,3],master[,6])])/set_length_6 +
       (!sets[cbind(master[,7],master[,2])])/set_length_2
    ,0)
  # 10: value_coincidence
  master[,10] <- ifelse(filter, coincidence[cbind(master[,2], master[,6])], 0)
  master[,12] <- ifelse(filter, master[,10]/set_length_2 + master[,10]/set_length_6, 0)
  # 11: value_products
  master[,11] <- ifelse(filter, n[master[,2]] * (n[master[,6]] - (master[,2]==master[,6])),
       0)
  master[,13] <- ifelse(filter, master[,11]/set_length_2 + master[,11]/set_length_6, 0)

} else
  stop("ERROR: Difference function not valid!")

value_difference <- matrix(0,n_sets*n_values,n_sets*n_values)
value_difference[upper.tri(value_difference,diag=TRUE)] <- master[,9]
value_difference <- value_difference + t(value_difference)
value_coincidence <- matrix(0,n_sets*n_values,n_sets*n_values)
value_coincidence[upper.tri(value_coincidence,diag=TRUE)] <- master[,10]
value_coincidence <- value_coincidence + t(value_coincidence)
diag(value_coincidence) <- diag(value_coincidence)/2

value_products <- matrix(0,n_sets*n_values,n_sets*n_values)
```

```
value_products[upper.tri(value_products,diag=TRUE)] <- master[,11]
value_products <- value_products + t(value_products)
diag(value_products) <- diag(value_products)/2

# difference matrix
if(detail>=2){
  group <- rep(1:ncol(sets),each=n_values)
  difference <- rowsum(t(rowsum(value_difference,group)),group)
  dimnames(difference) <- list(sets_alpha,sets_alpha)

  cat("Difference matrix by sets:\n")
  print(difference[order(colSums(sets)),order(colSums(sets))],digits=2)
  cat("\n")
}
if(diff_fct=="Krippendorff_set_theoretic_difference "){
  value_coincidence_dom <- matrix(0,n_sets*n_values,n_sets*n_values)
  value_coincidence_dom[upper.tri(value_coincidence_dom,diag=TRUE)] <- master[,12]
  value_coincidence_dom <- value_coincidence_dom + t(value_coincidence_dom)
  diag(value_coincidence_dom) <- diag(value_coincidence_dom)/2

  value_products_dom <- matrix(0,n_sets*n_values,n_sets*n_values)
  value_products_dom[upper.tri(value_products_dom,diag=TRUE)] <- master[,13]
  value_products_dom <- value_products_dom + t(value_products_dom)
  diag(value_products_dom) <- diag(value_products_dom)/2

  value_observed_nom <- tapply(rowSums(value_difference * value_coincidence),
       rep(1:n_values,ncol(sets)),sum)
  value_observed_dom <- tapply(rowSums(value_coincidence_dom),
       rep(1:n_values,ncol(sets)),sum)

  value_expected_nom <- tapply(rowSums(value_difference * value_products),
       rep(1:n_values,ncol(sets)),sum)
  value_expected_dom <- tapply(rowSums(value_products_dom), rep(1:n_values,ncol(sets)),sum)
} else {
  value_observed_nom <- tapply(rowSums(value_difference * value_coincidence),
       rep(1:n_values,ncol(sets)),sum)
  value_observed_dom <- tapply(rowSums(value_coincidence), rep(1:n_values,ncol(sets)),sum)
  value_expected_nom <- tapply(rowSums(value_difference * value_products),
       rep(1:n_values,ncol(sets)),sum)
  value_expected_dom <- tapply(rowSums(value_products), rep(1:n_values,ncol(sets)),sum)
}

value_Do <- value_observed_nom / value_observed_dom
value_De <- value_expected_nom / value_expected_dom

# observed disagreement
Do <- sum(value_observed_nom) / sum(value_observed_dom)

# expected disagreement
De <- sum(value_expected_nom) / sum(value_expected_dom)

kalpha <- 1 - Do / De
kalpha_value <- setNames(1 - (value_Do / value_De),values)

  if(detail>=1){
  cat("### Overall C-Alpha Results ###\n\n")
  cat("C-Alpha =",kalpha,"\n\n")

  if(detail>=2){
    cat("Do =",Do,"\n")
    cat("De =",De,"\n\n")
    cat("Do_nom =",sum(value_observed_nom),"\n\n")
    cat("Do_dom =",sum(value_observed_dom),"\n\n")
    cat("De_nom =",sum(value_expected_nom),"\n\n")
    cat("De_dom =",sum(value_expected_dom),"\n\n")
  }
}

# Confidence intervals & Type 1 error
N <- sum(value_observed_dom)
D <- sum(value_observed_nom)

if(D <= 5 | (N-D) <= 5){
  if(detail>=1){
    cat("Approximations for Alpha not possible due to small sample size\n\n")
  }
} else {
  # Confidence intervalls
```

```
    sig <- sqrt(D*(N-D)/N)
    alpha_confidence <- c(1-(D+qnorm(1-p_value/2)*sig)/(N*De), 1-(D-qnorm(1-
        p_value/2)*sig)/(N*De))

    if(detail>=1){
      cat("Approximations to the confidence intervals for C-Alpha-nominal:\n")
      cat((1-p_value)*100, "% confidence intervall for alpha: [",alpha_confidence[1],",
         ",alpha_confidence[2],"]\n\n",sep = "")
    }
    #Type 1 error for not reaching alpha_min
    D_min <- (1-alpha_min) * N * De
    z_min <- (D_min - D) / sig
    q <- 1 - pnorm(z_min)
    if(detail>=1){
      cat("Approximation to the probability q of the Type I error, that the observed C-Alpha
         does not fall below a chosen Alpha min = ",alpha_min,"\n")
      cat("q =",round(q,4),"\n\n")
    }
  }
  if(detail>=1){
    cat("Absolute set agreement: ",round((1-Do)*100,2), "%\n\n\n")

    cat("### Indivudial Value Alpha Results ###\n\n")
  }
  for(z in seq_along(kalpha_value)){
    if(detail>=1){
      cat("~~ Results for individual value \"",values[z],"\" ~~\n\n",sep = "")
      cat("Alpha for individual value \"",values[z],"\" = ",kalpha_value[z],"\n\n",sep = "")
    }
    N <- value_observed_dom[z]
    D <- value_observed_nom[z]

    if(D <= 5 | (N-D) <= 5){
      if(detail>=1){
        cat("Approximations for \"",values[z],"\" not possible due to small sample
         size\n\n",sep = "")
      }
    } else {
      # Confidence intervalls
      sig <- sqrt(D*(N-D)/N)
      alpha_confidence <- c(1-(D+qnorm(1-p_value/2)*sig)/(N*value_De[z]), 1-(D-qnorm(1-
         p_value/2)*sig)/(N*value_De[z]))

