Master thesis in Communication and Information sciences

Burnout complaints in the Dutch working population

Inductively unraveling relative predictor importance using Johnson's relative weights method, Breiman's random regression forests and Kruskal's LMG method

Author: Indy Wijngaards

Internal supervisor: dr. Elske van der Vaart
External supervisor: Jeldrik Bakker MSc

Acknowledgement

For many students, their thesis is the final product of their academic careers; happily entering in the private or public sector to finally apply the many skills they have learned in the Master's program Data Science: Business and Governance. I, however, am one of the exceptions. I will start as a PhD candidate; during this 4-year period I will link happiness economics, human resource management (HRM) and data science to study the wellbeing of employees in organizations. In writing this master's thesis, I have experienced what huge opportunities large datasets and data mining techniques can provide in gaining insights in the field of occupational wellbeing. An understanding that presumably will prove itself useful in my academic and professional future.

First and foremost, I would like to express my gratitude for dr. Elske van der Vaart for her support and guidance. Besides the careful thesis supervision, I am also grateful for her valuable lessons in the course 'Statistical programming in R'. This elementary course has allowed me use R in a confident manner during the whole thesis project. Many thanks!

Secondly, thanks to dr. Marie Postma for reviewing my thesis in the end, providing valuable input to improve the thesis and enabling me to graduate in August 2017. In addition, I would like to express my gratitude for the insights prof. dr. Ulrike Grömping provided with respect to some methods of variable importance.

Thirdly, I would like to thank the Dutch Central Bureau of Statistics (CBS), and, in particular, Jeldrik Bakker MSc. During the kick-off of the master thesis projects, I soon noticed that no topics related to HRM and occupational health. When I communicated my concerns, Tilburg University and the CBS allowed me to propose a research strategy that aligned my own preferences from scratch. Now, look where we are now. The data used in this study was highly confidential and many official procedures had to be started. Jeldrik helped me for a great deal by pressuring internal CBS stakeholders, giving me advice, and allowing to use computational resources of the CBS. Thank you, a lot.

Finally, I want to express my gratitude for the support my family and friends offered. In particular, I want to thank Laura for her unconditional love and trust.

Indy Wijngaards

Tilburg, June 2016

Abstract

The present study attempted to examine the relative importance of a large variety of variables in the prediction of burnout complaints. The relative importance of individual variables was examined within variable categories; five variable categories were distinguished: job, social, organizational, attitudinal, behavioral and personal characteristics. Besides examining within-category variable importance, this study aimed at ranking the variable categories in importance. In accordance, two experiments were conducted: i) using Johnson's (2000) relative weight method and Breiman's (2001) random regression forests within-category variable importance was computed and ii) using random regression forests and Kruskal's (1987) LMG method variable category importance was determined.

A highly representative sample of the Dutch working population of 80,586 respondents was used for these analyses. In total, 95 independent variables divided in five variable categories were considered in this study. Both unsurprising (e.g. employability, supervisor support, job control, workload, job insecurity, general health, occupational health and safety practices, organizational changes) and surprising variables (e.g. ethnicity, screen work, physical job demands, household composition, marital status, demotion, organizational change) were identified as essential in forecasting burnout complaints. In addition, the results indicated that especially attitudes and behaviors and social characteristics play an important part in predicting burnout complaints. The results showed that no method really outperformed the other. Instead, the random regression forest performed better in some cases, while the relative weights method and LMG method performed better in other cases.

In the end, several theoretical and methodological limitations are presented, and interesting future research recommendations are discussed. The present study shows the importance of inductive research in the field of occupational health and human resource management, may inspire scholars to use appropriate variable importance methods, and could function as a fruitful basis for future hypothesis building.

Table of Contents

1. Introduction

In an era of ever-increasing globalization, competition and digitalization, the nature of work changes constantly; often deteriorating employee wellbeing and health (Landy & Conte, 2016; Peeters, De Jonge, & Taris, 2013). According the World Health Organization, work stress is taking on epidemic proportions in modern workplaces, costing organizations businesses billions of dollars (Cardon & Patel, 2015). In 2015, Hooftman et al. (2016) reported that over one third of the working population in the Netherlands suffers from work stress. The same study showed that more than 13 percent experiences burnout complaints. Burnout is considered to be “a severe form of psychological response/consequence of stress” (Siu, Cooper, & Phillips, 2014, p.70) or, more formally, the mental condition of emotional exhaustion, depersonalization and decreased personal accomplishment at work (Maslach, Schaufeli, & Leiter, 2001). The costs of work stress and burnout are especially determined by its consequences; prevalent ones are sickness absence (Cooper & Dewe, 2008; Roelen, Koopmans, Notenbomer, & Groothoff, 2008; Siu et al., 2014), decreased job performance (González-Morales & Neves, 2015; Peeters et al., 2013), and increased turnover rates (Paris & Hoge, 2010). Sickness absence, defined as “non-attendance at work due to poor health and/or poor wellbeing” (Peeters et al., 2013, p. 367), imposes a heavy burden for organizations and societies as a whole (Cooper & Dewe, 2008; Roelen et al., 2008). Similarly, poor job performance and increased turnover rates can have disastrous consequences for organizational performance (Hancock, Allen, Bosco, McDaniel, & Pierce, 2013; T. W. Taris & Schreurs, 2009).

This study addresses the continuous interest of practitioners and academics in work-related health (Black & Frost, 2011; Schaufeli, Leiter, & Maslach, 2009), by conducting an explorative research that aims at discovering the most important personal, organizational, job, social, attitudinal, and behavioral determinants of the burnout complaints in the Netherlands. The relevance of this study is threefold.

First, although being aware of the fact that burnout is a well-established academic subject, this study does contribute to the literature by incorporating an uncommonly large and diverse variety of predicting variables, and analyzing a comprehensive, highly representative sample of the Dutch working population. As an illustration, the vast majority of studies in the field fail to systematically incorporate contextual and personal characteristics such as sector, industry and occupations in the prediction of burnout (Pawlowski, Kaganer, & Cater III, 2007; Sparks & Cooper, 2013; T. Taris, Houtman, & Schaufeli, 2013). Studies often draw broad theoretical conclusions based conveniently sampled respondents (Howitt & Cramer, 2007), originating from very specific segments of the labor force (e.g. nurses, teachers, highly educated professionals, geographic region). Besides, academics claim that studies’ sample sizes are often unreliably small for the prediction of health-related work outcomes such as burnout (Schaufeli, Bakker, & Van Rhenen, 2009), hereby lacking sufficient statistical power (Maxwell, Kelley, & Rausch, 2008). Even though many meta-analyses have attempted to diminish the above research flaws, they inevitably deal with caveats of their own (e.g. publication bias, search bias, selection bias, heterogeneity of results, Walker, Hernandez, & Kattan, 2008).

Second, this study aims at conducting relatively advanced linear relative importance and random regression forest analyses to determine what factors play the most vital role in predicting burnout complaints. This is important from a practical, theoretical and methodological viewpoint. Practically, Siu et al. (2014) pointed out that designing interventions is of vital importance in preventing or combatting negative health-related work outcomes, such as burnout. An abundance of studies suggests that having insights about relative importance of potential stressors is necessary for guaranteeing their relevance and cost-efficiency (Fragoso et al., 2016; Garrosa, Moreno-Jimenez, Liang, & González, 2008; Nahrgang, Morgeson, & Hofmann, 2011; Pawlowski et al., 2007). Theoretically speaking, it appears that discovering the most important predictors is crucial for drawing valid conclusions and building reliable theories (Piha, Laaksonen, Martikainen, Rahkonen, & Lahelma, 2009; Tonidandel & LeBreton, 2011; Tonidandel, LeBreton, & Johnson, 2009). Besides its use for deductive research, variable importance analysis also lends itself for exploratory research (Jebb, Parrigon, & Woo, 2016). In terms of methodology, the present study contributes to the literature by adopting reliable measures of variable importance. Scholars posited that many researchers conclude their studies with assumptions about relative importance based on inappropriately simplistic estimates such as standardized regression coefficients and stepwise regression analysis (Behson, 2012; Grömping, 2015; Tonidandel et al., 2009). Considering the present study's research design (i.e. many variables and instances), the selected methods could potentially outperform the simplistic ones, as they suffer less from multicollinearity and overfitting (Matsuki, Kuperman, & Van Dyke, 2016). Along the same line, the selected methods may prove particularly useful, as they do not rely on significance levels too much compared to the more simplistic methods (LeBreton, Hargis, Griepentrog, Oswald, & Ployhart, 2007). Even though using these parametric relative importance analyses seems to be a clever way to go, the exclusive use of these methods would be naive. It turns out that scholars in the fields of work psychology and occupational health often wrongly assume linear relationships between work characteristics and occupational health (Karanika-Murray & Cox, 2010), often unjustly conducting regression analyses which inherently assume so. To overcome this limitation, this study will conduct and evaluate both a method that assumes linearity (i.e. relative importance analysis) and a method that does not (i.e. random regression forest), and compare the results in the end.

Third, as this study intends to analyze a wide variety of predictors, this study embraces a more exploratory (i.e. inductive or data-driven) research approach. Because the social (Berente & Seidel, 2014), management (Eisenhardt, Graebner, & Sonenshein, 2016) and human resource management (HRM) sciences (Woo, O'Boyle, & Spector, 2016) are dominated by theory-driven, deductive research (Locke, 2007; Spector, 2015), this study hopes to contribute to the rebalancing of the scales. Moreover, as mentioned by Markoulli, Lee, Byington, and Felps (2016), adopting this type of research design allows this study to break the current habit of studying concepts in work psychology and HRM within narrowly defined clusters and silos; hereby ensuring a comprehensive and integrative approach.

In line with the study's research aim and its rationale in terms of societal, scientific and practical relevance, this study will address the following problem statement (PS):

PS: *What predictors are most important in predicting burnout complaints within the Dutch working population?*

In an effort to address the above problem statement, two research questions are formulated. The first one aims at discovering the most important predictors of burnout complaints within five variable categories: personal characteristics, organizational characteristics, job characteristics, social characteristics, and work attitudes and behaviors. As such, the consequent research question (RQ) states:

RQ1: *What are the most important personal, organizational, job, social and attitudinal/behavioral predictors of burnout complaints?*

Besides investigating variable importance within these categories, this study also intends at discovering a hierarchal ranking between the variables categories. The second research question therefore concerns the investigation of the most variable categories in the prediction of burnout complaints, and reads as follows:

RQ2: *Which of the variable categories predict burnout complaints best?*

The next section will present a literature review and contextualize the present study. In the subsequent section, a description of the experimental setup is described. Hereafter, the results of the experiments and its interpretations are exhibited. To finish, a conclusion is presented, and research limitations and recommendations for scholars and practitioners are discussed.

2. Related work

In this section, first, the inconsistency of research findings and meta-analytical studies will be reviewed. Second, the methods for variable importance assessment will be discussed. Finally, the categorization of the variables as used in this study will be explained. Considering the large variety of variables that is included in this study and the consequent impossibility to cover them in text, the variables' conceptualizations, operationalization and definitions are schematically summarized in Appendix A.

2.1 Limitations of primary research

The academic field of employee health and wellbeing is well-developed, ever-growing and mostly theory-driven (Danna & Griffin, 1999; Peeters et al., 2013; Piotrowski, 2012; Woo et al., 2016). Yet, the bulk of the published studies a) is based on rather small sample sizes, b) focuses on very specific branches or occupations, and c) presents contradictory results (Edwards, Burnard, Coyle, Fothergill, & Hannigan, 2000; Ones, Viswesvaran, & Schmidt, 2016; Stone & Rosopa, 2016).

Several studies and literature reviews are presented below that illustrate the inappropriate smallness of sample sizes, narrowness of sample contexts, contradictoriness of results, or some combination of the three. Studies by Xie and Johns (1995) and Pawlowski et al. (2007) contrast each other. Xie and Johns' (1995) study of 418 respondents suggested that overstimulation is more important than understimulation in predicting work stress among professionals. A qualitative study by Pawlowski et al. (2007) of 20 information technology professionals implied the exact opposite. A literature review of burnout studies among nurses by Adriaenssens, De Gucht, and Maes (2015) showed that studies are inconclusive with respect to the predictive power of a variety of demographic (e.g. gender, age), social (e.g. social support) and job characteristics (e.g. physical demands). For example, Sorour and El-Maksoud (2012) and Van Der Ploeg and Kleber (2003), respectively, researched 58 and 123 health care professionals, and found that emotional demands at work, again respectively, have a positive and negative effect on burnout prevalence. A review of Watts and Robertson (2011) concluded that empirical evidence about the effect of gender and social support on burnout complaints of university teaching staff is also not uniform.

Whereas these studies and reviews show that contradictory findings exist between primary studies with relatively small sample sizes and narrow sample contexts, this trend also extents to larger studies with supposedly more generalizable findings. Norlund et al. (2010) researched 1,000 Swedish employees and found out that age has a strong negative association with burnout complaints. In contrast, the study of Lindblom, Linton, Fedeli, and Bryngelsson (2006) of 3,000 Swedish employees indicated that a negative significant relationship exists between age and burnout occurrence. Most interestingly, a research based on a sample of 6,091 Swiss employees suggested that age is of no importance at all in predicting burnout complaints (Brauchli, Bauer, & Hämmig, 2011).

2.2 Meta-analyses as imperfect alternatives

To overcome the limitations of the primary studies, meta-analyses attempt to objectively summarize them, and draw conclusions based on higher statistical power (Stone & Rosopa, 2016). Citing the inventor of the term, a meta-analysis basically concerns "the analysis of analyses" (Glass, 1976, p.3). It allows researchers to see through the inconsistency of different studies, and address hypotheses in a reasonably conclusive way (Walker et al., 2008). The same occurred in the field of burnout (e.g. G. M. Alarcon, Eschleman, & Bowling, 2009; Berry, Lelchook, & Clark, 2012; Crawford, LePine, & Rich, 2010; Faragher, Cass, & Cooper, 2005; Purvanova & Muros, 2010). Peeters et al. (2013) reviewed the (meta-analytical) literature and concluded that work overload, time pressure, number of working hours, role problems, work-home interference, number of clients and recipients, emotional demands, lack of social support, lack of job control, lack of feedback and poor participation in decision making are the most important (job-related) antecedents of burnout.

Unfortunately, meta-analyses deal with methodological issues of their own. Stone and Rosopa (2016) and Walker et al. (2008) describe several; e.g. i) the selection, ii) validity, iii) sample sizes and quality, and iv) heterogeneity of methods of primary studies that meta-analyses include. Firstly, meta-analyses select studies based upon the availability of published literature (Stone & Rosopa, 2016), thereby sustaining the 'file drawer phenomenon' (Walker et al., 2008). Many relevant studies remain unpublished and stay in scholars' file drawer (e.g. results that do not confirm hypotheses, journal editors prefer work of more established researchers), in turn, biasing estimates and effect sizes in the meta-analyses. Secondly, meta-analyses depend on the validity of the studies they include in the analyses. Meta-analyses may include studies that suffer from low construct, external, internal and statistical conclusion validity (Stone & Rosopa, 2016). Accordingly, when including these low-quality primary studies, a meta-analyses is affected by the "garbage in, garbage out" principle (Egger, Smith, & Sterne, 2001). Thirdly, meta-analyses regularly include studies with samples that are small, unrepresentative or both. This has severe negative effects on the homogeneity of effect sizes, and may result in flawed or misleading interferences about the findings (Stone & Rosopa, 2016). Fourthly, as indicated by Murphy (2015), meta-analyses habitually incorporate studies that are dissimilar in terms of methods, measures, samples and contexts. The inclusion of these studies can strongly bias the meta-analytical results (Murphy, 2015; Stone & Rosopa, 2016).

As an illustration to the above, meta-analyses usually analyze the majority of studies in the field, and disregard the national culture and origin of the respondents as represented in the study (Schmidt & Hunter, 2014). Although burnout is a global phenomenon (Schaufeli, Leiter, et al., 2009), Halbesleben and Buckley (2004) and Adriaenssens et al. (2015) note that scholars should be prudent with generalization in terms of causes and manifestation of burnout across nations and cultures. Considering these cross-cultural and cross-national differences (Schaufeli, Leiter, et al., 2009), such meta-analyses might be severely biased. For instance, Carod-Artal and Vázquez-Cabrera (2013) explained that, even though employees from both developed and developing countries perform demanding and constraining activities, employees in developing countries are likely to experience more work stress and burnout complaints due to factors outside their work environment (e.g. poor nutrition and hygiene, illiteracy). Moreover, meta-analyses about burnout suffer from the selection bias, since studies from specific continents (e.g. Europe, North-America) and countries (e.g. Brazil, non-Muslim countries) are overrepresented in the body of scientific literature (Carod-Artal & Vázquez-Cabrera, 2013). They may also be biased by the heterogeneity of methods limitation; Halbesleben and Buckley (2004) provide a rationale. Reviewing the most-commonly used survey on burnout (i.e. Maslach Burnout Inventory), they postulate that there may be large differences across countries in the way respondents interpret the phenomenon of burnout, perceive social acceptance of publically expressing it, and answer survey questions in general.