      if(detail>=1){
        cat("Approximations to the confidence intervals for the individual value
         \"",values[z],"\"\n",sep = "")
        cat((1-p_value)*100, "% confidence intervall for ",values[z],"'s Alpha:
         [",alpha_confidence[1],", ",alpha_confidence[2],"]\n\n",sep = "")
      }
      #Type 1 error for not reaching alpha_min
      D_min <- (1-alpha_min) * N * value_De[z]
      z_min <- (D_min - D) / sig
      if(kalpha_value[z] > alpha_min)
        q <- 1 - pnorm(z_min)
      else
        q <- pnorm(z_min)
      if(detail>=1){
        cat("Approximation to the probability q of the Type I error, that the observed Alpha
         for \"",values[z],"\" does not fall below a chosen Alpha min = ",alpha_min,"\n",sep =
         "")
        cat("q =",round(q,4),"\n\n")
      }
    }
    if(detail>=2){
      cat("Absolute agreement for \"",values[z],"\": ",round((1-
         value_Do[z])*100,2),"%\n\n",sep = "")
    }
    if(detail>=1){
      cat("\n\n")
    }
  }
  results <- list()
  results[['krip.alpha.multi']] <- kalpha
  results[['krip.alpha.multi.values']] <- kalpha_value
  results[['krip.alpha.multi.coincidence']] <- coincidence
  results[['krip.alpha.multi.sets']] <- n_sets
  return(results)
}
```

Appendix D. Final LLWI Scales (Copyrighted by the authors)

Table D.1

Final LLWI scales (Copyrighted by the authors. Approach the authors for permission to use)

Code	English Items (translated from the validated German items using back-translation procedure by Brislin, 1970)
OC Organizational climate	
The organizational culture dimension includes the set standards and actions of an employer shaped by the mission and values of the organization. An organizational culture that fosters good management of employees just before and in retirement age especially promotes equal opportunities and a positive image for all age groups. Subscales are:	
OC1 *Equality of opportunity*: Initial conditions should be the same for every employee regardless of age. Further, no discrimination or stigmatization due to age should occur. Each employee therefore has the same opportunities, e.g., participation in training and professional qualification or in the need of downsizing.	
OC1Q1	In our organization regardless of age, all employees have the same opportunities.
OC1Q2	In our organization regardless of age, all employees have the same opportunities for further training.
OC1Q3	In our organization regardless of age, all employees have the same opportunities to develop their career.
OC2 *Positive image of age*: Prevailing beliefs and attitudes regarding older employees are shaped by a positive attitude within the organization. Aging should be understood as an individual change process of competencies, motivation, values, and behavior. Opportunities should be recognized, valued and realized. For example, by identifying and assigning tasks which correspond to the specific competencies of older individuals.	
OC2Q1	In our organization there is a positive attitude towards older employees.
OC2Q2	In our organization older employees are perceived as being able to adapt well to changes.
OC2Q4	In our organization older employees are perceived as competent.
OC3 *Open and target group-oriented communication*: The organization is characterized by a differentiated image of age that is communicated trough external and internal representation of the organization. This explicitly includes open and transparent exchange between employees and their managers regarding retirement and/or continued opportunities for work. Positive images representing all age groups within the employee magazine, on the intranet or website are another example.	
OC3Q1	In our organization the possibilities of working for older employees are openly communicated.
OC3Q2	In our organization "aging" is talked about openly.
OC3Q3	In our organization employees can openly talk about age-related challenges and issues (e.g., performance limitations, speed in using digital tools, changes in short-term memory).
OC3Q4	In our organization there is a great deal of understanding for the challenges of aging.
LE Leadership	
The leadership dimension includes the responsibility of organizational executives to harness the potential of employees at all ages and particularly just before and in retirement age. This is achieved through the consideration of each individual employee's strengths and by showing appreciation for their talents and contributions. The framework comprises two facets, however the empirical data suggest unidimensionality:	
Appreciation: Managers of an organization should have an appreciative attitude towards their employees of all ages, manifested through a consistent demonstration of respect and kindness. Managers should reward the experience and achievements of their employees by offering higher levels of job autonomy and responsibility. Celebrating milestones and farewells are another way to convey gratitude, particularly when an employee is going into retirement.	
Responsiveness to individuality: Managers of an organization should be sensitive to individual needs and events that occur at different life stages. They should also take into account each individual's personality and performance capability. Managers are responsible for recognizing and harnessing individual potential regardless of age and for creating performance-enhancing conditions. Among other factors, this includes the consideration of employees' wishes and suggestions regarding the design of their work space as well as the consideration of individual life circumstances, such as the need to care for family.	
LE1Q2	Managers of our organization …show appreciation both for current work results as well as for the overall performance of their employees.
LE1Q4	Managers of our organization give their employees freedom in designing their work.
LE2Q1	Managers of our organization invest time in their employees.
LE2Q2	Managers of our organization address the personal needs and living conditions of their employees.
LE2Q3	Managers of our organization sincerely support their employees in their professional and personal development.
LE2Q4	Managers of our organization are interested in the well-being of their employees.
WD Work design	
The work design dimension includes the adaptation of work location, times and physical space to fit the individual needs and abilities of employees, relieve strain and increase job satisfaction and efficiency. Subscales are:	
WD1 *Flexible work time arrangements*: The organization should allow employees to change their work time depending on individual needs. Specific solutions will depend on the nature of an employee's work. Options for flexibility could include a long or short term switch to part time, offering flextime, job sharing, the possibility of swapping shifts, and unpaid leaves.	
WD1Q1	Employees of our organization can adjust the beginning and the end of their daily working hours to their individual needs.

Code	English Items (translated from the validated German items using back-translation procedure by Brislin, 1970)
WD1Q2	Employees of our organization can reduce or increase the number of hours specified in their work contract according to their individual needs.
WD1Q3	Employees of our organization can adapt the timing and the length of their breaks to their individual needs.
WD1Q4	Employees of our organization have enough flexibility in their working time organization to appropriately address unforeseen events in their private lives.