In an attempt to overcome the limitations of meta-analyses and to ensure appropriate generalizability of the results, this study concentrates on a particular target population: the Dutch working population in 2014 and 2015. This research can be classified as a secondary study, as it concerns the reanalysis of data (Glass, 1976). The data is collected and initially analyzed by the central bureau of statistics (CBS) in the Netherlands. This government office has the task, among others, of carrying out research and publishing statistical data on working conditions, occupational accidents, work content and work experiences (Hooftman et al., 2016). Although data (e.g. large random sample, validated measurement instruments) is of outstanding quality, the CBS' analyses based on the NEA data are usually quite elementary (e.g. group comparisons and basic significance testing). Due to data confidentiality reasons, more advanced analyses are usually not performed on the data. An example of a more advanced data analysis is variable importance assessment.

2.3 Methods for assessing variable importance

Both primary studies and meta-analyses are often interested in making interferences about the relative importance of variables (Fragoso et al., 2016; Grömping, 2015; Luchman, 2014; Nimon & Oswald, 2013; Tonidandel & LeBreton, 2011). Nonetheless, scholars commonly rely on inappropriately simplistic and unreliable measures of variable importance, such as standardized regression coefficients (Braun & Oswald, 2011; Nimon & Oswald, 2013) and stepwise regression analysis (Grömping, 2006). According to Grömping (2006, 2007, 2015), these methods are appropriate for studies where regressors are uncorrelated, as in that case "each regressor's contribution is just the R^2 from univariate regression, and all univariate R^2-values add up to the full model R^2" (Grömping, 2006, p. 1). She, however, explains that sciences these days are predominantly grounded on observational data (e.g. survey); a data type variables characterized by correlated regressors. Lance, Lance, and Vandenberg (2010) confirm that this trend also exists within the social and organizational sciences. Thus, it seems that the use of simplistic methods is problematic when researchers are dealing with multicollinearity, the issue of moderately or highly inter-correlating regressors (Chiaburu, Oh, Wang, & Stoverink, 2016; Grömping, 2015).

Many researchers do not take these methodological issues into consideration when investigating relative importance of variables in the prediction of burnout complaints. For example, Garrosa et al. (2008) used standardized regression coefficients to determine the relative importance of a large variety of inter-correlating variables in prediction of burnout among nurses. Tsigilis, Zachopoulou, and Grammatikopoulos (2006) aimed at unraveling the relationship between job satisfaction and burnout in the Greek public and private sector by simply comparing beta coefficients and R^2-values. Kokkinos (2007) adopted stepwise regression analyses to investigate which of the inter-correlating regressors predicted teacher burnout best. Finally, Hauge, Skogstad, and Einarsen (2010) ignored the fact that regressors moderately correlated, and made assumptions about the relative importance of bullying in relation to burnout based on stepwise regression analyses.

Alternatively, scholars recommend more adequate measures of variable importance, such as linear relative importance analysis and random regression forests (Grömping, 2006, 2015; Johnson & LeBreton, 2004; Matsuki et al., 2016). Nahrgang et al. (2011) accurately followed these recommendations, and used Johnson's (2000) epsilon (denoted as ε), also called relative weights, to meta-analytically assess the relative importance of a variety of job demands and resources in relation to burnout. The relative weights method is based on the traditional linear regression model, and is characterized by the orthogonalization of the independent variables (i.e. transforming variables in such a way that they do correlate with each other). After this transformation, the orthogonal variables (Z_k) are regressed on the dependent variable (Y), resulting in a set of uncorrelated regression coefficients, denoted as β_k. Next, Z_k are related back to the original predictors (X_j), resulting in another set of regression coefficients denoted as λ_{jk} (Johnson, 2000; Johnson & LeBreton, 2004). In the process of regressing the original variables on the orthogonal coefficients, the problem of multi-collinearity becomes irrelevant (Grömping, 2015; Johnson & LeBreton, 2004). After all, the regression coefficients are ascribed to the uncorrelated orthogonal variables and not to the original correlated coefficients. Due to the orthogonalization, the regression coefficients (λ_{jk}) equal to the correlations between the original and transformed variables, and, in turn, each $\lambda_{jk}{}^2$ equals "the proportion of variance in Z_k that is accounted for by X_j" (Hawthorne, 2011, p. 8). The calculation of the relative weights for each predictor is based on the two sets of coefficients that result from these two regression procedures (i.e. β_k and λ_{jk}). In specific, the amount of variance for every Z_k by X_j is multiplied by the amount of variance accounted for by every Z_k, whereafter all these products are summed. Kath, Stichler, Ehrhart, and Sievers (2013) used Budescu (1993) dominance analysis – another reliable and computationally more intensive variable importance measure – to explore the relative importance of five inter-correlating workplace stressors for nurse manager perceptions of job stress. Other reliable methods closely related to dominance analysis were developed, such as LMG (Kruskal, 1987; Lindeman, Merenda, & Gold, 1980) and PMVD (Feldman, 1999). The two methods respectively concern "the unweighted or weighted average of sequential variances over all possible orderings of regressors" (Grömping, 2015, p. 143). In general, these methods both start with calculating the semi-partial correlations of each regressor in the model (i.e. to what extent does R^2 change when the regressor is added to the model). This is done for every order the variables can introduced in the model. For example, in case three variables a, b, and c are regressed on a dependent variable, the semi-partial correlations of a are computed for the thee orderings: $\{a,b,c\}$, $\{b,a,c\}$, $\{b,c,a\}$ (Bi, 2012; Lindeman et al., 1980). Subsequently, the average of all the semi-partial correlations is computed. It is argued that methods such as relative weights, LMG, PMVD and dominance analysis are statistically and substantially superior to the aforementioned simplistic measures (Budescu & Azen, 2004).

Yet, it is sometimes contended that the adoption of linear methods such as correlation, odds ratios and regressions to assess relative importance is incorrect for studying work-health relationships (Ferris et al., 2006; Karanika-Murray & Cox, 2010). This would include the seemingly reliable measures of variable importance such as relative weights, LMG, PMVD, and dominance analysis (Grömping, 2015). Even though contemporary organizations are dynamic, adaptive and complex, and many theories in the field explicitly and implicitly assume curvilinearity, homeostasis, systems approaches, complexity and nonlinearity, researchers often presume stable linear relationships between work and health (Karanika-Murray & Cox, 2010). Also, empirical evidence clearly indicates that many work-health relationships are potentially of nonlinear nature (e.g. Borg, Kristensen, & Burr, 2000; Ceja & Navarro, 2012; De Jonge & Schaufeli, 1998; Karanika-Murray & Cox, 2010; Ladstätter et al., 2016; Noblet & Rodwell, 2009; Rydstedt, Ferrie, & Head, 2006; Zivnuska, Kiewitz, Hochwarter, Perrewé, & Zellars, 2002). For example, De Jonge, Reuvers, Houtman, Bongers, and Kompier (2000) found that the model that included nonlinear relationships was preferable for predicting emotional exhaustion and depression (i.e. two dimensions of burnout). Ladstätter et al. (2016) examined the relationship between work characteristics and burnout in China using artificial neural networks and linear regression analysis, and concluded that the nonlinear method is evidently superior to the linear one. The authors found and visualized that, for example, the relationship between work overload, job control and emotional exhaustion, and the relationship between troubled interactions with doctors, patients and relatives, job control and emotional exhaustion resemble a nonlinear pattern. Camerino et al. (2010) explored relationship between work-family conflict and occupational health and safety (OHS) among Italian nurses using Bayesian networks, and concluded that the relationship is nonlinear. Thus, the literature suggests that besides using variable importance measures which inherently assume linearity (Nimon & Oswald, 2013), studies would surely benefit from using (machine learning) methods that do not inherently assume so (e.g. decision trees, random regression forests, artificial neural networks) (Grömping, 2009; Tonidandel, King, & Cortina, 2016; Ye, Yang, & Yang, 2016).
This is why random regression forests are used in this study to examine relative variable importance in a non-parametric way. The random regression forest algorithm allows researchers to effectively deal

with high dimensional data and tends to not overfit data (Grömping, 2009, 2015; Jin, 2014). Random regression forests are fundamentally different from linear regression methods as described in the previous paragraphs (Pretnar, 2015). Random forests are based on recursively partitioned regression trees. These trees are formed by asking flow-like questions about every single variable in the dataset, "subdividing a sample into groups that are as homogeneous as possible by minimizing the within-group variance, in order to determine a numerical response variable" (Smith, Ganesh, & Liu, 2013, p. 86). Random regression forests are built of trees of randomly selected variables and randomly selected observations, eventually basing its predictions on the average of all trees (Grömping, 2009) and forecasting the independent variable with the model with the smallest mean squared error (MSE) (Smith et al., 2013). Variable importance is determined by the random permutation of the values of one independent variable across all decision trees, and the refitting of the model (Grömping, 2009; Matsuki et al., 2016). Randomly permuting the values is done to 'break' the relationship between the independent and dependent variable. The importance of variable is determined by the extent to which the prediction accuracy drops when a particular relationship has disappeared (i.e. in regression tasks the percentage increase of the MSE). For important variables, the random permutation causes a severe increase in MSE. Variables with little or no importance, the MSE is stable or could even increase.

2.4 Exploratory research design and categorization of variables

Regarding the present study's research design, Woo et al. (2016) and Locke (2007) explain that today's organizational science focuses too much on hypothesis-driven, deductive reasoning. For instance, Spector (2015) compared the percentage of deductive scientific articles in the *Journal of Applied Psychology* in 1971 with the percentage in 2015. Over the years, the percentage increased from 28 percent to 100 percent. Woo et al. (2016) argue that instead academic disciplines should aim at discovering novel knowledge (i.e. induction), explaining and theorizing about these insights (i.e. abduction), and testing them in formal models (i.e. deduction). As this study has access to a large, representative and reliable dataset, it was decided to adopt a more data-driven or inductive research approach. The recommendations of Edmondson and McManus (2007) are loosely followed; merging the intermediate and mature archetypes of methodological fit in field research. Even though the burnout research field is already very mature, it seems still of relevance to conduct an "open-ended inquiry about a phenomenon of interest" and offer "an invitation for further work on the issue or set of issues opened up by the study" (Edmondson & McManus, 2007, p.2006). This is expected to be relevant, because this study adopts relatively advanced formal measures of variable importance (i.e. random regression forests and relative importance analysis), and incorporates a particularly comprehensive set of research constructs, some of which are rarely included within studies (e.g. sector, place of living, marital status).

Guiding the process of inductively discovering the most important predictors of burnout complaints, and avoiding purposeless and useless data analysis (Ketokivi & Mantere, 2010), it is important to use a broad theoretical framework (Woo et al., 2016). For this study, it was therefore decided to consider the widely-used job demands-resources (JD-R) model (Bakker & Demerouti, 2007; Schaufeli & Taris, 2014). The model is, among others, often used to predict work-related health outcomes such as burnout and stress (Bakker & Demerouti, 2007; Nahrgang et al., 2011; Schaufeli, Bakker, et al., 2009). In this model, variables can be assigned to one of five variable categories: job demands (e.g. long working hours, working overtime), job resources (e.g. job autonomy, social support), personal resources (e.g. optimism, resilience), attitudes related to work motivation (e.g. job engagement, work involvement) and work outcomes (e.g. turnover, job performance). Variables that do not fit the formal definition of the variable categories are usually included as control variables. As this study aims at regressing a substantial number of independent variables that cannot be assigned to one of these categories (e.g. gender, ethnicity, organizational size, province, independent contractor, general health, satisfaction with working conditions), it was decided to use the literature on the JD-R model to design a set of variable categories that appropriately fit the study's inductive research design. As a result, all independent variables were categorized in one of the following five variable categories: personal characteristics, organizational characteristics, job characteristics, social characteristics, and attitudes and behaviors. In Table 1, the independent and dependent variables and corresponding variable categories as used in the present study are displayed. A complete overview of all independent variables as used in the analyses, including categorization, operationalization and related studies is to be found in Appendix A.

Table 1
Overview independent variables and variable categories

Behaviors and attitudes	Job characteristics	Organizational characteristics	Personal characteristics	Social characteristics
Employability	Working time	Size	Gender	Supervisor support
Turnover intention	Shift work	Sector	Chronological age	Co-worker support
Under- or over qualification	Working in weekends	Organizational changes	Educational level	Interpersonal conflict
Job satisfaction	Working at nights	OHS practices	Ethnicity	Intimidation
Satisfaction with working conditions	Working overtime		Marital status	Physical violence
	Working at home		Province	Bullying
	Seniority in function		Independent contractor	
	Managing position		Organizational tenure	
	Dangerous work		Multiple jobs	
	On-call work		General heath	
	Workload		Occupational accident	
	Physical job demands (Environmental)		Household composition	
	Physical job demands (Heavy loads)			
	Physical job demands (Unusual or tiring body positions and movements)			
	Cognitive job demands			
	Job control			
	Screen work			
	Demotion			
	Promotion			
	Job change			
	Job enlargement			
	Job insecurity			
	Employment contract			

Note. OHS = Occupational Health & Safety

3. Method

First, the dataset will be described. Next, the methods for outlier detection and missing value imputation will be outlined. Then this study will elaborate on the variable creation, feature description and software that was used. After this, the findings that resulted from the exploratory data analysis (EDA) will be discussed. This section will end with the discussion of the used algorithms and methods that were used to conduct the experiments.

3.1 Dataset description, feature description and software

3.1.1 Dataset description

The study is conducted in supervision of the Dutch central bureau of statistics (CBS) and Tilburg University, and will base its conclusions on a Dutch employment survey *Natiale Enquête Arbeidsvoorwaarden* (NEA). The CBS annually conducts this large-scale survey. It yields a large and representative random sample of the Dutch working population on working conditions, occupational accidents, work content and work experience (Hooftman et al., 2016). From 2014 onward, the annual data collection results in around 45,000 valid respondents, gathered from a sample size of 140,000. The CBS choose to conduct a random sample stratified based on industry or branch. Within these strata, a systematic Probability Proportional to Size-sample was taken. This means that all people had the same likelihood to be sampled, except for younger people (15 to 23 years) and people who have a non-western migration background. People in these groups had a 50 percent higher chance to be sampled. In addition, people residing in institutional houses, living in a refugee center, that desire to be anonymous, or were already sampled for previous NEAs were excluded from the final sample. The data is open for the public upon request, but obviously in an anonymized version. For this study, the CBS allowed the analysis of the not anonymized data. This was permitted, as the researcher signed a temporary employment contract at CBS. The CBS also facilitated the use of demographic information (i.e. country of birth, marital status, and place/province of living) as an additional data source. This data is stored in the Dutch population register (in Dutch: *Gemeentelijke Basisadministratie Persoonsgegevens*). This study merged the surveys of 2014 and 2015; this resulted in a total sample size of N = 80,586 respondents (i.e. rows or instances).

With regard to the research procedure the CBS adopted, the respondents had two ways of participating: Computer Assisted Web Interviewing (CAWI) and Paper and Pencil Interviewing (PAPI). Respondents were not forced to answer questions; it was allowed to skip questions, and a ‘no answer’ answer category was provided. An outcome of this decision is that some questions dealt with significant non-response. For example, the question regarding satisfaction with working conditions in 2014 had a non-response between 11.8 and 18.0 percent. This problem was addressed during the pre-processing of the data (i.e. missing data imputation). The CBS conducted reliability analyses on several scales (i.e. summarized items), and concluded that all scales had sufficient to good reliabilities based upon rules of thumb about Cronbach α of de Heus, van der Leeden, and Gazendam (1995).

3.1.2 Outlier detection and missing value imputation

While the data was already thoroughly preprocessed by the CBS, it was checked for outliers, missing values and other distinct characteristics in this study. Although the data as provided by the CBS did not contain outliers, it did contain a substantial proportion of missing data (i.e. 62038 values, 0.8 percent of the data). Newman (2014) explained that researchers often use too simplistic methods for dealing with this problem, e.g. using list-wise deletion or substituting missing values by the column mean. An interesting imputation method from the machine learning field that effectively deals with large datasets and mixed-type data are random forests (Stekhoven & Bühlmann, 2012). Golino and Gomes (2016) concluded that the use of random forests for data imputation is highly appropriate in psychological research, and often outperforms more traditional imputation methods. Another reason for choosing this method is its extraordinary performance on ordinal data (Cugnata & Salini, 2017), an item type which happens to be prevalent in the NEA data. Therefore, it was decided to use random forests to impute missing data in this study. Following recommendations of Stekhoven and Bühlmann (2012) and considering the computational resources available, the number of trees per forest was set to 10, the maximum number of iterations was set to 6, and the number of variables randomly selected per split was set to 6. The normalized root mean squared error (NRMSE) was used to evaluate the imputation method. Very poor imputation of data approaches a NMRSE value of 0, while perfect imputation implies a value of 1 (Stekhoven & Bühlmann, 2012). Audigier, Husson, and Josse (2016) explained that values close to 0 resemble the mean imputation method. The NRMSE-score of the present imputation was 0.477. Although this value is not particularly high, it is preferred over mean imputation, as a NRMSE-score of approximately 0 would be even more undesirable. The method was also preferred over list-wise deletion of cases, a commonly adopted imputation method within psychological sciences, as would result in a substantive decrease in analyzable observations. The 62,038 missing values would create a total of 29,325 useless cases; this is more than one-third of the entire dataset.

3.1.3 Variable creation

Once the dataset was fully preprocessed, variable creation occurred. To make sure the proposed concepts were properly measured and the number of variables was reduced, this study summarized items in line with recommendations of the CBS. For example, the NEA contained five questions measuring burnout complaints, which in this study will be summarized by computing a weighted average:

$$burnout\ complaints = \frac{burnout1 + burnout2 + burnout3 + burnout4 + burnout5}{5}$$

Another method of summarizing data was creating dummy variables that indicate the presence or absence of a certain behavior, attitude or occurrence (e.g. working at home; yes or no). For example, the province variable was transformed into 12 dummy variables. Several variables were recoded to improve their interpretability during the experiments. After pre-processing, summarizing items and creating dummy variables, the total number of variables in the dataset equaled 96.

3.1.3 Feature description

The definition, categorization and operationalization of all variables is to be found in Appendix A. The feature descriptions in terms of mean, standard deviation, minima and maxima are displayed in Appendix B. It was decided to not select variables prior to the analyses, as this study is of explorative nature. Nevertheless, according to research, the majority of the variables do have a significant relationship with burnout complaints (see Appendix B).

3.1.4 Software

The software used in this study includes the programming language R and RStudio. In total, eleven packages were used for preprocessing, analyzing and visualizing the data. These include dplyr, qgraph, missForest, dummy, psych, car, stats, ggplot2, randomForest, relaimpo, and RFgroove. Some basic data manipulation and graph creation was done in Microsoft Excel 2016.

3.2 Exploratory data analysis

This section addresses the study's EDA strategy, and is divided in three parts. Firstly, as a supplement to the descriptive statistics in Appendix B, a histogram of the burnout complaints variable is displayed in Figure 1. Secondly, several visualizations of are presented, each illustrating a relationship between an independent variable and burnout complaints. It was decided to highlight the personal and organizational variables in this EDA, as research points out that the literature often fails to systematically investigate the relationships between these types of variables and burnout complaints (Pawlowski et al., 2007; Sparks & Cooper, 2013; T. Taris et al., 2013). This is not surprising, as often studies fail to obtain a representative sample of the population (e.g. sampling older, female nurses, or highly educated psychology students) due to financial or time constraints. Hence, a boxplot is provided in Figure 2 that analyzes burnout complaints grouped by gender. A line graph is exhibited in Figure 3 that visualizes burnout complaints prevalence across ages. A bar-chart is showed that depicts average burnout complaints per education level in Figure 4. In Figure 5, burnout complaints prevalence across sectors is showed by means of a bar-chart. A line chart is displayed in Figure 6 visualizing burnout complaints across organizational sizes. Thirdly, a correlation network is showed that presents all relationships between all variables in Figure 7.

3.2.1 Burnout complaints – The scale's distribution

The histogram shows a right-skewed distribution with more than 15,000 respondents reporting no burnout complaints at all. Also, the mean of 2.05 (i.e. red vertical line) suggests that the distribution of 7-point Likert scale is not normal.

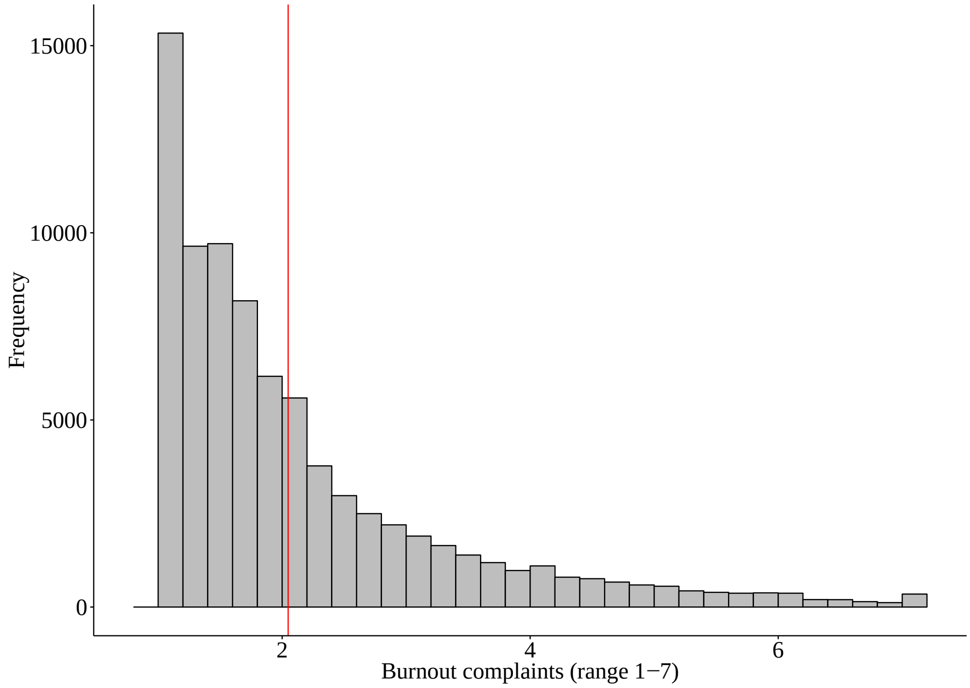

Figure 1. Histogram burnout complaints.

3.2.2 Burnout complaints and personal characteristics – Gender, age and education level

On the next two pages, three figures are displayed: a boxplot, line graph and bar-chart. The boxplot as showed in Figure 2 shows that the distribution for males and the distribution for females are close to identical, with no visible differences in medians, skewness and outliers. Furthermore, as displayed in Figure 3, burnout complaints occurrence seems to vary across educational levels to some extent. People with either primary education or an academic bachelor or higher professional education generally score higher on the burnout complaints scale. Employees with a higher secondary or lower professional education experience burnout complaints the least. Some substantial differences appear to exist across ages, as depicted in Figure 4. A peak is achieved around 30, whereafter a slight decline kicks in. The level of burnout complaints between 35 and 55 years is relatively stable. Between the ages 55 and 60 an increase in burnout complaints is noticeable. The steep decline in burnout complaints from 60 to 65 suggests that people approaching retirement experience significantly less burnout complaints than their younger counterparts. The large standard error at the end of the line is caused by a disproportionately small number of sampled respondents of 65 or older (i.e. a total of 1175 respondents).

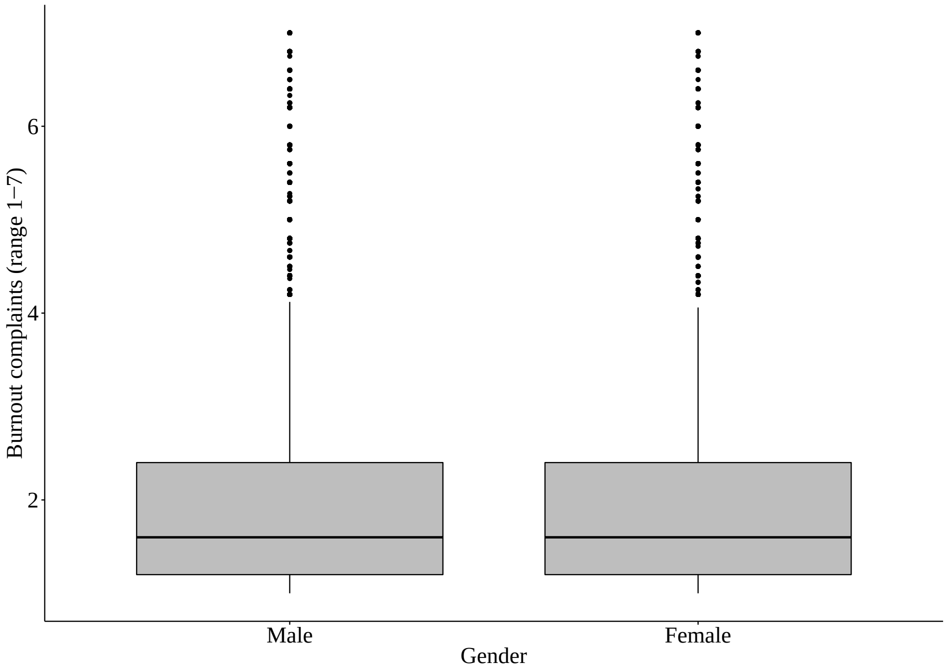

Figure 2. Burnout complaints for males and females.

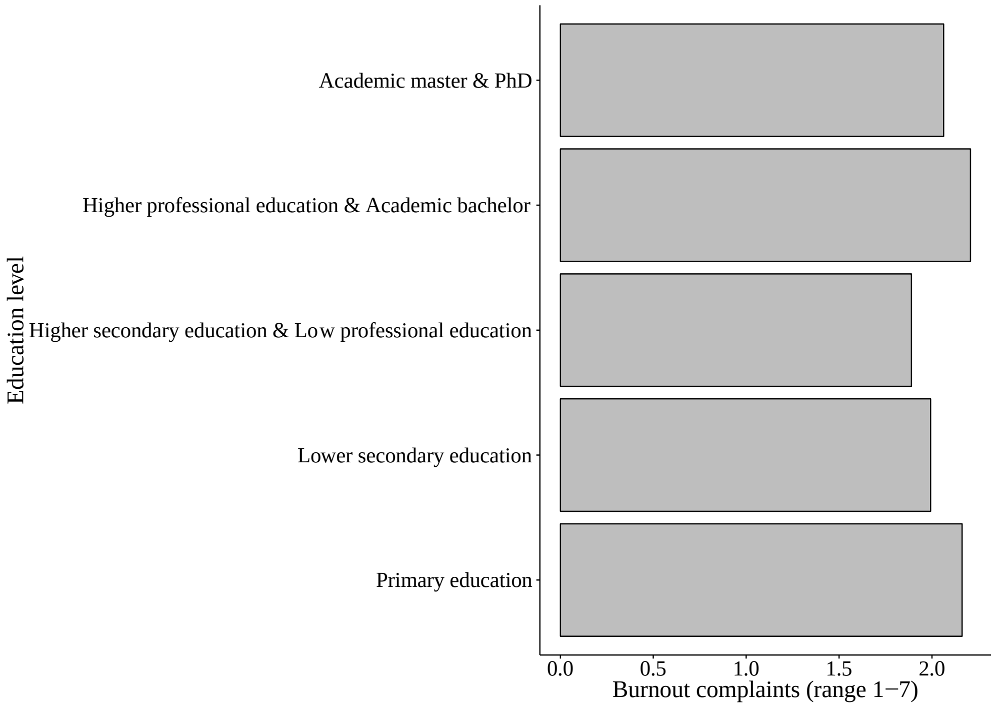

Figure 3. Burnout complaints grouped by educational level.

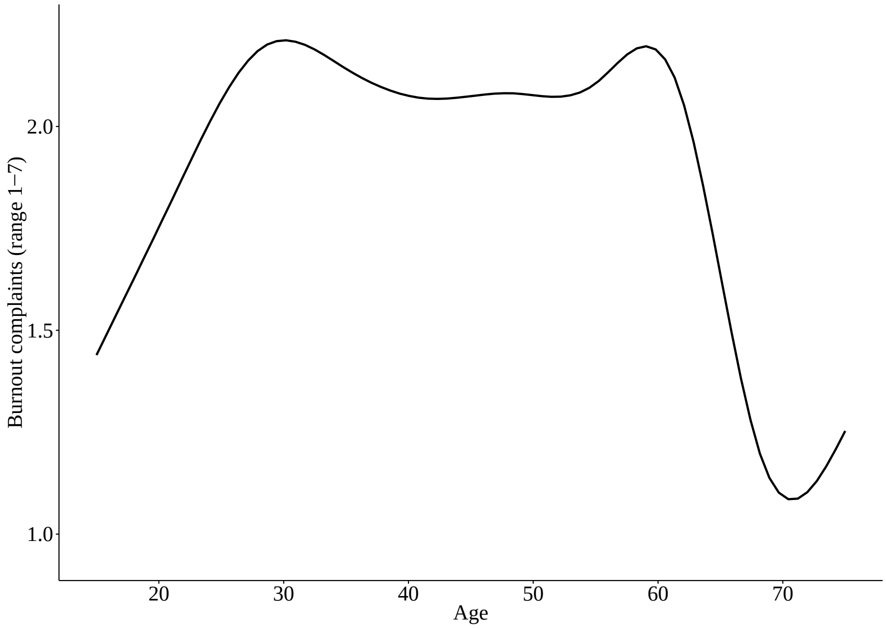

Figure 4. Burnout complaints across ages.

3.2.3 Burnout complaints and organizational characteristics – Organizational size and sector

The figures below visualize the relationship between two organizational characteristics and burnout complaints: organizational size and sector. First off, Figure 5 shows that some substantial differences exist in burnout complaints between sectors. Highlighting some interesting differences, employees in the education, and information and communication sector seem to be most prone to experience burnout complaints. Employees in health-care, financial services and public governance also appear to be more likely to experience burnout complaints. Interestingly, the forestry, fishery and agriculture sector and the hospitality sector show the lowest burnout levels in general. Figure 6 indicates that the larger the organization, the more likely employees experience burnout. More specifically, people working in organizations having fewer than 9 employees and people working in organizations having more than 250 employees experience respectively the least and the most burnout complaints.

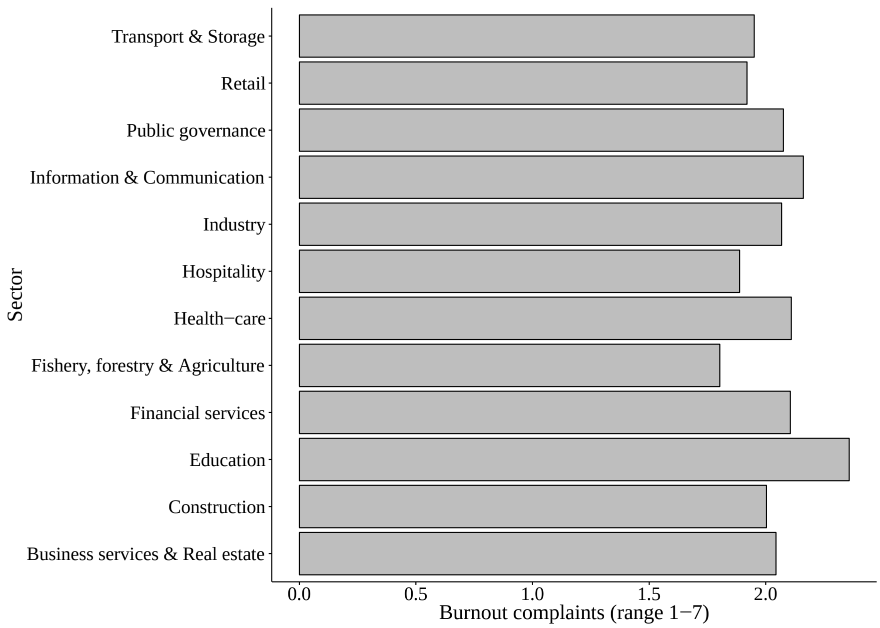

Figure 5. Burnout complaints across sectors.