WD2 *Flexible workplaces*: When possible, employees should be able to choose their work location based on their individual needs and what is most efficient. Examples include the facilitation and technical support of home-office-solutions or the installation of silent work places within the office.	
WD2Q1	Employees of our organization have the opportunity to work from home.
WD2Q2	Employees of our organization have the opportunity to flexibly adapt where they work in the organization to their current needs (e.g., quiet workplaces, standing workstations, project workrooms).
WD2Q3	Employees of our organization can choose their place of work to ensure a good balance between their work and private life (work-life balance).
WD3 *Work according to capabilities*: Employees should have adequate jobs corresponding to their individual physical and mental performance capability and resilience. If not the case, this could be realized through a temporary or permanent change to another role that is less straining. Swapping jobs or reconsidering and adapting work flows should also be taken into consideration.	
WD3Q2	In our organization managers change the tasks of their employees in the foreseeable future (e.g., within half a year) if the tasks no longer correspond to the employee's ability to perform and to withstand stress.
WD3Q3	In our organization job rotation (regular change of responsibilities) is provided in case of monotonous routines or high physical strain at the workplace.
WD3Q4	In our organization when tasks are cognitively over- or undemanding (e.g., asking employees to remember many things, to concentrate, to make difficult decisions) the assignment is changed in the foreseeable future (e.g., within half a year).
WD4 *Ergonomic working conditions*: The work place should be designed according to ergonomic requirements and should also take into account the individual circumstances of the employee. For example occupational safety measures should be taken and supportive equipment and/or tools should be provided.	
WD4Q1	In our organization workplaces are designed according to ergonomic recommendations.
WD4Q2	In our organization proposals by employees for ergonomic improvements are taken up and implemented as far as possible.
WD4Q4	In our organization employees can adapt the lighting conditions at their workplace to their individual needs.
WD4Q5	In our organization employees use the most appropriate tools to reduce the physical strain of their work.
HM Health management	
The health management dimension includes all organizational activities that aim to maintain and promote employees' health and work ability. Health management should be characterized by a holistic approach addressing not only specific interventions but also health-promoting work design and leadership. Subscales are:	
HM1 *Availability of physical exercise and nutrition opportunities*: Initiatives to strengthen health and work ability should be offered, such as company sports activities, active breaks and nutritional guidance.	
HM1Q2	Employees of our organization receive incentives and opportunities to eat healthy food (e.g., by lower prices or a greater variety compared to the less healthy alternatives).
HM1Q3	Employees of our organization are encouraged to move as much as possible in the workplace (e.g., use the stairs, talk a walk during lunch break, sports during lunch break, use the bicycle to work).
HM1Q4	Employees of our organization receive incentives and opportunities to do sports outside work (e.g., company sports groups, cooperation with gyms).
HM2 *Workplace medical treatment*: Measures should be taken to help employees avoid medical conditions and assistance to aid in the recovery of sick employees should be offered. Examples include company doctors, on-site medical check-ups and physical therapy, along with wellness programs.	
HM2Q1	In our organization employees regularly receive medical check-ups (e.g., vaccinations, stress tests, eye examinations, blood pressure).
HM2Q2	In our organization there are special programs to reintegrate employees into work after a long illness (e.g., medical therapies, mental or physical health therapies).
HM2Q4	In our organization employees receive therapeutic help in the workplace or in the immediate vicinity if required (e.g., physiotherapy in case of great physical stress and strain).
HM3 *Health promotion*: Measures should be taken to disseminate knowledge about healthy behaviors to help employees make responsible and healthy decisions. This could be done by providing information on healthy living. Moreover, managers should act as role models for healthy behaviors and promote a healthy work environment. This includes taking part in physical exercise, nutrition opportunities and related programs themselves, as well as encouraging a sustainable work-life balance.	
HM3Q2	In our organization employees are made aware of health-promoting behavior (e.g., through training, counseling, displays).
HM3Q3	In our organization managers and top management are committed to promoting a sustainable, healthy way of life and work for their employees.
HM3Q4	In our organization health aspects play an important role in organizational decisions (e.g., investment decisions or operational changes).

Code	English Items (translated from the validated German items using back-translation procedure by Brislin, 1970)
ID Individual development	
Employees should be supported in their professional and personal development during their entire work life. A special emphasis is put on the importance of lifelong learning through continued education and training. There should also be opportunities for career development through internal advancement and promotions. The framework comprises three facets, however the empirical data suggest unidimensionality:	
	Continuous development planning: Planning for each individual employee's future should be done on an ongoing basis at all ages and stages of the work life. This could be done through individual meetings between managers and employees and by providing professional workshops that allow for self-reflection on abilities, competencies, and goals.
	Appropriate solutions for training and development: The organization should provide further training and education aligned with the individual employee's professional, educational, and life experience as well as with organizational goals. Further, training content and methods should be targeted towards specific groups. Examples of appropriate training and development solutions are workshops, seminars and industry conferences, training for new technologies or equipment, cross-training, and internships for people of all ages. These training and development opportunities can be facilitated onsite or through reimbursement of tuition or fees.
	Enabling development steps and job changes: Modifications to an employee's current position, function or job should be made possible to reflect the specific competencies and development interests of an individual. For example, this could be achieved by increasing job responsibilities, inclusion into other projects, or a horizontal or vertical change of position, which could also mean an additional apprenticeship or a new job within a different department.
ID1Q1	In our organization development prospects and qualification requirements are identified for employees, regardless of age.
ID1Q2	In our organization managers have regular conversations with their employees, regardless of age, about their personal and professional objectives (e.g., annual meetings to discuss their developmental goals).
ID1Q5	In our organization employees, regardless of age, know about their potential for development.
ID2Q2	In our organization older employees are offered training to learn new competencies and develop their expertise.
ID2Q3	In our organization training methods are adapted to take into account the needs of older employees (e.g., more practical learning techniques instead of lecture formats).
ID3Q2	In our organization employees, regardless of age, are involved in projects according to their competencies and developmental interests.
ID3Q3	In our organization opportunities for career development into management or expert positions are possible for older employees.
ID3Q4	In our organization employees move to a different job or position if it better suits their specific skills and abilities.
KM Knowledge management	
The knowledge management dimension includes procedures for the transfer, exchange, and conservation of knowledge between different generations of employees. Subscales are:	
	KM1 *Institutionalized knowledge transfer*: Institutionalized structures that transfer knowledge from experienced employees to their successors should be in place. This can be achieved through mentoring and "buddy" programs or through a systematic knowledge transfer process before employees leave the organization for retirement.
KM1Q1	In our organization there are mentoring programs in which experienced employees support others with their knowledge.
KM1Q2	In our organization there are processes / procedures to systematically pass on the knowledge and experience of older employees to their younger colleagues before they leave the organization.
KM1Q3	In our organization there are IT systems that are also used by older employees for the documentation and dissemination of knowledge.
KM1Q4	In our organization there are regular opportunities for every employee to exchange experiences and knowledge (e.g., in regular meetings).
	KM2 *Inter-generative collaboration*: The organization should allow for mutual transfer of knowledge and experience between generations. This transfer goes in both directions, young to old, as well as old to young. Its structure is not necessarily determined by the organization. For example, collaboration can happen within intergenerational pairs or age-mixed teams.
KM2Q1	In our organization …older and younger employees are encouraged to share their knowledge and experience.
KM2Q2	In our organization managers support the exchange of knowledge between younger and older employees.
KM2Q3	In our organization employees pass on their knowledge to colleagues of other generations (younger or older).
TR Transition to retirement	
The transition into retirement dimension includes the necessary conversations, planning, and workplace solutions for any employee who is on the verge of retiring. Information and counseling should be provided to help the employee transition. Subscales are:	
	TR1 *Timely transition planning*: Managers should talk with employees about their personal plans for entering the retirement stage, including a succession plan. Potential transition scenarios should be actively discussed to find individual solutions, for example, through annual employee interviews.
TR1Q1	In our organization managers discuss early with their employees (e.g., from the age of 55) as to how to make the transition to retirement.
TR1Q2	In our organization managers take time to plan the transition to retirement for individual employees.

Code	English Items (translated from the validated German items using back-translation procedure by Brislin, 1970)
TR1Q3	In our organization succession planning for the employee who is retiring is begun long before the expected retirement date.

TR2 *Phased retirement and individualized transition solutions*: Generic solutions for the transition into retirement should be tailored according to employees' individual needs. Flexibility and imagination should be present when designing the employee's individual transition into retirement. Phased retirement through a gradual reduction of working time should be offered companywide. Phased retirement can take place over a shorter or longer period of time, depending on needs.

TR2Q1	In our organization employees have the option to reduce their weekly working hours during the last years before retirement (phased retirement).
TR2Q2	In our organization employees have the option to work full time (with 50 percent pay), followed by a period of non-working (also with 50 percent pay) over a period of 2-3 years each before retirement.
TR2Q4	In our organization employees can adjust their working hours before retirement (e.g., flextime or, if shift work, no night shifts).
TR2Q6	In our organization the transition to retirement is flexibly shaped according to employee needs.

TR3 *Counselling for retirement life preparation*: Organizations should support their employees in preparing mentally for the life change of retirement by providing advising and counseling. Employees should be motivated to actively design their retirement life prior to transition. For example, individual preparation can be fostered through a structured approach that reflects individual expectations and plans. There may also be opportunities to establish alternative activities beyond employment.

TR3Q3	Our organization offers counseling to employees who are about to retire so they can reflect upon their expectations and plans for retirement.
TR3Q4	Our organization encourages employees who are about to retire to develop alternative activities for a meaningful daily routine after retirement (e.g., family, volunteering, traveling).
TR3Q7	Our organization provides employees with information about retirement (e.g., articles, brochures, books, internet/intranet sites).

TR4 *Continuous inclusion and maintaining contact*: Tools should be in place to maintain contact with employees even after their retirement and to help them stay engaged as part of the organization. This could be facilitated through an active management of relationships by means of an alumni network, invitations to organizational events or by allowing for voluntary work.

TR4Q1	Our organization maintains active contact with retired employees (e.g., by an alumni network).
TR4Q2	Our organization informs retired employees about current developments in the organization (e.g., newsletter, alumni newsletter).
TR4Q4	Our organization allows retired employees to catch up with each other regularly (e.g., at meetings of an alumni network).
TR4Q7	Our organization is still in active contact with most of its former employees, even 5 years after their retirement.

CE Continued employment

The continued employment dimension includes the organizational design and employment options for employees at retirement age. This includes former employees of the organization as well as external employees looking for continued employment. Subscales are:

CE1 *Individualized employment options*: Employment options for individuals, who would otherwise be fully retired, should be offered systematically. To ensure employment options are meaningful for both the organization and the employee, integration of those employees into the organization should be strategically planned. For example, they might be brought in on a temporary basis at peak production times. Tasks and work time should be adaptable to the individual employee. This can be achieved through alternative contract forms such as consulting and mentoring work or flexible work time arrangements with generally fewer hours than a full-time position.

CE1Q1	In our organization employees may work beyond the conventional retirement age if they wish so.