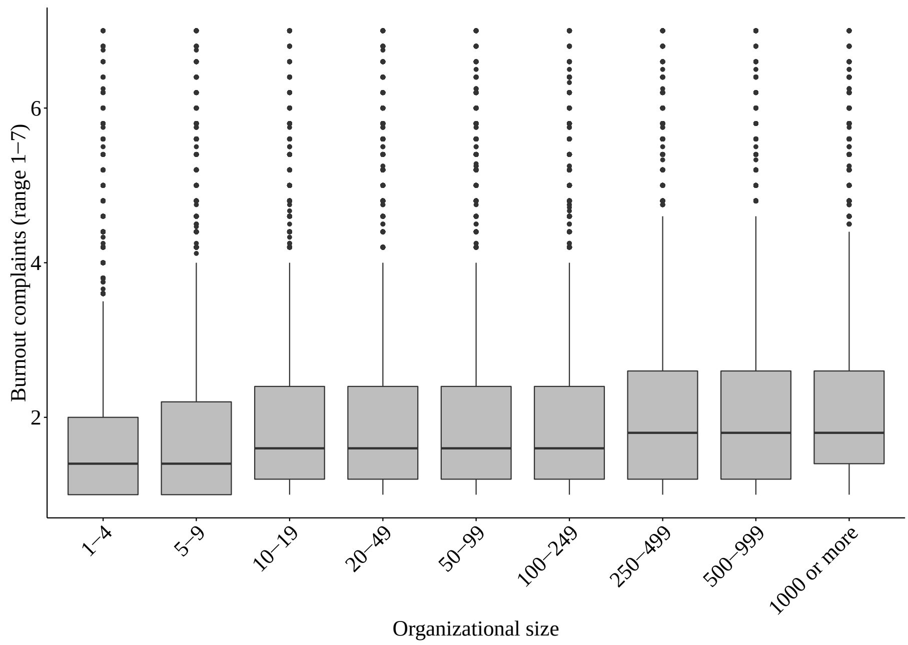

Figure 6. Burnout complaints across organizational sizes.

3.2.4 Correlation network

As an overview of the correlation between all variables, a correlation network was presented on the next page (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012). It was decided to not use a correlation matrix, as the total number of variables as used in this study exceeded the number of coefficients that logically can be presented in a one- or two-page correlation matrix. Within this correlation network, every variable is represented by a node and each correlation represented by an edge. The thickness of an edge corresponds to the height of a Pearson's correlation coefficient, with the minimal absolute correlation coefficient value of 0.15 (i.e. correlations below this value are not shown). For clarity, the five variable groups were visualized in clusters of nodes; all variables names are displayed below the figure. The correlation network showed that a large variety of variables correlate with each other, both within variable category as well as between variable categories. This means that the issue of multicollinearity is present within the present study's data.

It should be noted that some edges are not relevant to the study's purpose or even meaningful in general. Some nominal variables are transformed into more than two dummy variables (e.g. marital status, province, sector, and ethnicity). As these dummies were included in the bivariate correlation analysis, the correlations between these dummies were computed and visualized in the correlation network. Even though the significant correlations between these dummies are statistically valid, interpretation of the coefficients (or in this case: edges) becomes trivial. For example, the bivariate correlation between the industry and construction sector is -0.0901, and the correlation between dummy variables 'Non-western migration background' and 'Western migration background' is -0.0804. These negative associations between such dummies can simply be explained by the fact that a dummy variable will always have an opposite score from the other dummies in the group (e.g. when 'Construction' = 1, then 'Industry' = 0). In other words, each dummy has some potential to predict the others, plainly because each dummy is not similar to the others.

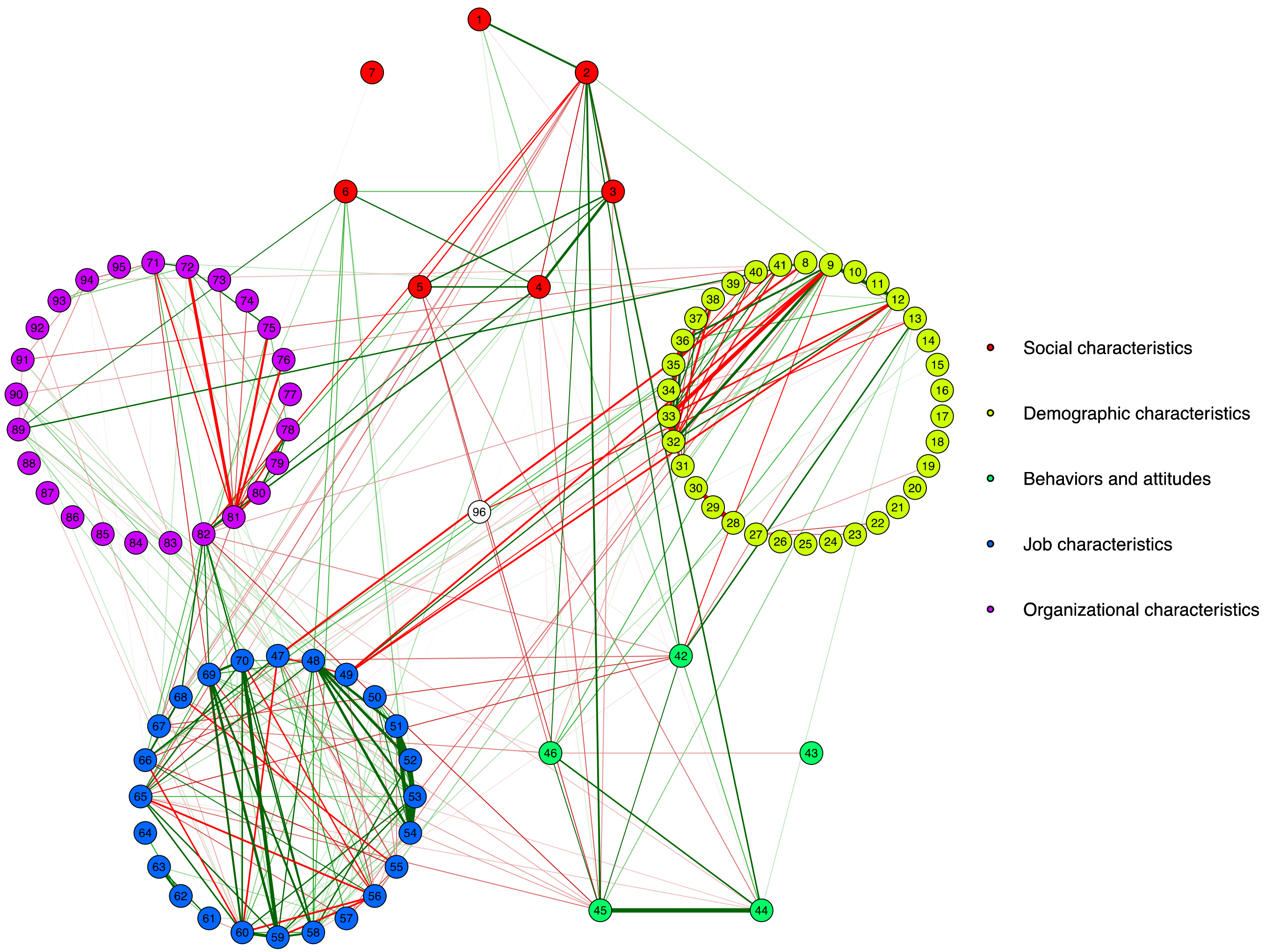

Figure 3

Correlation network

1 = Colleague support; 2 = Supervisor support; 3 = Bullying; 4 = Intimidation; 5 = Intra-personal conflict; 6 = Physical violence; 7 = Work-home interference; 8 = Gender; 9 = Age; 10= Educational level; 11 = Independent contractor; 12 = Organizational tenure; 13 = General health; 14 = Occupational accident occurrence; 15 = Multiple jobs; 16 = Drenthe; 17 = Flevoland; 18 = Friesland; 19 = Gelderland; 20 = Groningen; 21 = Limburg; 22 = Noord-Brabant; 23 = Noord-Holland; 24 = Overijssel; 25 = Utrecht; 26 = Zeeland; 27 = Zuid-Holland; 28 = No migration background; 29 = Non-western migration background; 30 = Western migration background; 31 = Divorced; 32 = Married; 33 = Never married; 34 = Widowed; 35 = Married with children; 36 = married without children; 37 = One-parent household; 38 = One person; 39 = Other or unknown; 40 = Unmarried with children; 41 = Unmarried without children; 42 = Employability; 43 = Under- and overqualification; 44 = Job satisfaction; 45 = Satisfaction with working conditions; 46 = Turnover intention; 47 = Working hours; 48 = Shift work; 49 = Employment contract; 50 = On-call work; 51 = Working 19:00-00:00; 52 = Working 00:00-06:00; 53 = Working Saturdays; 54 = Working Sundays; 55; Working overtime; 56 = Working at home; 57 = Managerial position; 58 = Dangerous work; 59 = Physical job demands – Heavy loads; 60 = Screen work; 61 = Job enlargement; 62 = Job change; 63 = Promotion; 64 = Demotion; 65 = Job control; 66 = Cognitive job demands; 67 = Job insecurity; 68 = Workload; 69 = Physical job demands – Environmental; 70 = Physical job demands – Unusual body positions; 71 = Organizational size; 72 = Reorganization; 73 =Acquisition of own company; 74 = Acquisition of other company; 75 = Organizational downsizing with forced layoffs; 76 = Organizational downsizing without forced layoffs; 77 = Merger; 78 = Outsourcing; 79 = Off-shoring; 80 = Automation; 81 = No change; 82 = OHS; 83 = Business services & Real estate; 84 = Construction; 85 = Culture, sport, recreation & other; 86 = Education; 87 = Financial services; 88 = Agriculture, fishery and forestry; 89 =Health-care ; 90 = Hospitality; 91 = Industry; 92 = Information & communication; 93 = Public governance; 94 = Retail; 95 = Transport & storage; 96 = Burnout complaints.

3.3 Experiments and analysis

This section describes the techniques adopted to investigate and evaluate the relative importance of independent variables in prediction of burnout complaints. First, the general methods for examining variable importance within the variable categories will be discussed. Second, the methods for creating a variable category ranking will be elaborated on. Third, the evaluation method is described. Finally, the two experiments will be explained.

3.3.1 Methods for within category variable importance

This study examined variable importance within the variable categories in two ways: relative weights and random regression forest. Firstly, Grömping (2015) explained that researchers who compute LMG scores with many explanatory variables might run in resource feasibility problems. Ye et al. (2016) indicated that this might already happen with 20 independent variables. Although Grömping (2015) clearly expressed her preference for computationally more demanding methods such as LMG and PVMD, she indicated that Johnson's (2000) relative weights method is the best choice for studies that incorporate many independent variables (*p*) and instances (*n*), and aim at decomposing variance and determining goodness of fit. Accordingly, besides the conceptual reason to break up the independent variables into categories, computational issues also are part of the rationale. After all, analyzing almost 100 independent variables would certainly cause resource feasibility problems, regardless the method selected. Because resource feasibility issues actually arose with the LMG method, this study used the relative weights method to examine relative importance in a parametric way. Section 2.3 provides information about how the relative weights method works. To evaluate this method, bootstrapping was used, a competitor of cross-validation (Grömping, 2006), and characterized by computer-intensively re-sampling test data to mimic the original data (Davison & Hinkley, 1997). Bootstrapping especially considers variable stability (Beyene, Atenafu, Hamid, To, & Sung, 2009), and overcomes a potential limitation of relative weights analysis (Tonidandel et al., 2009). Tonidandel et al. (2009) explained that the Johnson's relative weights method does not provide information about the significance of the relative weights, since the precise sampling distribution of the epsilons is unknown. When adopting bootstrapping (i.e. computing relative weights within resampled datasets, and aggregating these weights across all datasets), scholars can empirically derive an overall sampling distribution, create confidence intervals, and start considering the statistical significance of the relative weights. It has to be noted that the values in the confidence intervals can never include zero, as it is practically impossible to yield zero correlations in every single bootstrap (Johnson, 2004). This means that statistical significance of *individual* relative weights cannot be determined using the bootstrapping procedure (Tonidandel et al., 2009). As this study did not aim to do so and considered significance of the *differences* between the relative weights, this limitation of bootstrapping was considered irrelevant (Tonidandel et al., 2009). Bootstrapping for random regressors was used, a method where "the complete observation rows – consisting of regressors and response – are resampled" (Grömping, 2006, p.14). It was advised to used bootstrapping for random regressors instead of bootstrapping for fixed regressors, as the results of this study are based on a random sample of the overall population.

Secondly, random regression forests were used to assess relative importance of the explanatory variables. As an alternative to the parametric relative weights method, random regression forests seemed an exceptionally suitable method for examining variable importance (Grömping, 2009, 2015; Jin, 2014) and analyzing a large variety of independent variables (Grömping, 2009; Ye et al., 2016). Although already carefully explained in Section 2.3, a short overview of the random regression forest algorithms is presented below. Random regression forests are built of regression trees; objects "built by recursively partitioning the sample (i.e. the 'root node') into more and homogeneous groups, so-called nodes, down to 'terminal nodes'" (Grömping, 2009, p. 208). Each split of the regression tree is substantiated on the values of a single variable, and selected based upon a splitting criterion. The random regression forest algorithm, based on Breiman (2001) classification and regression trees (CART), is a machine learning method, and is characterized by the use of a large number of trees constructed from a limited number of variables and random selection of instances (Grömping, 2015). The increase in MSE was used as importance metric, as node impurity metric is considered inaccurate and biased (Grömping, 2009).

3.3.2 Variable importance between categories

Besides investigating within-category variable importance, it was also interesting to examine which category is most important in terms of predicting burnout complaints. Therefore, the variable categories were compared to each other, once with LMG (Grömping, 2007), and once with random regression forests (Gregorutti, Michel, & Saint-Pierre, 2015).

The LMG method was in this case a viable option, as no individual variable contribution was computed. A bootstrapping procedure was conducted. As already outlined, LMG for individual variable importance uses "the unweighted average of sequential variances over all possible orderings of regressors" (Grömping, 2015, p. 143), and calculates the semi-partial correlations of each predictor in the all different orderings. LMG for group-based variable importance (here: between-group variable importance) assessment works a bit different. Building on to the example in Section 2.3, let's again consider the thee variables: a, b, and c. Because the conventional LMG method combines the sums of squares of individual variables in different orderings, the following sums of squares would be relevant to consider: the model without any variables, the model with only a, the model with only b, the model with only c, and the model with all variables. Now, let's assume that variable a and b are grouped together. This means that the individual sums of squares of a and b become irrelevant, and the sum of squares of the model that includes a and b becomes the only relevant one. Hence, in terms of calculating semi-partial correlations for different orderings, the total number of combinations for the individual variable LMG method would equal six (i.e. 3! = 3 x 2 x 1 = 6), while the group-based LMG only considers two (i.e. $\{ab, c\}$, $\{c, ab\}$). Even though both computations would be viable in this hypothetical example, the group-based LMG option was clearly preferred over its traditional counterpart within this study. After all, the number of combinations in the traditional LMG method would equal 1.0323 x 10^{20} (95!), whereas the group-based method would only have to consider 120 combinations (5!).

Besides assessing variable category importance in a parametric way, Gregorutti et al. (2015) provided a way to do so in a non-parametric way using random regression forests. Inspired by the Recursive Feature Elimination (RFE) algorithm by Guyon, Weston, Barnhill, and Vapnik (2002), they developed an backward grouped elimination algorithm that allows researchers to assess grouped variable importance. The algorithm starts with training a random forest model, whereafter the error is calculated based on a validation sample (Gregorutti, Michel, & Saint-Pierre, 2013). Next, the grouped variable importance measure (i.e. permutation importance measure) is computed by minimizing the validation error. Finally, the least important group is discarded. These steps are repeated until one most important group remains.

3.3.3 Model evaluation

Before any analyses were conducted, the data was randomly divided in two: 70 percent for training (i.e. 56,410 observations) and 30 percent for testing purposes (i.e. 24,176 observations). The experiments were performed on the training data. The evaluation of the methods was based on the total variance explained (R^2). R^2 is a metric that can be used to measure to what extent a model fits the data; presenting the proportion of variance that the model explains (Field, 2009). This measure was selected because both the relative weights method and random regression forests outputted it. Only the between-group variable importance measure of Gregorutti et al. (2015) lacks a model evaluation measure. Therefore, it was decided to base the performance of the between-category regression forests on the R^2 of a random regression forest model that included all explanatory variables.