CE1Q2	In our organization employment opportunities for people in retirement age are clearly defined and structured (e.g., by integration into strategic workforce planning).
CE1Q3	In our organization managers are well-informed about the possibilities of working beyond the conventional retirement age.
CE1Q4	In our organization working conditions (time and type of activity) for employees in retirement age are flexibly adapted to their wishes.

CE2 *(Re-) hiring of older employees*: Older individuals, particularly including already and almost retired employees should be specifically addressed by job marketing, hiring and re-employment processes. This is achieved through age-friendly communication of job offers and the use of alternative marketing paths to address external as well as internal individuals. This explicitly includes employees with long careers in other industries or companies.

CE2Q1	In our organization older applicants are hired as well.
CE2Q2	In our organization age-neutral language is used in recruitment (e.g., job advertisements).
CE2Q4	In our organization people of all ages apply for job vacancies.

CC Health & retirement coverage

Organizations should support their employees with retirement savings and insurance coverage, if not sufficiently provided by public systems. Requirements vary due to different regulations and social systems. The support may be a direct financial benefit or put into practice as individual planning and assistance. Subscales are:

Code	English Items (translated from the validated German items using back-translation procedure by Brislin, 1970)
CC1 R*etirement savings and pensions*: Employees should be offered options for retirement savings, if not sufficiently covered by public systems. Organizations may include pensions and retirement saving accounts into their full compensation packages, offer optional saving possibilities to be opened by the employees individually, and support their employees in timely planning and organization of their retirement savings.	
CC1Q1	Our organization thoroughly informs employees about the components of a retirement plan (e.g., federal or state retirement systems, retirement plans offered by employer, private savings and investments, continued employment during retirement).
CC1Q2	Our organization offers employees comprehensive opportunities to save money for their retirement.
CC1Q3	Our organization offers employees good personal advice on financial security in later life.
CC2 *Insurances and financial emergency support*: Organizations should offer health related insurance coverage, if not sufficiently covered by public systems. This includes (additional) health-, disability-, care- or life insurances, which particularly cover risks that increase with age. Additional financial support may be offered in case of family emergencies, as e.g., in a case of nursing care or child sickness.	
CC2Q1	Our organization keeps employees well-informed about meaningful private supplemental insurance covering age-related risks (e.g., supplements to health or long-term care insurance, occupational accident insurance).
CC2Q2	Our organization offers employees private supplemental insurance as part of the total remuneration package (e.g., additions to health or long-term care insurance, occupational disability).

Note. Items were rated on a 5-point Likert response format, except for *organizational climate* and *leadership*, for which a 7-point format was used. The Likert response format was anchored at "Does not at all apply in our organization (for no employee or to no extent)" (1) and "Does fully apply in our organization (for all employees to the fullest extent)" (5 or 7). For *leadership* and *Individual development* results did not support the hypothesized factorial model with subscales, so that a unidimensional scale was developed. Construct definitions initially published by Wilckens et al. (2020).

Appendix E. Final German LLWI Scales (Copyrighted by the authors)

Table E.1

Final LLWI scales in German (Copyrighted by the Authors. Approach the authors for permission to use)

Code	German Items
OC Organisationsklima	
OC1 Chancengleichheit	
OC1-1	In unserer Organisation haben unabhängig vom Alter alle Beschäftigten die gleichen Möglichkeiten.
OC1-2	In unserer Organisation haben unabhängig vom Alter alle Beschäftigten die gleichen Chancen auf Weiterbildung.
OC1-3	In unserer Organisation haben unabhängig vom Alter alle Beschäftigten die gleichen Chancen auf Entwicklung ihrer Karriere.
OC2 Positives Altersbild	
OC2-1	In unserer Organisation herrscht eine positive Einstellung gegenüber älteren Beschäftigten.
OC2-2	In unserer Organisation werden ältere Beschäftigte als fähig wahrgenommen, sich Veränderungen gut anzupassen.
OC2-3	In unserer Organisation werden ältere Beschäftigte als kompetent wahrgenommen.
OC3 Offene und zielgruppengerechte Kommunikation	
OC3-1	In unserer Organisation werden Möglichkeiten des Arbeitens im Alter offen kommuniziert.
OC3-2	In unserer Organisation wird über das "Altern" offen gesprochen.
OC3-3	In unserer Organisation können Beschäftigte altersbedingte Herausforderungen und Probleme offen ansprechen (z.B. Leistungseinschränkungen, Schnelligkeit in der Bedienung digitaler Tools, Merkfähigkeit des Kurzzeitgedächtnis).
OC3-4	In unserer Organisation gibt es viel Verständnis für die Herausforderungen des Alterns.
LE Führung	
LE-1	Führungskräfte unserer Organisation zeigen Anerkennung sowohl für aktuelle Arbeitsergebnisse als auch für die Gesamtleistung ihrer Mitarbeiter.
LE-2	Führungskräfte unserer Organisation gewähren Ihren Mitarbeitern Freiraum in der Gestaltung der Arbeit.
LE-3	Führungskräfte unserer Organisation nehmen sich Zeit für Ihre Mitarbeiter.
LE-4	Führungskräfte unserer Organisation gehen auf persönliche Bedürfnisse und Lebensumstände ihrer Mitarbeiter ein.
LE-5	Führungskräfte unserer Organisation unterstützen ihre Mitarbeiter aufrichtig darin, sich beruflich und persönlich weiter zu entwickeln.
LE-6	Führungskräfte unserer Organisation sind an dem Befinden ihrer Mitarbeiter interessiert.
WD Arbeitsgestaltung	
WD1 Flexible Arbeitszeiten	
WD1-1	Die Beschäftigten unserer Organisation können den Beginn und das Ende ihrer täglichen Arbeitszeit an ihre individuellen Bedürfnisse anpassen.
WD1-2	Die Beschäftigten unserer Organisation können die Anzahl ihrer vertraglich vereinbarten Arbeitsstunden entsprechend ihrer individuellen Bedürfnisse reduzieren oder erhöhen.
WD1-3	Die Beschäftigten unserer Organisation können die Lage und die Länge ihrer Pausen an ihre individuellen Bedürfnisse anpassen.
WD1-4	Die Beschäftigten unserer Organisation haben ausreichend Flexibilität in der Arbeitszeitgestaltung, um auf unvorhergesehene Ereignisse im Privatleben angemessen reagieren zu können.
WD2 Flexible Arbeitsorte	
WD2-1	Die Beschäftigten unserer Organisation haben die Möglichkeit von zu Hause aus zu arbeiten.
WD2-2	Die Beschäftigten unserer Organisation haben die Möglichkeit, ihren Arbeitsort im Betrieb flexibel an ihre aktuellen Bedürfnisse anzupassen (z. B. stille Arbeitsplätze, Steharbeitsplätze, Projektarbeitsräume).
WD2-3	Die Beschäftigten unserer Organisation können ihren Arbeitsort so wählen, dass die Arbeit mit ihrem Privatleben gut zu vereinbaren ist (Work-Life Balance).
WD3 Arbeit gemäß Leistungsfähigkeit	
WD3-1	In unserer Organisation verändern Führungskräfte die Tätigkeiten ihrer Beschäftigten in absehbarer Zeit (z.B innerhalb eines halben Jahres), sofern sie ihrer Leistungsfähigkeit und Belastungsfähigkeit nicht mehr entsprechen.
WD3-2	In unserer Organisation wird bei einseitigen oder hohen körperlichen Belastungen an Arbeitsplätzen auf eine entlastungsorientierte Rotation (regelmäßiger Arbeitsplatzwechsel) geachtet.