The evaluation method was used for two purposes: comparing the performance of the two algorithms, and checking the individual algorithms for overfitting. Per experiment the most important predictors were selected amongst all regressors based on the training data. This selection was based on the procedure as outlined by Guyon, Gunn, Nikravesh, and Zadeh (2008). For every variable importance metric, ordered in descending order and denoted as $\mathrm{M}(\mathrm{x}_j)$, the difference $\Delta M(x_j) = M(x_j) - M(x_{j+1})$ was calculated. This was done for the rankings produced by the relative weights method or LMG, and the random regression forests. Based on these two lists of differences, the optimal cut-off point was determined, i.e. the smaller the difference between two ranked variables (e.g. most important and second most important variable) is, the less suitable cut-off point will be. Notably, this procedure is mostly used in the context of one single method that assesses variable importance. Because the two methods sometimes differed in the cut-off points they presented, the selected cut-off points were considered not perfectly reliable. Based on the cut-off points, the selected set of most important predictors of each method was regressed on the outcome variable two times; once on the training and once on the test set. This means that the comparison of algorithm performance on the training and test set was based on different sets of variables, in case the two methods presented different top ranked variables.

3.3.4 Hyper-parameter optimization

Since the linear regressions and consequent bootstrapping had no adjustable hyper-parameters, no hyper-parameter optimization was possible. Random regression forests, on the other hand, had several optimizable parameters. Due to time and resource feasibility constraints, it unfortunately was not possible to optimize the parameters of all random regression forests used in this study. Liaw and Wiener (2002) suggested that this is no very serious problem, as variable importance measures turn out to be relatively stable for different combinations of hyper-parameters. Liaw and Wiener (2002) and Oshiro, Perez, and Baranauskas (2012) noted several 'best practices' for parameter tuning. For example, Oshiro et al. (2012) explained that *ntree* (i.e. the total number of trees per random forest) would best lie between 64 and 128. Liaw and Wiener (2002) reported that the optimal value for *mtry* (i.e. the number of variables tried at every node) is best investigated by trying the default (*p*/3), *p*/6 and *p*/1.5, where *p* is the number of variables. Based on these best practices, it was decided to use 96 and *p/*3 as respectively the values for *ntree* and *mtry*.

To investigate the validity of Liaw and Wiener (2002) their claim about the stability of variable importance metrics and the consequent ranking, the hyper-parameter tuning procedure was performed once. For this, the attitudes and behaviors variable category was selected. Performing 5-fold cross-validation and grid-search, several values for *ntree* (64, 96, and 128) and *mtry* (*p*/3, *p*/6, and *p*/1.5) were tried, and the best combination was selected based on root MSE. As a result, the nine combinations were compared in terms of R^2 and differences in the variable ranking.

3.3.5 Experiments

The first experiment concerned the examination of variable importance within the five variable categories: attitudes and behaviors (5 variables), job characteristics (24 variables), social characteristics

(6 variables), organizational characteristics (25 variables, including dummy reference categories), and personal characteristics (33 variables, including dummy reference categories). In total, ten rankings were produced; five rankings based on the relative weights method and five based on random regression forests. The second experiment was performed to uncover variable category importance. The relative weights and random regression forest method both produced one ranking.

Table 2

Overview of experiments performed in this study

#	Experiment	Method 1	Method 2
1	Variable importance within variable categories		
1a	Behavioral and attitudinal characteristics – Burnout complaints	Relative weights	Random regression forest
1b	Job characteristics – Burnout complaints	Relative weights	Random regression forest
1c	Social characteristics – Burnout complaints	Relative weights	Random regression forest
1d	Organizational characteristics – Burnout complaints	Relative weights	Random regression forest
1e	Personal characteristics – Burnout complaints	Relative weights	Random regression forest
2	Variable importance between variable categories		
-	Behavioral and attitudinal, job, social, organizational and personal characteristics – Burnout complaints	LMG	Random regression forest

4. Results

The results of the two experiments will be explained below. Due to space constraints and only in case the number of independent variables in a variable category exceeds six, only the most important predictors will be reported in the in-text tables. The complete tables can be found in Appendix C. For readability, the algorithms' output metrics will not explicitly be stated in-text.

4.1 Experiment 1: Determining variable importance within variable categories

For each of the five variable categories the most important predictors of burnout complaints are determined. In general, each section starts with a comparison of the rankings as provided by the two algorithms, whereafter the regressors' relationships with burnout complaints are elaborated on. Each section ends with an assessment of model performance.

4.1.1 Behavioral and attitudinal characteristics and burnout complaints

The five behavioral or attitudinal characteristics were only partly ranked in the same way by the relative weights and random regression forest method, as showed in Table 3a. More specifically, the top three differed for both methods, while the ranking for the least important variables was the same. Even though the relative contribution of employability, job satisfaction and satisfaction with working conditions turned out to be different and the overall ranking of the variables is somewhat less robust, both methods suggested that these three variables are vital in predicting burnout complaints. In terms of ranking comparison, the relative weights method showed that employability is by far the most important predictor of burnout complaints, accounting for almost one third of the explained variance. The second, third and fourth ranked variables explained roughly the same percentage of variance. In contrast, the random regression forest revealed that satisfaction with working conditions is the most important predictor. Besides, it considered employability and turnover intention as far less important variables than the parametric method implied. In terms of interpretation, the results suggested that employees with high employability and low turnover intention will be less likely to experience burnout complaints than employees who have less employability and high turnover. Employees who are satisfied with their job and its accompanying working conditions seem to be probable of dealing with burnout complaints better than less satisfied employees.

Based on Table D1 and Figure D1a and D1b (see Appendix D for clarification), the top four of most important predictors was used to evaluate the performance of the two algorithms. The results showed that the random regression forest has a superior performance over the relative weights method, with R^2-scores exceeding those of the relative weights with at least 4 percent. Both methods did not suffer from overfitting too much, as the differences between training and test set performance were negligible.

The results in Table 3b showed that the rankings, importance metrics and performance are relatively stable for different values of *ntree* and *mtry*. Moreover, although a five-fold cross-validation procedure showed that the model where *mtree* equals 3 and *ntree* equals 64 was preferable based on root MSE, the differences between the best combination of parameters and the rest were trivial.

Table 3a

Within variable category variable importance: Behaviors and attitudes

	Relative weights					Random regression forest		
Ranking	Variable	s	%	LLCI	ULCI	Variable	s	% Increase in MSE
1	Employability (-)	*	33.12	31.41	34.78	Satisfaction with working conditions	*	0.189
2	Satisfaction with working conditions (-)	*	23.88	22.46	25.33	Job satisfaction	*	0.169
3	Job satisfaction (-)	*	21.63	20.34	23.03	Employability	*	0.136
4	Turnover intention (+)	*	20.62	19.25	22.03	Turnover intention	*	0.079
5	Under-or overqualification (-)		0.774	0.51	1.03	Under-or overqualification		0.019
OP training set	R^2 = 20.91%					R^2 = 28.77%		
PBP training set	R^2 = 20.22%					R^2 = 28.76%		
PBP test set	R^2 = 21.12%					R^2 = 28.07%		

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); s = Selected for model evaluation; * = Selected as most important variable; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages); MSE = Mean Squared Error.

Table 3b

Ranking, % Increase in MSE and R^2 for different values of ntree and mtry

Ranking	Variable	ntree 64 mtry 1	ntree 64 mtry 2	ntree 64 mtry 3	ntree 96 mtry 1	ntree 96 Mtry 2	ntree 96 mtry 3	ntree 128 mtry 1	ntree 128 mtry 2	ntree 128 mtry 3
1	Employability	0.187	0.246	0.266	0.190	0.240	0.266	0.186	0.244	0.269
2	Satisfaction with working conditions	0.172	0.232	0.251	0.169	0.220	0.252	0.173	0.230	0.250
3	Job satisfaction	0.136	0.165	0.176	0.136	0.163	0.173	0.138	0.164	0.174
4	Turnover intention	0.085	0.107	0.118	0.079	0.106	0.115	0.083	0.107	0.116
5	Under-or overqualification	0.020	0.021	0.026	0.019	0.022	0.023	0.019	0.023	0.023
	R^2	28.84	28.32	27.70	28.77	28.38	27.05	28.90	28.43	27.18

Note. R^2 = Total variance explained (in percentages). MSE = Mean Squared Error.

4.1.2 Job characteristics and burnout complaints

The ten best predicting job characteristics were reported in Table 4. Both methods clearly suggested that workload is the most important job characteristic in predicting burnout complaints. Even though the two methods presented a totally different ranking for the rank 2 to 10, some interesting similarities are apparent. The parametric and the nonparametric method both showed that job control is an important determinant of burnout complaints. In addition, the three types of physical demands are almost all present in the both rankings, with working in unusual body positions being in the top five of both methods.

The differences between the two rankings became particularly evident when considering screen work, job insecurity, demotion, and working overtime. Firstly, the random regression forest method regarded the number of hours working in front of a screen as the third most important job characteristic, while the relative weights method reported a much lower ranking. Secondly, the relative weights method suggested that job insecurity is the second most important predictor of burnout complaints, whereas the random regression forest ranked it far less important. Thirdly, an interesting difference in the two rankings can be found in the rank of demotion; it came in at the ninth place in the relative weights method's ranking, while the random regression forest ranked it last. Finally and in a similar vein as the previous, the relative weights method ranked working overtime seventh and the random regression method as seventeenth.

The results showed that people who are dealing with high workloads, high job insecurity, low job control, high cognitive job demands and high physical job demands are more likely to experience burnout complaints. Employees who are working long hours before screens, deal with long working hours or have gone through a demotion are also more vulnerable for work stress and burnout.

The overall performance of the relative weights method and random regression forest on the training set was quite similar. In contrast, clear differences in performance were noticeable when regressing the five best performing variables on both the training and test set (see Table D2, Figure D2a and Figure D2b in Appendix D for clarification), with the relative weights method outperforming the random regression forest. The differences between performance on training and test set are rather small, which suggests that the rankings were not influenced by overfitting to a large extent.

Table 4
Within variable category variable importance: Job characteristics (shortened)

	Relative weights					Random regression forest		
Ranking	Variable	s	%	LLCI	ULCI	Variable	s	% Increase in MSE
1	Workload (+)	*	40.43	38.72	42.09	Workload	*	0.255
2	Job insecurity (+)	*	13.16	11.92	14.38	Job control	*	0.134
3	Job control (-)	*	9.63	8.68	10.58	Screen work	*	0.111
4	Cognitive job demands (+)	*	9.36	8.53	10.21	PJD – Unusual body positions	*	0.091
5	PJD – Unusual body positions (+)	*	5.43	4.79	6.12	Working hours	*	0.081
6	Screen work (+)		3.71	3.21	4.26	Job insecurity		0.066
7	Working overtime (+)		3.60	3.10	4.10	Cognitive job demands		0.060
8	Working hours (+)		2.80	2.38	3.23	Working at home		0.050
9	Demotion (+)		1.64	1.17	2.16	PJD – Environmental		0.050
10	PJD – Environmental (+)		1.58	1.26	2.02	PJD – Heavy loads		0.041
OP training set	R^2 = 20.91%					R^2 = 20.56%		
PBP training set	R^2 = 19.70%					R^2 = 17.38%		
PBP test set	R^2 = 19.26%					R^2 = 16.55%		

Note. OP = Overall performance; PBP = Performance Best Predictors; PJD = Physical job demands; % = Percentage (i.e. percentage of total variance explained); s = Selected for model evaluation; * = Selected as most important variable; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages), MSE = Mean Squared Error.

4.1.3 Social characteristics and burnout complaints

As reported in Table 5, both the relative weights and random regression forest method indicated that all variables but physical violence play a significant role in predicting burnout complaints. Remarkably, the two methods presented a similar ranking and an almost equal R^2-score on the training set. Supervisor support, intra-personal conflict and intimidation play the most principal roles in predicting burnout complaints. This means that employees who experience low supervisory support, high levels of inter-personal conflict or significant intimidation at the workplace will be relatively prone to burnout complaints.

Based on a selection procedure to be found in Appendix D (Table D3, Figure D3a and Figure D3b), it was decided to use the top four social characteristics to evaluate the performance of the two methods. The performance of both methods on the training and test set were again quite similar, with the random regression forest performing slightly better on all datasets. The results showed that the methods and consequent rankings did not suffer from overfitting too much.

Table 5

Within variable category variable importance: Social characteristics

	Relative weights					Random regression forest		
Ranking	Variable	s	%	LLCI	ULCI	Variable	s	% Increase in MSE
1	Supervisor support (-)	*	31.76	29.77	33.69	Supervisor support	*	0.1378
2	Intra-personal conflict (+)	*	22.72	20.94	24.44	Intra-personal conflict	*	0.0648
3	Intimidation (+)	*	20.21	18.69	21.82	Intimidation	*	0.0643
4	Bullying (+)	*	19.05	17.4	20.76	Bullying	*	0.0501
5	Colleague support (-)		4.89	44.17	5.66	Colleague support		0.0290
6	Physical violence (+)		1.37	1.02	1.82	Physical violence		0.0000
OP training set	R^2 = 17.75%					R^2 = 18.13%		
PBP training set	R^2 = 18.03%					R^2 = 18.46%		
PBP test set	R^2 = 17.59%					R^2 = 18.05%		

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); s = Selected for model evaluation; * = Selected as most important variable; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages); MSE = Mean Squared Error.

4.1.4 Organizational characteristics and burnout complaints

As outlined in Table 6, the results showed that the dissatisfaction with OHS practices is the most important organizational determinant of burnout complaints. It is interesting to see that the decrease in importance metric when converting from the first (i.e. OHS) to the second ranked variable (respectively, education sector and no organizational change) was much stronger in the relative weight's ranking than in the random forest's. Although way less important, organizational changes, organizational size and sector also seem to play a key role in predicting burnout complaints. The two rankings showed roughly the same pattern. In both rankings, reorganizations and organizational downsizing with and without forced layoffs rated fairly high. Interestingly, organizational size rates much higher in the random regression forest compared to the relative weights method, whereas the education sector rated higher in the relative weights method's ranking.

It seems that organizations that fails to effectively implement OHS practices at the workplace, deal with a workforce with more burnout complaints. Additionally, employees in the health-care and education sector are way more likely to experience burnout complains compared to employees in the business services and real-estate sector. Organizations going through a reorganization, downsizing, outsourcing or automation change seem to have a workforce with more burnout complaints compared to organizations who do not go through an organizational change. In addition, employees in larger organizations appear to cope with burnout complaints more often than employees in smaller organizations.

The relative weights method's overall performance on the training was better than the random regression forest's. Nevertheless, the performance of the random regression forest on the training and test set when regressing the three most important variables was superior to the performance of the relative weights method (see appendix D, Table D4, Figure D4a and D4b). Notably, in case of both methods, the performance on the training set was worse than the performance on the test set. This means that the methods did not show clear signs of overfitting the data.

Table 6

Within variable category variable importance: Organizational characteristics (shortened)

	Relative weights					Random regression forest		
Ranking	Variable	s	%	LLCI	ULCI	Variable	s	% Increase in MSE
1	OHS (-)	*	83.07	81.43	84.25	OHS	*	0.481
2	Sector: education (+)	*	2.57	1.91	3.29	Change: no change	*	0.074
3	Change: reorganization (+)	*	2.46	1.93	3.05	Organizational size	*	0.070
4	Change: downsizing with forced layoffs (+)		2.01	1.50	2.58	Change: reorganization		0.046
5	Change: downsizing without forced layoffs (+)		1.73	1.24	2.23	Change: downsizing without forced layoffs		0.027
6	Change: outsourcing (+)		1.72	1.27	2.28	Change: downsizing with forced layoffs		0.026
7	Change: automation (+)		1.15	0.82	1.59	Sector: education		0.021
8	Organizational size (+)		0.91	0.84	1.01	Sector: health-care		0.016
9	Sector: health-care (+)		0.63	0.51	0.81	Sector: retail		0.013
OP training set	R^2 = 16.42%					R^2 = 14.42		
PBP training set	R^2 = 14.66%					R^2 = 15.05		
PBP test set	R^2 = 15.38%					R^2 = 15.66		

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); s = Selected for model evaluation; * = Selected as most important variable; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages); Reference category '(Organizational) change' dummy-variable = No organizational change; Reference category 'Sector' dummy-variable = Business services & Real-estate; OHS = Occupational Health and Safety.

4.1.5 Personal characteristics and burnout complaints

As displayed in Table 7, general health most strongly contributed to the total variance explained. Both methods pointed out that the province in which an employee lives does not contribute (strongly) to the total variance explained. Besides the first ranked variable and the unimportance of the province variable, the two rankings differ in many ways. For example, age came in as second most important variable in the random regression forest's ranking, while it ranked much lower in the relative weights ranking. The differences became especially clear when considering the differences in metrics of the first ranked variable, general health, and age. The random regression forest showed that age is approximately half as important as general health. The importance of age in the relative weights ranking is almost negligible, as it accounts for 0.60 percent of the explained variance as compared to general health's 82.11 percent. In the same line, the difference between the first and second ranked variable in the random regression forest was a lot smaller (i.e. 0.407 and 0.197) in comparison to the first and second ranked as indicated by the relative weights method (i.e. 83.11 percent and 4.97 percent). The relative weights method suggested that occupational accidents play a central role in forecasting burnout complaints, while the random regression forest did not do so at all. Furthermore, some interesting results were outputted with regard to ethnicity, marital status and household composition. Both method indicated that having a migration background, living alone, being married and having children are important determinants of burnout complaints.

The linear model suggested that employees with a negative perception of their own general health or who experienced an occupational accident, are more likely to experience burnout complaints. Employees with a migration background are more vulnerable for burnout complaints compared to people who do not have such a background. The same is true for people who are living alone, are unmarried or do not have children. Remarkably, the results showed that older employees experience less burnout complaints than their younger counterparts, whereas people with a higher organizational tenure deal with less burnout complaints compared to starters.

In terms of overall performance, the relative weights method proved to be superior to the random regression forest. However, when regressing the best four predictors on the dependent variable, the random regression forest outperformed the relative weights method on both the training and test set. For a clarification about how the best predictors were selected, see Appendix D (Table D5, Figure D5a and D5b).

Table 7

Within variable category variable importance: Personal characteristics (shortened)

	Relative weights					Random regression forest		
Ranking	Variable	s	%	LLCI	ULCI	Variable	s	% Increase in MSE
1	General health (-)	*	83.11	81.13	84.48	General health	*	0.407
2	Occupational accident occurrence (+)	*	4.97	3.84	6,21	Age	*	0.197
3	Household: one-person (+)	*	3.69	2.97	4.55	Organizational tenure	*	0.098
4	Ethnicity: non-western migration background (+)	*	2.03	1.41	2.70	Marital status: married	*	0.090
5	Household: unmarried without children (+)		1.80	1.19	2.36	Marital status: never married		0.061
6	Age (-)		0.60	.054	0.70	Household: married with children		0.056
7	Organizational tenure (+)		0.52	0.34	0.76	Educational level		0.052
OP training set	R^2 = 15.03%					R^2 = 12.68%		
PBP training set	R^2 = 14.49%					R^2 = 15.27%		
PBP test set	R^2 = 14.32%					R^2 = 15.00%		

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); s = Selected for model evaluation; * = Selected as most important variable; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages); Reference category 'Household' dummy-variable = Married with children; Reference category 'Marital status' dummy-variable = Married; Reference category 'Province' dummy-variable = Noord-Holland; Reference category 'Ethnicity' dummy-variable = No migration background; MSE = Mean Squared Error.

4.1 Experiment 2: Determining variable importance between variable categories

The second experiment assessed of the importance of the variable categories: personal characteristics, organizational characteristics, job characteristics, social characteristics, and work attitudes and behaviors, and is summarized in Table 8. The two rankings showed some similarities but mostly differences. In terms of similarities, both methods ranked personal and organizational characteristics, respectively, fourth and fifth. A first difference between the two rankings was found in the top three; there were no matches. For example, job characteristics were ranked first by the relative weights method, while it was ranked third by the random regression forest. Attitudes and behaviors were by far the most essential variable category according to the random regression forest, while this was not the case according to the LMG method. In addition, the decrease in importance when swapping from the first ranked to the second ranked variable category was much more serious in the random regression forest (0.055 to 0.012) than it was in the LMG method (25.53 to 22.83).

The overall performance of the random regression forest was slightly better than that of the LMG method. Similarly, the random regression forest outperformed the LMG method on the training and test set when regressing the thee best predicting variable categories (see Appendix D, Table D6, Figure D6a and D6b). The performance on the training set was almost equal to the performance on the test set for both methods. This implies that the two methods did not overfit the data.

Table 8
Between variable category importance

	LMG					Random regression forest		
Ranking	Variable category	s	%	LLCI	ULCI	Variable	s	Importance metric
1	Job characteristics	*	25.53	24.65	26.39	Attitudes and behaviors	*	0.055
2	Attitudes and behaviors	*	22.83	21.99	23.51	Social characteristics	*	0.012
3	Social characteristics	*	18.94	18.07	19.83	Job characteristics	*	0.011
4	Personal characteristics		18.11	17.28	19.03	Personal characteristics		0.005
5	Organizational characteristics		14.60	13.99	15.35	Organizational characteristics		0.004
OP training set	R^2 = 42.36%					R^2 = 44.37%		
PBP training set	R^2 = 35.83%					R^2 = 38.75%		
PBP test set	R^2 = 36.96%					R^2 = 38.40%		

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); s = Selected for model evaluation; * = Selected as most important variable; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; R^2 = Total variance explained (in percentages).

5. Conclusion and discussion

5.1 Conclusion

The purpose of this study was twofold. It aimed at empirically discovering the most important predictors of burnout complaints in the Dutch working population within five categories: personal characteristics, organizational characteristics, job characteristics, social characteristics, and work attitudes and behaviors. Furthermore, this study attempted to discover an importance ranking of these variable categories. The CBS randomly sampled the Dutch working population and gathered data about 80,586 respondents. In turn, this data was used for three types of variable importance analysis: Johnson's (2000) relative weight method, Breiman's (2001) random regression forests, and Kruskal's (1987) LMG. By means of this study, interesting new insights about the relative variable importance of a wide variety of variables was obtained. In addition, both this study's exploratory research design and relatively advanced methodological approach for assessing variable importance contribute to the literature by providing a fruitful basis for future deductive and inductive research.

Although the random regression forest and relative weights method not always presented a matching variable ranking, some general findings from the first experiment can be presented. The results showed that satisfaction with working conditions, employability, and job satisfaction are the most important attitudinal or behavioral predictors of burnout complaints. Workload, job insecurity, job control, physical job demands, cognitive job demands, long working hours and hours worked in front of a screen were determined to be the most essential job characteristics in the estimation of burnout complaints. With respect to social characteristics, supervisor support, intra-personal conflict, intimidation and bullying are, respectively, ranked first, second, third and fourth. Although satisfaction with OHS practices was by far the most vital organizational determinant, sector (e.g. education, health-care), organizational changes (e.g. reorganization, downsizing, outsourcing, automation) and organizational size seem to be principal antecedents of burnout complaints. The analyses suggested that general health, age, education level, occupational accident occurrence, household composition (e.g. one-person household, unmarried without children), marital status (e.g. never married, married) and ethnicity (e.g. non-western migration background) all play an important role in predicting burnout complaints. The present study's results do not indicate that one method was superior to the other. Instead, the relative weights method outperformed the random regression method in some analyses, while the random regression forest performed better in others. Sometimes the superiority even shifted from the one to the other; with overall performance on the one hand and top predictors performance on the training and test set on the other hand. In general, both the random regression forest and relative weights method did not really seem to overfit the data.

The second experiment was concerned with the variable category importance and eventually produced two rankings. The LMG method indicated that job characteristics, attitudes and behaviors and social characteristics respectively were the first, second and third most important categories; the random regression forest ranked attitudes and behaviors first and social and job characteristics respectively second and third. Both methods indicated that personal and organizational characteristics were the least important. The overall performance of the random regression forest was better than the LMG method's. The random regression method also outperformed the LMG method on the training and test set when regressing the three most important variable categories on the dependent variable. Both methods did not show clear signs of overfitting the data.

5.2 Discussion

It proved rather hard to compare the present study's results to evidence provided by the scientific community, as, to the knowledge of the author, no comparable study has yet been conducted. One or a combination of the following aspects underlie reason of their incomparability with the current study: i) the study only covers a very small segment of the total working population (e.g. nurses), ii) the study's samples size are inappropriately small (i.e. seldom more than 500 respondents), iii) the adopted relative importance measures are too inaccurate (e.g. standardized regression coefficients), and iv) the study only considers a very small number of variables (e.g. solely considering job characteristics). Although a comparison is hard to make, the literature has provided plenty of rationale of why and how variables relate to burnout and work stress. Hence, the most important variables according to the present findings are related to the literature in the next five paragraphs. Hereafter, the between-category rankings are discussed as well as the performance of the algorithms.

First off, the literature is rather uniform with respect to the predictive power and type of relationship employability (Gene M Alarcon, 2011; Aybas, Elmas, & Dundar, 2015; De Cuyper, Raeder, Van der Heijden, & Wittekind, 2012), satisfaction with working conditions (Maslach et al., 2001) and job satisfaction (Faragher et al., 2005) have with burnout complaints. This study is no exception and also found out that all above variables have a strong negative correlation with burnout complaints.

Second, a number of the most important predictors as found by the current study were found in the review about best predictors of burnout by Peeters et al. (2013): workload, working hours and job control. They pointed out that both workload and working hours have the potential to increase burnout complaints, while job control does the exact opposite. Interestingly, physical job demands, cognitive job demands and screen work were identified as essential predictors of burnout complaints. Although Nahrgang et al. (2011) and Fragoso et al. (2016) provide a rationale of why job demands are more important than job resources such as job control, feedback, task variety, it is surprising to see that especially these job demands play such an important role in predicting burnout complaints. Even more remarkable is the substantial importance of screen work in forecasting burnout complaints. While some scholars in the field of ergonomics have discovered a positive association between screen work and burnout complaints (Korpinen, Pääkkönen, & Gobba, 2013, 2016), a solid basis of evidence as well as a theoretical explanation remain absent.

Third, the most important social characteristics found in this study are overall consistent with existing literature. For instance, meta-analytical findings of Nahrgang et al. (2011) and longitudinal results of De Beer, Pienaar, and Rothmann Jr (2013) pointed out that supervisor support is the most essential social characteristic in the prediction of burnout complaints. Some (biased) primary studies (see Section 2.3) reported that conflictive interactions (Garrosa et al., 2008), bullying (Hauge et al., 2010), intimidation (Aybas et al., 2015), and interpersonal hostile behaviors in general (Ben-Zur & Yagil, 2005) have the potential to disastrously damage employee health (e.g. cause burnout complaints).

Fourth, satisfaction with OHS practices proved to be the most important organizational factor in the estimation of burnout complaints. This seems rather logical, as the very purpose of OHS practices is to improve occupational health (e.g. providing more job control) and diminish occupational health risks (e.g. lower the number of job stressors) (Frick, Jensen, Quinlan, & Wilthagen, 2000). Another reason for this unsurprising finding lies in OHS' the long history, and the many theoretical and practical developments that have occurred in the field (Peeters et al., 2013). The results of the analyses indicated that working in the education and health-care sector are key predictors of burnout. This does also not come as a surprise, as it is a well-known that people in contact professions are more vulnerable to burnout than people who are not (Bakker, Van Der Zee, Lewig, & Dollard, 2006; Maslach, 2001; Schaufeli, Leiter, et al., 2009). Since most professions in the education and health-care sector require a great deal of human interaction, they could be classified as risk-sectors with respect to burnout complaints (Adriaenssens et al., 2015; Kokkinos, 2007; Van Der Ploeg & Kleber, 2003; Van Droogenbroeck, Spruyt, & Vanroelen, 2014). In addition, Brown and Quick (2013), Crawford et al. (2010) and Datta, Guthrie, Basuil, and Pandey (2010) provide empirical evidence and a theoretical rationale with respect to the reasons why organizational changes are strong determinants of burnout complaints. They explain that especially reorganizations, downsizing and outsourcing often cause job insecurity, lower job involvement and conflicts; all indicators of higher burnout risk. Furthermore, the strong predictive power of organizational size came as a surprise, as no studies, to the author of the present study's knowledge, have found a significant positive (or for that matter, negative) correlation between organizational size and burnout complaints.

Fifth, the present study has identified some interesting important personal characteristics in relation to burnout complaints. Corroborating with past studies (Ahola, Salminen, Toppinen-Tanner, Koskinen, & Väänänen, 2013; Farsi, Nayeri, & Sajadi, 2013; Khamisa, Peltzer, & Oldenburg, 2013), the current findings suggested that general health and occupational accidents are the most essential personal predictors of burnout complaints. Straightforward theoretical reasons underlie the findings. Burnout is simply defined as mental health problem (Maslach, 2001), and burnout pretty much comes down to a mental occupational accident (Ahola et al., 2013). It, however, is interesting to see that the question as used in this study collected data on general – physical and mental – health, and does not specifically refer to mental health. The results further indicated that both age and organizational tenure play an important part in the prevalence of burnout complaints, respectively having a negative and positive association with it. Although no uniformity exists on the relationship age and tenure have with burnout complaints, this finding is still rather surprising. Most studies report either a positive or a negative relationship between burnout complaints and *both* variables (e.g. Lasalvia et al., 2009; Thomas, Kohli, & Choi, 2014). The reason for this consistency in the literature is that a strong positive correlation exists between age and organizational tenure. Older people usually have a longer organizational tenure than younger people. Finally, the findings corroborate with the work of Smulders, Houtman, Rijssen, and Mol (2013), who reviewed a handful of studies that theorized about the importance of personal determinants such as ethnicity, household composition and marital status in the prediction of burnout. Both their research and the present study indicate that the employees having a migration background or living alone are more likely to experience burnout complaints compared to people who do not have such a background or household composition.

Sixth, while numerous studies have proven the importance of all five variable categories (e.g. Nahrgang et al., 2011; Peeters et al., 2013; Schaufeli, Leiter, et al., 2009; Smulders et al., 2013), it seems that so far no study has attempted to rank variable categories so comprehensively as this study did. This is why no real theoretical explanation for the rankings can be provided. Considering the fact that attitudes and behaviors and social characteristics categories contain much fewer variables compared to personal, organization and job characteristics, it seems that in particular these two variable categories play an exceptionally important role.

Seventh, with respect to the algorithm performance, the random regression forests did not consistently outperform the relative weights and LMG method. Instead, the results suggested that in some models the random regression forest acted as a better method, while in others the linear models proved to be superior. Based on overall model performance, the random regression forest outperformed the relative weights method in the attitudes and behaviors and social characteristics category, and outclassed the LMG method on between-category importance. The linear model explained more variance in the organizational, job and personal characteristics categories than the random regression forest did. Besides performance in terms of explained variance, the rankings that the two methods presented were not uniform. As no study has compared parametric and non-parametric relative importance methods with each other in terms of performance and rankings on such a large scale, it is unfortunately not possible to present any concrete reasons for these differences. Still, some rationale for why the random regression forests' rankings should be preferred over the linear methods is provided. First, as there exists strong evidence about the potential non-linear relationships between work characteristics and employee health, the assumption of normality that the relative weights and LMG method make could possibly be faulty (Karanika-Murray & Cox, 2010). In addition, the parametric methods did not incorporate interactions between independent variables, while the literature clearly stresses the significance of moderators in the prediction of burnout (Fila, Purl, & Griffeth, 2017; Halbesleben & Buckley, 2004; Turnipseed, 1994; Zapf, 2002). As random regression forests consider the non-linearity of the data and incorporates interaction effects by nature (Breiman, 2001; Grömping, 2009), it would be most logical to prefer the random regression forest over the relative weight and LMG method (Pretnar, 2015).

5.3 Limitations and future research directions

Several limitations and consequent research directions should be noted. Firstly, the present study's cross-sectional research design might be a potential limitation, as no causality claims could be made (Lazarus, 2000). As this study aimed at 'predicting' and 'forecasting' burnout complaints, one might have assumed that the regressors in fact *cause* burnout complaints. Even though many variables do have causal relationships with burnout complaints (Hakanen, Schaufeli, & Ahola, 2008), there is also evidence for reversed and reciprocal relationships (T. Taris et al., 2013). For example, De Beer et al. (2013) and Demerouti, Bakker, and Bulters (2004), respectively, found out that supervisor support has a reversed causal and work pressure has a reciprocal relationship with burnout. Moreover, this study did not explicitly incorporate interactions or check for indirect effects, while the literature noticeably indicates that moderators and mediators play an important role in understanding the concept of burnout (Bakker & Demerouti, 2007). For instance, meta-analytical evidence from the nursing sector indicates that the relationship between shift work and burnout is positively moderated by age (Vargas, Cañadas, Aguayo, Fernández, & Emilia, 2014). Additionally, a recent meta-analysis by Fila et al. (2017) suggested that occupations, gender and nationality seem to act as important moderating variables in the relationship between job demands and burnout complaints. Besides evidence about moderating factors, the JD-R model indicates that job demands can indirectly influence burnout complaints (Bakker & Demerouti, 2007). Future studies are therefore encouraged to incorporate longitudinal, interaction and indirect effects in their studies, so that true insights are gained about what factors to consider in combatting work stress and burnout in the workplace.