WD3-3	In unserer Organisation wird die Tätigkeit bei kognitiver Über- oder Unterforderung (sich viele Dinge merken, sich konzentrieren, schwierige Entscheidungen treffen müssen) in absehbarer Zeit verändert (z.B innerhalb eines halben Jahres).
WD4 Ergonomische Arbeitsplatzbedingungen	
WD4-1	In unserer Organisation werden Arbeitsplätze nach ergonomischen Empfehlungen gestaltet.
WD4-2	In unserer Organisation werden Vorschläge der Beschäftigten zu ergnomischen Verbesserungen aufgegriffen und möglichst umgesetzt.
WD4-3	In unserer Organisation können Beschäftigte die Lichtverhältnisse an ihrem Arbeitsplatz an ihre individuellen Bedürfnisse anpassen.
WD4-4	In unserer Organisation verwenden die Beschäftigten die am besten geeigneten Hilfsmittel, um körperliche Belastungen durch die Arbeit zu verringern.

Code	German Items
HM Gesundheitsmanagement	
HM1 Bewegungs- und Ernährungsangebote	
HM1-1	Die Beschäftigten unserer Organisation erhalten Anreize und Möglichkeiten, sich gesund zu ernähren (z.B. über Vergünstigungen oder ein größeres Angebot im Vergleich zu den weniger gesunden Alternativen).
HM1-2	Die Beschäftigten unserer Organisation werden dazu ermutigt, sich möglichst viel am Arbeitsplatz zu bewegen (z.B. Nutzung der Treppen, Spaziergänge in der Mittagspause, kurzes Sportangebot in der Mittagspause, Nutzung des Fahrrads auf dem Arbeitsweg).
HM1-3	Die Beschäftigten unserer Organisation erhalten Anreize und Möglichkeiten, sich außerhalb der Arbeit sportlich zu betätigen (z.B. Betriebssportgruppen, Kooperationen mit Vereinen oder Fitnessstudios)
HM2 Medizinische Angebote	
HM2-1	In unserer Organisation erhalten die Beschäftigten regelmäßig medizinische Vorsorgeuntersuchungen (z.B. Schutzimpfungen, Belastungstests, Sehtest, Blutdruck).
HM2-2	In unserer Organisation gibt es spezielle Programme, um Beschäftigte nach längerer Krankheit gezielt wieder in den Beruf einzugliedern (z.B. medizinische oder therapeutische Angebote).
HM2-3	In unserer Organisation erhalten die Beschäftigten am Arbeitsplatz oder in der direkten Umgebung bei Bedarf therapeutische Hilfe (z.B. Physiotherapie bei körperlicher Überbeanspruchung oder Fehlbelastung).
HM3 Gesundheitsförderung	
HM3-1	In unserer Organisation werden die Beschäftigten für gesundheitsförderliches Verhalten sensibilisiert (z.B. durch Schulungen, Beratungsangebote, Aushänge).
HM3-2	In unserer Organisation setzen sich Führungskräfte und die Geschäftsführung für eine nachhaltig gesunde Lebens- und Arbeitsweise ihrer Mitarbeiter ein.
HM3-3	In unserer Organisation spielen gesundheitliche Aspekte in betrieblichen Entscheidungen eine relevante Rolle (z.B. bei Investitionsentscheidungen oder operativen Veränderungen)
ID Persönliche Entwicklung	
ID-1	In unserer Organisation werden für Beschäftigte jeden Alters Entwicklungsperspektiven und Qualifizierungsbedarfe identifiziert.
ID-2	In unserer Organisation führen Führungskräfte mit ihren Mitarbeitern jeden Alters regelmäßig Gespräche hinsichtlich ihrer beruflichen und persönlichen Perspektiven (z.B. Jahresgespräche).
ID-3	In unserer Organisation wissen Beschäftigte jeden Alters, wie sie sich weiterentwickeln können.
ID-4	In unserer Organisation werden auch älteren Beschäftigten Trainings zum Erlernen neuer Kompetenzen und Expertise angeboten.
ID-5	In unserer Organisation werden Trainingsmethoden so angepasst, dass auch die Bedürfnisse älterer Beschäftigter berücksichtigt werden (z.B. mehr praktische Lerntechniken anstelle von Vorlesungsformaten).
ID-6	In unserer Organisation werden Beschäftigte jeden Alters entsprechend ihrer Kompetenzen und Entwicklungsinteressen in Projekte eingebunden.
ID-7	In unserer Organisation sind bis ins hohe Alter Aufstiegsmöglichkeiten in Führungsfunktionen oder Spezialistenfunktionen möglich.
ID-8	In unserer Organisation wechseln die Beschäftigten in eine andere Tätigkeit oder Position, wenn diese ihren spezifischen Kompetenzen und Fähigkeiten besser entspricht.
KM Wissensmanagement	
KM1 Institutionalisierter Wissenstransfer	
KM1-1	In unserer Organisation existieren Mentoring Programme, in denen erfahrene Beschäftigte andere mit ihrem Wissen unterstützen.
KM1-2	In unserer Organisation existieren Prozesse / Abläufe, um die Kenntnisse und Erfahrungen älterer Beschäftigter vor ihrem Ausscheiden aus der Organisation systematisch an jüngere Kollegen weiterzugeben.
KM1-3	In unserer Organisation existieren IT Systeme, die auch von älteren Beschäftigten für die Dokumentation und Verbreitung von Wissen genutzt werden.
KM1-4	In unserer Organisation gibt es für jeden Beschäftigten regelmäßig Gelegenheit, Erfahrungen und Kenntnisse auszutauschen (z.B. Erfahrungsaustauschrunden).
KM2 Inter-generative Zusammenarbeit	
KM2-1	In unserer Organisation sind ältere und jüngere Beschäftigte dazu angehalten, ihr Wissen und ihre Erfahrungen untereinander auszutauschen.
KM2-2	In unserer Organisation unterstützen die Führungskräfte den Wissensaustausch zwischen jüngeren und älteren Beschäftigten.
KM2-3	In unserer Organisation geben die Beschäftigten ihr Wissen an Kollegen anderer Generationen (jünger oder älter) weiter.

Code	German Items
TR Übergang in den Ruhestand	
TR1 Frühzeitige Übergangsplanung	
TR1-1	In unserer Organisation besprechen Führungskräfte mit ihren Beschäftigten frühzeitig (z.B. ab einem Alter von 55 Jahren), wie der Übergang in den Ruhestand gestaltet werden soll.
TR1-2	In unserer Organisation nehmen sich Führungskräfte Zeit, um den Übergang in den Ruhestand einzelner Beschäftigter zu planen.
TR1-3	In unserer Organisation ist die Nachfolge für den Beschäftigten, der in den Ruhestand geht, frühzeitig geplant.
TR2 Altersteilzeit und individuelle Übergangslösungen	
TR2-1	In unserer Organisation haben Beschäftigte die Möglichkeit, die letzten Jahre vor Eintritt in den Ruhestand ihre wöchentliche Arbeitszeit zu reduzieren (Teilzeit).
TR2-2	In unserer Organisation können Beschäftigte durch geblockte Altersteilzeit früher in den Ruhestand gehen.
TR2-3	In unserer Organisation können Beschäftigte vor Eintritt in den Ruhestand ihre Arbeitszeit individuell gestalten (z.B. Gleitzeit oder bei Schichtarbeit keine Nachtschichten).
TR2-4	In unserer Organisation wird der Übergang in den Ruhestand flexibel nach den Bedürfnissen der Beschäftigten gestaltet.
TR3 Beratung zur Vorbereitung des Lebens im Ruhestand	
TR3-1	Unsere Organisation bietet Beschäftigten, die kurz vor dem Eintritt in den Ruhestand stehen, Beratungsangebote, um ihre Erwartungen und Pläne für den Ruhestand zu reflektieren.
TR3-2	Unsere Organisation ermutigt Beschäftigte, die kurz vor dem Eintritt in den Ruhestand stehen, alternative Aktivitäten für eine sinnvolle Tagesgestaltung im Ruhestand aufzubauen (z.B. Ehrenamt, Reisen, Familie).
TR3-3	Unsere Organisation bietet Beschäftigten Informationen zum Thema Ruhestand (z.B. Artikel, Broschüren, Bücher, Internet- / Intranetseiten).
TR4 Fortlaufende Einbindung und Kontaktpflege	
TR4-1	Unsere Organisation hält zu ehemaligen Beschäftigten im Ruhestand aktiven Kontakt (z.B. in Form eines Alumni Netzwerkes).
TR4-2	Unsere Organisation informiert ehemalige Beschäftigte im Ruhestand über die aktuellen Entwicklungen im Unternehmen (z.B. Newsletter, Alumni-Newsletter).
TR4-3	Unsere Organisation ermöglicht es ehemaligen Beschäftigten im Ruhestand sich regelmäßig auszutauschen (z.B. bei Treffen eines Alumni-Netzwerkes).
TR4-4	Unsere Organisation steht mit dem Großteil der ehemaligen Beschäftigten auch 5 Jahre nach deren Eintritt in den Ruhestand noch in aktivem Kontakt.
CE Weiterbeschäftigung nach Renteneintritt	
CE1 Individualisierte Beschäftigungslösungen	
CE1-1	In unserer Organisation können (ehemalige) Beschäftigte über das Rentenalter hinaus tätig sein, sofern dies ihrem Wunsch entspricht.
CE1-2	In unserer Organisation sind Beschäftigungsmöglichkeiten für Personen im Rentenalter klar definiert und strukturiert (z.B. durch Integration in die strategische Personalplanung).
CE1-3	In unserer Organisation sind Führungskräfte über die Möglichkeiten einer Weiterbeschäftigung ihrer Mitarbeiter im Rentenalter gut informiert.
CE1-4	In unserer Organisation werden die Arbeitsbedingungen (Zeit und Art der Tätigkeit) für Beschäftigte im Rentenalter flexibel an deren Wünsche angepasst.
CE2 (Wieder-) Einstellung von älteren Beschäftigten	
CE2-1	In unserer Organisation werden auch ältere Bewerber eingestellt.
CE2-2	In unserer Organisation wird in der Personalwerbung (z.B. Stellenanzeigen) auf eine altersunabhängige Formulierung geachtet.
CE2-3	In unserer Organisation bewerben sich auf ausgeschriebene Stellen Erwerbstätige aller Altersgruppen.
RC Versicherungen und Vorsorge	
RC1 Altersvorsorge	
RC1-1	Unsere Organisation informiert die Beschäftigten gut über die Bestandteile einer finanziellen Versorgung im Alter (z.B. gesetzlich, betrieblich, privat, Weiterbeschäftigung im Rentenalter).
RC1-2	Unsere Organisation bietet seinen Beschäftigten umfassende Möglichkeiten, Gelder für das Rentenalter anzusparen.
RC1-3	Unsere Organisation bietet den Beschäftigten eine gute persönliche Beratung zu ihrer finanziellen Versorgung im Alter an.
RC2 (Kranken-) Versicherungen und finanzielle Unterstützung in Notlagen	
RC2-1	Unsere Organisation informiert die Beschäftigten gut über sinnvolle private Zusatzversicherungen, die altersbedingte Risiken abdecken (z.B. Ergänzungen zu Kranken- oder Pflegeversicherung, Berufsunfähigkeitsversicherung).
RC2-2	Unsere Organisation bietet den Beschäftigten private Zusatzversicherungen als Teil des Gesamt-Vergütungspaketes (z.B. Ergänzungen zu Kranken- oder Pflegeversicherung, Berufsunfähigkeit).

Note.

Appendix F. Most Relevant Existing Scales Considered During Item Generation

Table F.1

Most relevant existing scales considered during item generation in Study 1

Measures, scales, and inventories	Authors
General age-related human resources practices	
Human resources practices specifically tailored to employees 50 and over	Armstrong-Stassen & Lee (2009)
Human resources activities directed at recruiting and retaining older managerial and professional employees	Armstrong-Stassen & Templer (2006)
Human resources practices for the post- retirement employment experience of older workers	Armstrong-Stassen (2008)
Bundles of human resources practices for aging workers	Kooij, Jansen, Dikkers, & de Lange (2014)
Age-diversity human resources practices	Boehm, Kunze, & Bruch (2014)
Human resource practices that late-career workers find valuable	Taneva & Arnold (2018)
Age discrimination and climate	
Work-related age-based stereotypes (WAS)	Marcus, Fritzsche, Le, & Reeves (2016)
Psychological Age Climate Scale (PACS)	Noack (2009)
Opinions about characteristics of older workers	Henkens (2005)
Workplace Age Discrimination Scale (WADS)	Marchiondo, Gonzales, & Ran (2016)
Nordic Age Discrimination Scale (NADS)	Furunes & Mykletun (2010)
Leadership	
Respectful leadership	van Quaquebeke & Eckloff (2010)
Ethical leadership	Kalshoven, Den Hartog, & De Hoogh (2011)
Developmental leadership	Rafferty & Griffin (2006)
Content area specific human resources practices	
Work-time control measure	Valcour (2007)
People management scale	Knies, Leisink, & van de Schoot (2017)
European survey of enterprises on new and emerging risks (ESENER 2)	European Agency for Safety and Health at Work (2015)
Worksite Health Promotion	Della, DeJoy, Goetzel, Ozminkowski, & Wilson (2008)
Knowledge exchange and combination scale	Collins & Smith (2006)

Appendix G. Study 2 Item Statistics

Table G.1 *LLWI item statistics from Study 2*

Code	N	*M*	*SD*	Likert scale maximum	Skew	Kurtosis	Final scale
OC Organizational climate							
OC1Q1	595	5.26	1.51	7	-0.78	0.19	X
OC1Q2	589	5.36	1.58	7	-0.85	0.09	X
OC1Q3	585	4.94	1.60	7	-0.45	-0.49	X
OC1Q4	586	5.53	1.40	7	-0.80	0.11	
OC2Q1	592	5.48	1.35	7	-0.80	0.44	X
OC2Q2	585	5.29	1.41	7	-0.75	0.37	X
OC2Q3	598	5.93	1.13	7	-1.02	0.96	
OC2Q4	596	5.91	1.16	7	-1.06	0.97	X
OC3Q1	528	4.60	1.89	7	-0.39	-0.86	X
OC3Q2	572	4.32	1.83	7	-0.22	-0.94	X
OC3Q3	548	4.74	1.74	7	-0.44	-0.68	X
OC3Q4	557	4.71	1.71	7	-0.39	-0.62	X
LE Leadership							
LE1Q1	602	5.00	1.51	7	-0.57	-0.15	
LE1Q2	601	4.97	1.54	7	-0.63	-0.11	X
LE1Q3	601	4.89	1.43	7	-0.50	-0.09	
LE1Q4	602	5.03	1.50	7	-0.71	0.22	X
LE2Q1	599	4.74	1.50	7	-0.33	-0.40	X
LE2Q2	600	4.81	1.53	7	-0.50	-0.27	X
LE2Q3	597	4.77	1.49	7	-0.46	-0.23	X
LE2Q4	601	4.61	1.53	7	-0.41	-0.28	X
LE2Q5	501	3.22	1.86	7	0.37	-0.98	
WD Work design							
WD1Q1	602	3.06	1.36	5	-0.19	-1.14	X
WD1Q2	594	3.10	1.22	5	-0.19	-0.84	X
WD1Q3	597	2.91	1.35	5	-0.02	-1.19	X
WD1Q4	595	3.29	1.20	5	-0.33	-0.72	X
WD1Q5	501	2.83	1.33	5	0.08	-1.15	
WD2Q1	603	2.24	1.30	5	0.63	-0.80	X
WD2Q2	594	2.49	1.24	5	0.33	-0.94	X
WD2Q3	592	2.54	1.22	5	0.25	-0.93	X
WD3Q1	572	3.10	1.08	5	-0.21	-0.66	
WD3Q2	536	2.63	1.06	5	0.22	-0.53	X
WD3Q3	520	2.48	1.23	5	0.42	-0.83	X
WD3Q4	520	2.47	1.16	5	0.34	-0.80	X
WD4Q1	587	3.10	1.24	5	-0.15	-0.93	X
WD4Q2	576	3.09	1.18	5	-0.23	-0.77	X
WD4Q3	566	2.71	1.31	5	0.12	-1.17	
WD4Q4	588	2.87	1.28	5	-0.04	-1.03	X
WD4Q5	565	3.11	1.16	5	-0.26	-0.62	X
HM Health management							
HM1Q1	584	2.41	1.32	5	0.50	-0.98	
HM1Q2	587	2.30	1.34	5	0.64	-0.85	X
HM1Q3	588	2.55	1.29	5	0.31	-1.03	X
HM1Q4	591	2.63	1.44	5	0.26	-1.32	X
HM2Q1	591	2.96	1.50	5	-0.09	-1.44	X
HM2Q2	552	2.96	1.47	5	-0.05	-1.39	X
HM2Q3	573	2.48	1.16	5	0.32	-0.69	
HM2Q4	560	2.18	1.27	5	0.76	-0.61	X
HM3Q1	558	2.52	1.36	5	0.38	-1.11	
HM3Q2	572	2.51	1.23	5	0.29	-0.97	X
HM3Q3	573	2.42	1.26	5	0.45	-0.88	X
HM3Q4	534	2.39	1.21	5	0.41	-0.89	X

Code	N	*M*	*SD*	Likert scale maximum	Skew	Kurtosis	Final scale
ID Individual development							
ID1Q1	566	3.12	1.16	5	-0.24	-0.71	X
ID1Q2	592	3.35	1.33	5	-0.37	-1.03	X
ID1Q3	585	3.29	1.15	5	-0.39	-0.56	
ID1Q4	583	3.47	1.13	5	-0.47	-0.42	
ID1Q5	581	3.40	1.11	5	-0.44	-0.33	X
ID2Q1	570	3.44	1.18	5	-0.52	-0.48	
ID2Q2	575	3.48	1.15	5	-0.52	-0.41	X
ID2Q3	535	2.83	1.20	5	0.07	-0.88	X
ID2Q4	516	2.72	1.19	5	0.08	-0.92	
ID3Q1	554	3.02	1.08	5	-0.13	-0.48	
ID3Q2	585	3.46	1.11	5	-0.51	-0.27	X
ID3Q3	563	3.25	1.19	5	-0.31	-0.71	X
ID3Q4	575	3.05	1.11	5	-0.20	-0.60	X
KM Knowledge management							
KM1Q1	547	2.63	1.31	5	0.28	-1.02	X
KM1Q2	542	2.60	1.31	5	0.32	-1.03	X
KM1Q3	540	2.91	1.37	5	0.02	-1.23	X
KM1Q4	578	2.84	1.34	5	0.09	-1.14	X
KM2Q1	589	3.32	1.29	5	-0.40	-0.88	X
KM2Q2	576	2.99	1.28	5	-0.06	-1.00	X
KM2Q3	594	3.55	1.13	5	-0.62	-0.21	X
KM2Q4	594	4.15	0.94	5	-1.09	1.04	
TR Transition to retirement							
TR1Q1	484	2.41	1.34	5	0.47	-1.05	X
TR1Q2	488	2.45	1.31	5	0.39	-1.12	X
TR1Q3	535	2.69	1.31	5	0.21	-1.08	X
TR2Q1	518	3.22	1.34	5	-0.31	-1.00	X
TR2Q2	477	2.83	1.44	5	0.09	-1.28	X
TR2Q3	483	2.16	1.36	5	0.80	-0.73	
TR2Q4	505	2.57	1.35	5	0.28	-1.17	X
TR2Q6	498	2.55	1.28	5	0.26	-1.04	X
TR3Q3	459	2.27	1.32	5	0.64	-0.85	X
TR3Q4	477	1.99	1.23	5	1.07	0.06	X
TR3Q7	487	2.06	1.22	5	0.86	-0.41	X
TR3Q10	478	2.28	1.30	5	0.57	-0.95	
TR4Q1	510	2.35	1.34	5	0.55	-0.95	X
TR4Q2	494	2.12	1.31	5	0.87	-0.48	X
TR4Q4	502	2.13	1.29	5	0.80	-0.62	X
TR4Q7	495	2.22	1.31	5	0.68	-0.77	X
CE Continued Employment							
CE1Q1	500	3.15	1.34	5	-0.21	-1.06	X
CE1Q2	448	2.52	1.30	5	0.33	-1.07	X
CE1Q3	449	2.71	1.27	5	0.13	-1.06	X
CE1Q4	473	2.82	1.24	5	0.01	-0.93	X
CE2Q1	561	3.47	1.12	5	-0.48	-0.39	X
CE2Q2	495	3.68	1.25	5	-0.71	-0.47	X
CE2Q3	488	3.20	1.23	5	-0.29	-0.82	
CE2Q4	517	3.81	1.02	5	-0.68	-0.01	X
CC Health and retirement coverage							