Secondly, the categorization of the variables might have been biased, as the JD-R model was not strictly followed. Although adopting a novel approach might have been logical (e.g. no existing theoretical model covers such a large variety of independent variables) and could function as valuable starting point for future (inductive) research, it might be interesting to hold on to the more formal classification as delineated in the JD-R model (e.g. social and job characteristics would be mixed into different categories) (Bakker & Demerouti, 2007). While this study inductively researched a wide variety of potential antecedents, there still exist potentially important variables such as work-family conflict (Vargas et al., 2014), emotional job demands (Peeters et al., 2013) and occupations (T. Taris et al., 2013) that would be valuable additions to further inductive investigations. Such research could also benefit from adopting the multi-dimensional burnout approach. Choosing a unidimensional approach just like the current study may be too simplistic, as evidence suggests that various work and personal factors relate differently to the distinctive dimensions of burnout: emotional exhaustion, depersonalization and reduced personal accomplishment (Maslach, 2003; T. Taris et al., 2013; Toppinen-Tanner, 2011).

Thirdly, by means of a comparable inductive research design, scholars are also urged to investigate variable importance in predicting other potentially disastrous work outcomes such as turnover intention (Cohen, Blake, & Goodman, 2016), sickness absence (Schaufeli, Bakker, et al., 2009) and occupational accidents (Nahrgang et al., 2011). For this, it is highly recommended to use a large and highly representative dataset (e.g. national census data, CBS data, and open data) and narrowly define one's target population.

Fourthly, although the present study adopted far more advanced and appropriate measures of variable importance compared to the bulk of the existing HRM and occupational health research (Karanika-Murray & Cox, 2010; Tonidandel & LeBreton, 2011), certain weaknesses in research methodology arose due to serious resource feasibility issues. Post-hoc literature search suggested that parameter tuning could be crucial for the robustness and stability of the random forests' imputation results and variable importance metrics (Grömping, 2009; Stekhoven & Bühlmann, 2012). The search also showed that classical regression forests based on CART trees are sometimes biased in case regressors correlate with each other, or strongly differ in their measurement scales (Grömping, 2009; Hothorn, Hornik, & Zeileis, 2006; Strobl, Boulesteix, Zeileis, & Hothorn, 2007). As continuous, ordinal and binary variables were used in this study and many variables were inter-correlated, the results may have been biased to some extent. Luckily, Hapfelmeier and Ulm (2013) showed this bias only exists in the absolute importance scores (i.e. underestimating the importance), does not affect the relative ranking of the variables, and does therefore not seriously damage the present study's interferences. Still, to be sure of unbiased results, researchers are encouraged to build random forests from conditional inference trees (Hothorn et al., 2006), use the permutation-based MSE reduction as main variable importance metric (Strobl et al., 2007), and tune the random forest algorithm (Grömping, 2009).

Fifthly, besides the recommendations for non-linear variable importance assessment, scholars are also encouraged to adopt even more reliable linear variable importance methods such as LMG, PMVD or dominance analysis, if resources allow it (Grömping, 2015). As a less computationally demanding alternative, researchers are advised to use machine learning method called sparsity-oriented importance learning (SOIL), which performs at least as well as LMG and PMVD (Ye et al., 2016). SOIL allows studies to deal with high-dimensional contexts, outputs highly robust and reliable variable importance metrics, and conveniently scales variable importance on an absolute scale from 0 to 1.

A final interesting research direction concerns the operationalization of burnout complaints. While the current study's follows the general trend of continuously operationalizing burnout, future research might benefit from using cut-off points (e.g. low, average, and high) (Toppinen-Tanner, 2011). By switching from a regression to a classification task, the results might be of more use to those who diagnose burnout in real-life (Schaufeli, Leiter, et al., 2009).

5.4 Recommendations for practitioners

Within the aforementioned limitations, some recommendations for practitioners in the field of organizational health and HRM can be stated. Because organizations could have a large impact on an employee wellbeing via individual-level and organization-level interventions (Siu et al., 2014; Van De Voorde, Paauwe, & Van Veldhoven, 2012), practitioners can learn a great deal from the current research findings. By thoroughly studying and effectively implementing the most important variables and variable categories the current study has put forward, professionals in the field could improve their policies, and warrant the cost-efficiency and relevance of interventions (Fragoso et al., 2016; Garrosa et al., 2008; Nahrgang et al., 2011; Pawlowski et al., 2007). For instance, the present study's findings suggest that practitioners would benefit from establishing an employability culture (Nauta, Vianen, Heijden, Dam, & Willemsen, 2009), promoting constructive relationships between supervisors and employees (De Beer et al., 2013), and ensuring satisfactory OHS practices and good workforce health (Peeters et al., 2013). The results also showed that practice needs to consider factors that are underreseached or seem less obvious. For example, it is often ignored that factors such as migration background, marital status and household composition could play a principal role in the prevalence of burnout (Maslach et al., 2001; T. Taris et al., 2013). While job characteristics are more widely researched, even in that field some determinants are often wrongfully overlooked. T. Taris et al. (2013) noted that scholars pay inappropriately much attention to psychological stressors and systematically forget to incorporate physical job demands in their studies.

This study showed that job, social, attitudinal, behavioral, organizational and personal characteristics all play an important part in predicting burnout in the Dutch working population. In particular, the near optimal generalizability and large sample size allowed this study to reliably rank a wide variety of variables in importance. Amongst the most important ones are OHS practices, sector, organizational changes, social support, interpersonal hostile behaviors, employability, satisfaction with the job and working conditions, occupational accidents, general health, age workload, job insecurity, and job control. The current study's exploratory, comprehensive research design could function as an inspiring foundation for both further inductive research in which more certainty can be gained about which factors predict burnout best, and future deductive in which researchers can hypothesize about *why, how* and *when* factors most strongly affect burnout.

Appendix 1: Overview of variable categorizations, definitions, operationalization and relevance

Variable	Category	Definition	Operationalization	Related studies
Employability	Behaviors and attitudes	The extent to which a respondent regards oneself as capable to satisfy the physical and psychological work demands.	*Categorical variable.* A weighted average is calculated over the two questions with four answer categories ranging from totally disagree to totally agree.	G. M. Alarcon et al. (2009); Aybas et al. (2015); De Cuyper et al. (2012)
Turnover intention	Behaviors and attitudes	'A conscious and deliberate willingness to leave the organization' (Tett & Meyer, 1993, p.263).	*Categorical variable.* A weighted average is calculated over three questions with two answer categories no and yes.	De Croon, Sluiter, Blonk, Broersen, and Frings-Dresen (2004)
Under- or over qualification	Behaviors and attitudes	'The situation where individuals have qualifications such as education and skills that exceed job requirements (Erdogan & Bauer, 2009, p.557).	*Categorical variable.* Three categories under-qualification, qualification, and over-qualification.	Navarro, Mas, and Jiménez (2010)
Job satisfaction	Behaviors and attitudes	'An internal state that is expressed by affectively and/or cognitively evaluating and experienced job with some degree of favor or disfavor' (Brief, 1998, p.68).	*Categorical variable.* Five categories ranging from very unsatisfied to very satisfied.	Faragher et al. (2005)
Satisfaction with working conditions	Behaviors and attitudes	The extent to which an employee is satisfied with his or her working conditions.	*Categorical variables.* Five categories ranging from very unsatisfied to very satisfied.	Maslach et al. (2001)
Working time	Job characteristics	The total number of hours an employee works.	*Continuous variable.* Measured per week. Four additional categories are distinguished that contextualize the number of working hours (per week, per month, per year, teaching hours per week). For example, '160 hours' with 'per month' becomes '40 hours per week'.	Van den Heuvel, Geuskens, Hooftman, Koppes, and Van den Bossche (2010)
Shift work	Job characteristics	'The scheduling of work according to a particular time period' (Landy & Conte, 2016, p.382).	*Categorical variable.* Three categories: no, sometimes, regularly.	Costa (2010); Van den Heuvel et al. (2010)
Abnormal work hours	Job characteristics	'Working outside normal or standard hours' (Peeters et al., 2013, p.206)	*Categorical variable.* A weighted average is calculated over the four questions with answer categories no, sometimes and regularly.	Peeters et al. (2013)
Working overtime	Job characteristics	'All work hours that an employee works on top of his/her contractual work hours' (Beckers et al., 2004, p.18).	*Categorical variable.* Three categories: no, sometimes, regularly.	Bannai (2014); Van den Heuvel et al. (2010)
Working at home	Job characteristics	Often associated with teleworking (Bailey & Kurland, 2002). Teleworking is 'an alternative work arrangement in which employees perform tasks elsewhere that are normally done in a primary or central workplace, for at least some portion of their work schedule, using electronic media to interact with others inside and outside the organization' (Gajendran & Harrison, 2007, p.1525).	*Dummy variable.* 0 = No, 1 = Yes.	Peeters et al. (2013)
Managing position	Job characteristics	Whether or not an employee leads or manages subordinates.	*Categorical variable.* 0 = No, 1 = Yes.	Van den Heuvel et al. (2010)
Dangerous work	Job characteristics	Work 'in which a physically and/or psychologically harmful event has some probability of occurring' (Jermier, Gaines, & McIntosh, 1989, p.16)	*Categorical variable.* Three categories: no, sometimes, regularly.	Peeters et al. (2013)
On-call work	Job characteristics	'Work where workers are called to work either between regular hours or during set on-call periods' (Nicol & Botterill, 2004, p.1).	*Categorical variable.* Three categories: no, sometimes, regularly.	Nicol and Botterill (2004)

Workload	Job characteristics	'The amount and pace of work' (Peeters et al., 2013, p.117).	*Categorical variable.* A weighted average is calculated over the six questions with answer categories no, sometimes and regularly.	Peeters et al. (2013); Schaufeli, Bakker, et al. (2009)
Physical demands (Environmental)	Job characteristics	'Work-related tasks that require physical effort' (Van den Tooren, 2011, p.8) Here: physical, biological, chemical (Peeters et al., 2013).	*Categorical variable.* A weighted average is calculated over the six questions with answer categories no, sometimes and regularly.	Johns and Saks (2005)
Physical demands (Heavy loads)	Job characteristics	'Work-related tasks that require physical effort' (Van den Tooren, 2011, p.8). Here: heavy loads such as lifting loads, carrying loads (Peeters et al., 2013).	Categorical variable. Three categories: no, sometimes, regularly.	Johns and Saks (2005)
Physical demands (Unusual or tiring body positions and movements)	Job characteristics	'Work-related tasks that require physical effort' (Van den Tooren, 2011, p.8). Here: unusual or tiring body positions and movements such as standing and repetitive movements (Peeters et al., 2013).	*Categorical variable.* A weighted average is calculated over the two questions with answer categories no, sometimes and regularly.	Johns and Saks (2005)
Cognitive demands	Job characteristics	'Work-related tasks that require cognitive effort (e.g. finding solutions for complex problems)' (Van den Tooren, 2011, p.8)	*Categorical variable*. A weighted average is calculated over the four questions with answer categories never, sometimes, often, always.	Nahrgang et al. (2011); Peeters et al. (2013); Van den Heuvel et al. (2010)
Job control	Job characteristics	'The degree to which employees have a say about activities and the conditions under which they work so that they correspond most closely to their needs and goals' (Peeters et al., 2013, p.172).	*Categorical variable*. A weighted average is calculated over the six questions with answer categories no, sometimes and regularly.	Peeters et al. (2013); Schaufeli, Bakker, et al. (2009); Van den Heuvel et al. (2010)
Screen work	Job characteristics	Working in front of a screen, for example a smartphone, laptop or tablet.	*Continuous variable*. Measured in hours per day.	Korpinen et al. (2013)
Demotion	Job characteristics	'Job transitions to a lower hierarchical level' (Dohmen, Kriechel, & Pfann, 2004, p.201)	*Dummy variable.* 0 = No, 1 = Yes.	Crawford et al. (2010)
Promotion	Job characteristics	'Job transitions to a higher hierarchical level' (Dohmen et al., 2004, p.201)	*Dummy variable.* 0 = No, 1 = Yes.	Josten and Schalk (2010)
Job change	Job characteristics	Whether or not a person's job has changed in the past two years.	*Dummy variable.* 0 = No, 1 = Yes.	Crawford et al. (2010)
Job enlargement	Job characteristics	Horizontal job loading, vertical job loading or both (Ramlall, 2004).	*Dummy variable.* 0 = No, 1 = Yes.	Crawford et al. (2010)
Job insecurity	Job characteristics	'An overall concern about the continued existence of the job in the future' (Sverke, Hellgren, & Näswall, 2002, p.243).	*Categorical variable.* A weighted average is calculated over two questions with two answer categories no and yes.	Crawford et al. (2010); Johns and Saks (2005); Peeters et al. (2013)
Employment contract	Job characteristics	An agreement in which the one party, the employee, commits to be serve the other party, the employer, with salary as a return for a certain period of time (CBS, 2016).	*Dummy variable.* 0 = Fixed contract, 1 = Temporary contract.	Benach, Amable, Muntaner, and Benavides (2002); Van den Heuvel et al. (2010)
Organizational size	Organizational characteristics	The total number of employees working at the company.	*Categorical variable.* Based on Eurostat (2016), a categorical variable is created with categories small (50-), middle (50-250) and large (250+).	-
Sector	Organizational characteristics	The industry the organization is working in.	*Dummy variables.* The ten sectors are included as dummies.	Berry et al. (2012); Morse et al. (2005)

Organizational changes	Organizational characteristics	'A set of behavioral science-based theories, values, strategies, and techniques aimed at the planned change of the Organizational work setting for the purpose of enhancing individual development and improving organizational performance, through the alteration of organizational members' on-the-job behaviors' (Porras & Robertson, 1992, 723).	*Dummy variable.* Eleven dummies; one for every organizational change (e.g. 0 = No organizational changes have occurred, 1 = Reorganization)	Crawford et al. (2010); Johns and Saks (2005)
Occupational Health & Safety (OHS) practices	Organizational characteristics	'A limited number of mandated principles for systematic management of OHS, applicable to all types of employers including the small ones' (Frick et al., 2000, p.3). In this study, the *dis*satisfaction with these practices is considered.	*Continuous variable.* A weighted average is calculated over ten questions with four answer categories.	Peeters et al. (2013)
Burnout complaints	Outcome variable	The mental condition of emotional exhaustion, depersonalization and decreased personal accomplishment at work (Maslach et al., 2001).	*Categorical variable.* A weighted average is calculated over five questions with seven answer categories ranging from never to daily.	
Sex	Personal characteristics	Biological differences between male and female: the visible difference in genitalia, the related difference in procreative function' (Oakley, 1985, p.16).	*Dummy variable.* Female = 0, male = 1.	Bonde (2008); Purvanova and Muros (2010); Van den Heuvel et al. (2010)
Chronological age	Personal characteristics	A respondent's calander age.	*Continuous variable.* Ranging from 18 to 80.	Bonde (2008); Van den Heuvel et al. (2010)
Educational level	Personal characteristics	The highest achieved diploma or degree.	*Categorical variable.* Ranging from primary education to master/PhD.	Bonde (2008)
Ethnicity	Personal characteristics	The status of someone as regarding their citizenship of a country.	*Dummy variables.* A classification is made between people with a non-western migration background, western migration background and no migration background. The migration background label is assigned in case at least one parent is born abroad (CBS, 2016) (e.g. 0 = No migration background, 1 = non-western migration background).	Smulders et al. (2013)
Marital status	Personal characteristics	Formal position of a person in reference to marriage and civil partnership (CBS, 2016).	*Dummy variables.* Four dummies: never married, married, divorced, widowed.	Bonde (2008),
Household composition	Personal characteristics	The way the household of a person is designed.	*Dummy variables.* Married with children, unmarried with children, married without children, unmarried without children, one-parent household, one-person, other/unknown.	Smulders et al. (2013)
Province	Personal characteristics	The province the respondent is living in.	*Dummy variables.* As the Netherlands contains 12 provinces, 12 dummies will be included. (e.g. 0 = all other provinces, 1 = Noord-Holland).	-
Independent contractor	Personal characteristics	A person who performs labor at own expense or risk – within a private company or practice, or – as director – main shareholder, or – as miscellaneous independent actor, and - thereby, not having any personnel employed (CBS, 2016).	*Dummy variable.* 0 = No, 1 = Yes.	Cardon and Patel (2015)
Organizational tenure	Personal characteristics	The time an employee has worked in the organization.	*Continuous variable.* Measured in years. As a data is provided in the survey, the final score will be calculated by Year-survey – Year seniority in organization. For example, 2014 – 2010 = 4 years.	Chen and Kao (2012)