CC1Q1	561	3.03	1.29	5	-0.12	-1.02	X
CC1Q2	555	3.07	1.32	5	-0.19	-1.06	X
CC1Q3	545	2.68	1.29	5	0.18	-1.06	X
CC1Q4	571	3.65	1.44	5	-0.75	-0.80	
CC2Q1	551	2.45	1.32	5	0.43	-1.02	X
CC2Q2	539	2.35	1.37	5	0.55	-1.04	X
CC2Q4	454	2.77	1.42	5	0.09	-1.32	

Appendix H. LLWI Indicator Correlation Tables

Table H.1

LLWI indicator statistics from Study 2

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1. OC	5.15	1.21	(.92)																												
2. OC1	5.20	1.44	.87	(.91)																											
3. OC2	5.56	1.18	.87	.69	(.88)																										
4. OC3	4.60	1.58	.87	.59	.64	(.90)																									
5. LE	4.82	1.35	.71	.60	.61	.67	(.95)																								
6. WD	2.76	0.84	.40	.28	.30	.48	.59	(.91)																							
7. WD1	3.09	1.06	.30	.21	.25	.32	.43	.79	(.84)																						
8. WD2	2.41	1.04	.21	.14	.16	.27	.36	.84	.62	(.78)																					
9. WD3	2.50	1.01	.39	.27	.27	.52	.53	.79	.42	.52	(.88)																				
10. WD4	3.03	1.02	.41	.31	.31	.46	.57	.80	.46	.53	.61	(.86)																			
11. HM	2.53	1.04	.26	.19	.12	.37	.40	.60	.36	.48	.53	.61	(.92)																		
12. HM1	2.48	1.16	.23	.17	.11	.32	.36	.56	.36	.48	.47	.54	.91	(.82)																	
13. HM2	2.70	1.18	.18	.12	.06	.28	.30	.48	.29	.37	.40	.50	.89	.69	(.78)																
14. HM3	2.42	1.12	.31	.23	.17	.41	.46	.61	.35	.44	.58	.64	.92	.77	.73	(.90)															
15. ID	3.24	0.91	.65	.57	.49	.63	.72	.64	.41	.42	.60	.63	.55	.46	.44	.58	(.90)														
16. KM	2.98	1.01	.49	.39	.38	.55	.60	.56	.31	.38	.58	.57	.56	.50	.44	.60	.72	(.89)													
17. KM1	2.72	1.06	.43	.32	.29	.51	.53	.57	.34	.41	.56	.57	.63	.57	.50	.65	.67	.92	(.81)												
18. KM2	3.28	1.11	.49	.41	.40	.51	.59	.46	.25	.30	.50	.48	.41	.35	.31	.45	.65	.93	.70	(.88)											
19. TR	2.37	1.02	.43	.30	.27	.57	.55	.66	.44	.47	.65	.62	.70	.61	.60	.71	.67	.65	.67	.53	(.94)										
20. TR1	2.49	1.19	.47	.34	.33	.58	.57	.57	.33	.39	.59	.55	.58	.48	.48	.62	.65	.65	.63	.58	.88	(.90)									
21. TR2	2.77	1.15	.38	.28	.22	.48	.45	.61	.48	.46	.52	.52	.59	.51	.52	.57	.56	.50	.54	.40	.84	.68	(.87)								
22. TR3	2.08	1.15	.29	.17	.14	.45	.39	.57	.35	.42	.56	.54	.64	.57	.54	.64	.51	.52	.59	.37	.88	.70	.66	(.92)							
23. TR4	2.18	1.21	.36	.25	.26	.46	.46	.52	.33	.35	.53	.50	.60	.55	.51	.60	.56	.56	.57	.46	.85	.65	.56	.68	(.94)						
24. CE	3.19	0.87	.55	.49	.44	.53	.50	.38	.23	.20	.45	.36	.35	.30	.28	.37	.57	.50	.44	.48	.55	.52	.49	.41	.46	(.84)					
25. CE1	2.77	1.11	.46	.35	.34	.52	.47	.39	.23	.21	.47	.34	.38	.35	.30	.39	.53	.50	.47	.46	.60	.52	.51	.49	.54	.88	(.87)				
26. CE2	3.65	0.93	.50	.48	.42	.38	.38	.22	.14	.08	.24	.24	.15	.12	.13	.18	.42	.30	.26	.32	.30	.32	.29	.17	.21	.83	.47	(.76)			
27. RC	2.63	1.13	.33	.25	.19	.42	.44	.57	.40	.44	.51	.54	.67	.60	.58	.65	.56	.53	.57	.42	.75	.65	.64	.68	.61	.41	.43	.21	(.90)		
28. RC1	2.91	1.16	.35	.29	.21	.42	.47	.54	.38	.40	.46	.51	.61	.54	.54	.59	.57	.52	.54	.42	.68	.62	.61	.59	.55	.40	.37	.27	.92	(.88)	
29. RC2	2.37	1.26	.24	.18	.12	.35	.34	.49	.33	.38	.46	.46	.61	.54	.54	.60	.45	.44	.49	.34	.70	.58	.58	.68	.55	.35	.39	.13	.93	.73	(.87)

Note. $N = 609$. *M* and *SD* are used to represent mean and standard deviation, respectively. Internal consistency coefficients, Cronbach's alphas are reported in the parentheses on the diagonal.

Table H.2

LLWI indicator statistics from Study 3

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1. OC	5.22	1.24	(.93)																												
2. OC1	5.27	1.45	.86	(.90)																											
3. OC2	5.57	1.22	.89	.67	(.89)																										
4. OC3	4.81	1.55	.89	.59	.72	(.91)																									
5. LE	4.74	1.44	.75	.62	.65	.71	(.95)																								
6. WD	2.68	0.82	.45	.35	.34	.50	.59	(.90)																							
7. WD1	3.00	1.11	.32	.26	.22	.37	.43	.80	(.84)																						
8. WD2	2.33	1.04	.19	.17	.12*	.23	.36	.82	.66	(.77)																					
9. WD3	2.40	0.99	.41	.31	.33	.45	.53	.70	.29	.39	(.84)																				
10. WD4	2.95	1.04	.45	.32	.39	.48	.55	.79	.45	.47	.53	(.86)																			
11. HM	2.57	1.02	.30	.22	.23	.35	.45	.59	.36	.43	.48	.60	(.91)																		
12. HM1	2.41	1.19	.23	.15	.19	.29	.37	.53	.33	.41	.42	.52	.90	(.81)																	
13. HM2	2.84	1.16	.23	.19	.18	.27	.31	.46	.27	.33	.36	.48	.86	.62	(.77)																
14. HM3	2.45	1.10	.34	.26	.27	.39	.50	.58	.35	.41	.50	.58	.90	.74	.65	(.87)															
15. ID	3.20	0.99	.69	.61	.59	.63	.76	.67	.44	.42	.58	.63	.55	.44	.44	.57	(.92)														
16. KM	3.13	1.01	.51	.38	.45	.49	.63	.55	.29	.35	.53	.58	.57	.47	.43	.59	.81	(.90)													
17. KM1	2.85	1.10	.44	.34	.39	.44	.57	.55	.29	.38	.53	.54	.57	.47	.45	.57	.79	.94	(.83)												
18. KM2	3.44	1.05	.53	.42	.49	.51	.64	.49	.25	.27	.48	.53	.45	.38	.32	.51	.71	.93	.76	(.87)											
19. TR	2.26	0.98	.46	.33	.39	.49	.59	.63	.39	.45	.64	.56	.62	.53	.49	.66	.66	.65	.64	.59	(.94)										
20. TR1	2.33	1.15	.44	.32	.38	.48	.56	.56	.29	.37	.63	.50	.49	.42	.37	.53	.63	.65	.64	.59	.87	(.89)									
21. TR2	2.65	1.15	.38	.28	.30	.41	.48	.59	.48	.40	.46	.49	.53	.42	.45	.53	.57	.51	.50	.47	.83	.65	(.85)								
22. TR3	1.90	1.09	.36	.27	.30	.38	.44	.54	.31	.44	.55	.44	.54	.47	.41	.60	.51	.51	.51	.44	.86	.67	.59	(.89)							
23. TR4	2.16	1.21	.35	.25	.31	.38	.47	.49	.28	.37	.53	.39	.55	.50	.40	.57	.50	.52	.51	.47	.85	.63	.56	.68	(.93)						
24. CE	3.14	0.97	.46	.35	.40	.46	.47	.42	.28	.25	.46	.34	.29	.28	.20	.31	.49	.50	.46	.48	.56	.57	.39	.45	.47	(.86)					
25. CE1	2.69	1.23	.43	.30	.37	.46	.46	.35	.19	.21	.50	.26	.27	.24	.18	.32	.45	.51	.46	.48	.59	.55	.38	.51	.49	.90	(.89)				
26. CE2	3.64	0.97	.38	.33	.35	.34	.37	.38	.29	.24	.27	.35	.26	.25	.19	.26	.42	.37	.34	.35	.36	.38	.28	.22	.28	.85	.53	(.72)			
27. RC	2.61	1.20	.28	.19	.22	.34	.44	.54	.33	.41	.51	.47	.65	.57	.56	.62	.50	.53	.54	.47	.66	.59	.51	.60	.59	.39	.39	.27	(.91)		
28. RC1	2.88	1.24	.30	.22	.24	.33	.43	.50	.32	.35	.46	.48	.64	.54	.56	.61	.51	.55	.55	.48	.64	.55	.52	.53	.56	.37	.34	.30	.93	(.89)	
29. RC2	2.41	1.31	.23	.14*	.17	.30	.37	.47	.27	.38	.47	.38	.57	.51	.48	.53	.41	.43	.44	.38	.59	.54	.43	.57	.53	.37	.38	.20	.94	.75	(.87)

Note. $N = 349$. *M* and *SD* are used to represent mean and standard deviation, respectively. Internal consistency coefficients, Cronbach's alphas are reported in the parentheses on the diagonal.

Appendix I. LLWI Reliability

Table I.1

Cronbach's alpha confidence intervals

	Study 2		Study 3	
Scale	Cronbach's α	95% *CI* (2000 bootstrap runs)	Cronbach's α	95% *CI* (2000 bootstrap runs)
1. OC	.92	[.91 ; .94]	.93	[.92 ; .94]
2. OC1	.91	[.89 ; .92]	.90	[.87 ; .92]
3. OC2	.88	[.86 ; .91]	.89	[.87 ; .91]
4. OC3	.90	[.88 ; .92]	.91	[.88 ; .93]
5. LE	.95	[.94 ; .96]	.95	[.94 ; .96]
6. WD	.91	[.90 ; .92]	.90	[.88 ; .91]
7. WD1	.84	[.82 ; .86]	.84	[.80 ; .87]
8. WD2	.78	[.75 ; .82]	.77	[.72 ; .82]
9. WD3	.88	[.85 ; .90]	.84	[.80 ; .88]
10. WD4	.86	[.84 ; .88]	.86	[.84 ; .89]
11. HM	.92	[.91 ; .93]	.91	[.89 ; .92]
12. HM1	.82	[.78 ; .85]	.81	[.76 ; .85]
13. HM2	.78	[.74 ; .81]	.77	[.72 ; .81]
14. HM3	.90	[.88 ; .92]	.87	[.84 ; .90]
15. ID	.90	[.89 ; .92]	.92	[.90 ; .93]
16. KM	.89	[.87 ; .90]	.90	[.88 ; .92]
17. KM1	.81	[.77 ; .83]	.83	[.79 ; .86]
18. KM2	.88	[.86 ; .90]	.87	[.84 ; .89]
19. TR	.94	[.94 ; .95]	.94	[.92 ; .95]
20. TR1	.90	[.87 ; .92]	.89	[.85 ; .91]
21. TR2	.87	[.85 ; .89]	.85	[.81 ; .87]
22. TR3	.92	[.90 ; .94]	.89	[.84 ; .92]
23. TR4	.94	[.93 ; .95]	.93	[.91 ; .95]
24. CE	.84	[.83 ; .87]	.86	[.83 ; .89]
25. CE1	.87	[.85 ; .89]	.89	[.86 ; .92]
26. CE2	.76	[.71 ; .80]	.72	[.64 ; .78]
27. RC	.90	[.89 ; .92]	.91	[.90 ; .93]
28. RC1	.88	[.86 ; .90]	.89	[.87 ; .92]
29. RC2	.87	[.83 ; .90]	.87	[.82 ; .91]