Multiple jobs	Personal characteristics	Whether or not an employee has multiple jobs.	*Dummy variable.* 0 = No, 1 = Yes.	-
General heath	Personal characteristics	'A state of complete physical, mental and social well-being and not merely the absence of disease or infirmity' (WHO, 1946).	*Categorical variable.* Five categories ranging from very bad to very good.	Maslach (2001)
Occupational accident	Personal characteristics	'An accident that happens in the workplace' (Niza, Silva, & Lima, 2008, p.968)	*Dummy variable.* 0 = No, 1 = Yes.	Peeters et al. (2013)
Supervisor support	Social characteristics	The extent to which supervisors "concern for the wellbeing of their subordinates, helping employees with their career development, and valuing work of those who report to them" (Paterson, Luthans, & Jeung, 2014, p.5).	*Categorical variable.* A weighted average is calculated over the two questions with four answer categories ranging from totally disagree to totally agree.	Peeters et al. (2013); Schaufeli, Bakker, et al. (2009)
Co-worker support	Social characteristics	The care and consideration employees receive from organizational members at the same organizational level (Mossholder, Settoon, & Henagan, 2005).	*Categorical variable.* A weighted average is calculated over the two questions with four answer categories ranging from totally disagree to totally agree.	Peeters et al. (2013); Schaufeli, Bakker, et al. (2009)
Interpersonal conflict	Social characteristics	'One person's actions interfere, obstruct, or in some way get in the way of another's action' (Tjosvold, 2008, p.24).	*Categorical variable*. A weighted average is calculated over the three questions with five answer categories no, short-term and long-term.	Crawford et al. (2010); Johns and Saks (2005)
Intimidation	Social characteristics	All those acts that repeatedly and persistently aimed to torment, wear down, or frustrate a person, and all repeated behaviors which ultimately would provide, frighten, intimidate or bring discomfort to the victim (Einarsen, 2000, p.380).	*Categorical variable.* A weighted average is calculated over the two questions with four answer categories never, incidentally, often, very often.	Crawford et al. (2010); Lim, Cortina, and Magley (2008)
Physical violence	Social characteristics	'Physical attacks directed toward individuals, their property, or both' (Rogers & Kelloway, 1997, p.64).	*Categorical variable.* A weighted average is calculated over the two questions with four answer categories never, incidentally, often, very often.	Crawford et al. (2010); Lim et al. (2008); Van den Heuvel et al. (2010)
Bullying	Social characteristics	'Harassing, offending, or socially excluding someone or negatively affecting someone's work. In order for the label bullying (or mobbing) to be applied to a particular activity, interaction or process, the bullying behavior has to occur repeatedly and regularly (e.g. weekly) and over a period of time (e.g., about six months)' (Einarsen, Hoel, Zapf, & Cooper, 2011, p.23).	*Categorical variable.* A weighted average is calculated over the two questions with four answer categories never, incidentally, often, very often.	Crawford et al. (2010); Lim et al. (2008); Van den Heuvel et al. (2010)

Appendix B: Descriptive statistics

Table B1

Behaviors and attitudes

Feature	Description	Mean	SD	Min.	Max.
Employability	Range 1-4	2.88	0.59	1	4
Turnover intention	Range 1-2	1.57	0.26	1	2
Under- or over qualification	1 = Underqualification, 2 = Balanced qualification, 3 = Overqualification	2.28	0.55	1	3
Job satisfaction	Range 1-5	3.82	0.898	1	5
Satisfaction with working conditions	Range 1-5	3.74	0.91	1	5

Note. SD = Standard Deviation; Min. = Minimum, Max. = Maximum.

Table B2

Job characteristics

Feature	Description	Mean	SD	Min.	Max.
Working time	Working time in hours per week	29.64	11.87	0	95
Shift work	Range 1-3	1.31	0.69	1	3
Working between 19:00 and 00:00	Range 1-3	1.81	0.82	1	3
Working between 00:00 and 6:00	Range 1-3	1.24	0.57	1	3
Working Saturdays	Range 1-3	1.81	0.81	1	3
Working Sundays	Range 1-3	1.58	0.78	1	3
Working overtime	Range 1-3	2.02	0.78	1	3
Working at home	Range 1-3	2.20	0.66	1	3
Managing position	1 = No, 2= Yes	1.27	0.44	1	2
Dangerous work	Range 1-3	1.26	0.52	1	3
On-call work	Range 1-3	1.41	0.69	1	3
Workload	Range 1-3	2.37	0.67	1	3
Physical demands (Environmental)	Range 1-4	1.37	0.54	1	4
Physical demands (Heavy loads)	Range 1-3	1.58	0.79	1	3
Physical demands (Unusual or tiring body positions and movements)	Range 1-3	1.69	0.65	1	3
Cognitive demands	Range 1-4	3.00	0.70	1	4
Job control	Range 1-3	1.64	0.50	1	3
Screen work	Working in front of a screen in hours per day	9.01	2.96	0	13
Demotion	1 = No, 2= Yes	1.04	0.20	0	1
Promotion	1 = No, 2= Yes	1.13	0.34	0	1
Job change	1 = No, 2= Yes	1.18	0.38	0	1
Job enlargement/enrichment	1 = No, 2= Yes	1.41	0.49	0	1
Job insecurity	Range 1-2	1.28	0.39	1	2
Employment contract	1= Fixed contract, 2 = Temporary contract	1.23	0.42	1	2

Note. SD = Standard Deviation; Min. = Minimum, Max. = Maximum.

Table B3

Organizational characteristics

Feature	Description	Mean/%	SD	Min.	Max.
Size	Range 1-9	5.34	2.41	1	9
Sector	Agriculture, fishery and forestry	1,19%			
	Industry	15.31%			
	Construction	4.30%			
	Retail	16.95%			
	Transport and storage	4.87%			
	Hospitality	4.38%			
	Financial services	3.99%			
	Business services and real estate	13.59%			
	Public governance	6.80%			
	Education	6.61%			
	Health-care	15.56%			
	Culture, sport, recreation and other	3.04%			
Organizational changes[1]	Acquisition of another company	4.02%			
	Acquisition of own company	4.00%			
	Reorganization	18.68%			
	Downsizing with forced layoffs	13.36%			
	Downsizing without forced layoffs	12.45%			
	Merger	4.12%			
	Outsourcing	8.53%			
	Off-shoring	3.09%			
	Automation	9.94%			
	No change	53.35%			
Occupational Health & Safety (OHS) practices	Range 1-4	1.60	0.411	1	4

Note. SD = Standard Deviation; Min. = Minimum, Max. = Maximum.

[1] Note: These percentages do not add up to 100 percent, as it was possible to select more organizational changes.

Table B4
Personal characteristics

Feature	Description	Mean/%	SD	Min.	Max.
Sex	Male = 1, Female = 2	1.46	0.50	1	2
Chronological age	Range 15- 75	42.27	13.78	15	75
Educational level	Range 1-5	2.30	1.10	1	5
National status	No migration background	85.09%			
	Non-western migration background	7.83%			
	Western migration background	7.83%			
Marital status	Unmarried	40.01%			
	Married/partnership	51.59%			
	Divorced	7.40%			
	Widowed	1.00%			
Province	Zeeland	2.17%			
	Noord-Brabant	16.26%			
	Limburg	7.10%			
	Gelderland	12.22%			
	Drenthe	2.74%			
	Zuid-Holland	20.26%			
	Noord-Holland	15.31%			
	Flevoland	2.15%			
	Friesland	3.50%			
	Groningen	3.21%			
	Overijssel	7.10%			
	Utrecht	7.98%			
Household composition	Married couple with children	42.80%			
	Unmarried couple with children	7.73%			
	Married couple without children	18.11%			
	Unmarried couple without children	10.20%			
	One-parent household	5.43%			
	One person household	15.04%			
	Unknown or other	0.69%			
Independent contractor	1 = No, 2= Yes	1.047	0.21	1	2
Seniority in organization	Seniority in organization in years; range 0-100	11.11	10.65	0	61
Multiple jobs	1 = No, 2 = Yes	1.07	0.25	1	2
General heath	Range 1-5	4.03	0.69	1	5
Occupational accident	1 = No, 2 = Yes	1.03	0.17	1	2

Note. SD = Standard Deviation; Min. = Minimum, Max. = Maximum.

Table B5
Social characteristics

Feature	Description	Mean	SD	Min.	Max.
Supervisor support	Range 1-4	2.99	0.70	1	4
Co-worker support	Range 1-4	3.30	0.57	1	4
Interpersonal conflict	Range 1-4	1.17	0.32	1	4
Intimidation	Range 1-4	1.17	0.33	1	4
Physical violence	Range 1-4	1.03	0.14	1	4
Bullying	Range 1-4	1.07	0.23	1	4

Note. SD = Standard Deviation; Min. = Minimum, Max. = Maximum.

Table B6
Outcome variables

Feature	Description	Mean	SD	Min.	Max.
Burnout complaints	Range 1-5	2.05	1.02	0	7

Note. SD = Standard Deviation; Min. = Minimum, Max. = Maximum.

Appendix C: Experiment 1 – Complete tables

Table C1

Within variable category variable importance: Job characteristics

	Relative weights				Random regression forest	
Ranking	Variable	%	LLCI	ULCI	Variable	% Increase in MSE
1	Workload (+)	40.43	38.72	42.09	Workload	0.255
2	Job insecurity (+)	13.16	11.92	14.38	Job control	0.134
3	Job control (-)	9.63	8.68	10.58	Screen work	0.111
4	Cognitive job demands (+)	9.36	8.53	10.21	PD – Unusual body positions	0.091
5	PD – Unusual body positions (+)	5.43	4.79	6.12	Working hours	0.081
6	Screen work (+)	3.71	3.21	4.26	Job insecurity	0.066
7	Working overtime (+)	3.60	3.10	4.10	Cognitive demands	0.060
8	Working hours (+)	2.80	2.38	3.23	Working at home	0.050
9	Demotion (+)	1.64	1.17	2.16	PD – Environmental	0.050
10	PD – environmental (+)	1.58	1.26	2.02	PD – Heavy loads	0.041
11	Dangerous work (+)	1.54	1.18	1.95	Working 19:00 to 00:00	0.034
12	Working at home (-)	1.43	1.16	1.72	Working Sundays	0.033
13	Employment contract (-)	1.25	0.98	1.57	Working Saturdays	0.032
14	PD – Heavy loads (-)	0.94	0.86	1.04	Dangerous work	0.027
15	Job enlargement (+)	0.70	0.52	0.94	Shift work	0.020
16	Working Saturdays (-)	0.54	0.40	0.74	Managerial position	0.019
17	Managerial position (-)	0.45	0.40	0.53	Working overtime	0.019
18	Job change (+)	0.41	0.24	0.61	Working 00:00 to 06:00	0.016
19	Promotion (-)	0.39	0.26	0.57	Employment contract	0.014
20	Working 00:00 to 06:00 (-)	0.29	0.18	0.46	Job change	0.008
21	Shift work (-)	0.24	0.21	0.60	On-call wok	0.007
22	Working 19:00 to 00:00 (-)	0.20	0.18	0.24	Job enrichment	0.006
22	On-call work (+)	0.16	0.09	0.27	Promotion	0.005
24	Working Sundays (+)	0.14	0.14	0.18	Demotion	0.004
OP training set	R^2 = 20.91				R^2 = 20.56	
PBP training set	R^2 = 18.16				R^2 = 14.98	
PBP test set	R^2 = 17.41				R^2 = 14.08	

Note. OP = Overall performance; PBP = Performance Best Predictors; PD = Physical job demands; % = Percentage (i.e. percentage of total variance explained); Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages)' MSE = Mean Squared Error.

Table C2

Within variable category variable importance: Organizational characteristics

	Relative weights				Random regression forest	
Ranking	Variable	%	LLCI	ULCI	Variable	% Increase in MSE
1	OHS (-)	83.07	81.43	84.25	OHS	0.481
2	Sector: education (+)	2.57	1.91	3.29	Change: no change	0.074
3	Change: reorganization (+)	2.46	1.93	3.05	Organizational size	0.070
4	Change: downsizing with forced layoffs (+)	2.01	1.50	2.58	Change: reorganization	0.046
5	Change: downsizing without forced layoffs (+)	1.73	1.24	2.23	Change: downsizing without forced layoffs	0.027
6	Change: outsourcing (+)	1.72	1.27	2.28	Change: downsizing with forced layoffs	0.026
7	Change: automation (+)	1.15	0.82	1.59	Sector: education	0.021
8	Organizational size (+)	0.91	0.84	1.01	Sector: health-care	0.016
9	Sector: health-care (+)	0.63	0.51	0.81	Sector: retail	0.013
10	Sector: retail (-)	0.62	0.41	0.90	Sector: public governance	0.010
11	Change: acquisition own organization (+)	0.54	0.30	0.89	Sector: information & communication	0.009
12	Sector: industry (+)	0.36	0.28	0.50	Sector: industry	0.009
13	Sector: transport & storage (-)	0.35	0.15	0.61	Change: outsourcing	0.008
14	Sector: information & communication (+)	0.31	0.17	0.52	Change: automation	0.007
15	Change: merger (+)	0.30	0.15	0.57	Change: merger	0.006
16	Sector: agriculture, fishery & forestry (-)	0.25	0.12	0.42	Change: acquisition of other company	0.005
17	Change: off-shoring (+)	0.23	0.10	0.45	Sector: financial services	0.005
18	Sector: hospitality (-)	0.18	0.08	0.33	Change: acquisition of own company	0.001
19	Sector: construction (-)	0.16	0.07	0.32	Sector: business services & real estate	0.001
20	Sector: public governance (-)	0.16	0.09	0.30	Sector: hospitality	0.004
21	Sector: financial services (-)	0.13	0.10	0.23	Change: off-shoring	0.003
22	Change: acquisition other organization (+)	0.12	0.02	0.29	Sector: transport & storage	0.000
23	Sector: culture, sport, recreation & other (-)	0.04	0.01	0.12	Sector: construction	0.000
24	-				Sector: culture, sport, recreation & other	0.000
25	-				Sector: agriculture, fishery & forestry	0.000
OP training set	R^2 = 16.42				R^2 = 14.42	
PBP training set	R^2 = 15.17				R^2 = 16.02	
PBP test set	R^2 = 12.94				R^2 = 14.00	

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); * = Identified as best predictor; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages); Reference category 'Household' dummy-variable = Married with children; Reference category 'Marital status' dummy-variable = Married; Reference category 'Province' dummy-variable = Noord-Holland; Reference category 'Ethnicity' dummy-variable = No migration background; MSE = Mean Squared Error.

Table C3

Within variable category variable importance: Personal characteristics

	Relative weights				Random regression forest	
Ranking	Variable	%	LLCI	ULCI	Variable	% Increase in MSE
1	General health (-)	83.11	81.13	84.48	General health	0.407
2	Occupational accident occurrence (+)	4.97	3.84	6,21	Age	0.197
3	Household: one-person (+)	3.69	2.97	4.55	Organizational tenure	0.098
4	Ethnicity: non-western migration background (+)	2.03	1.41	2.70	Marital status: married	0.090
5	Household: unmarried without children (+)	1.80	1.19	2.36	Marital status: never married	0.061
6	Age (-)	0.60	.054	0.70	Household: married with children	0.056
7	Organizational tenure (+)	0.52	0.34	0.76	Educational level	0.052
8	Ethnicity: western migration background (+)	0.52	0.26	0.87	Household: married without children	0.027
9	Marital status: divorced (-)	0.49	0.26	0.80	Occupational accident	0.021
10	Multiple jobs (-)	0.46	0.24	0.72	Household: one person	0.020
11	Marital status: never married (-)	0.32	0.27	0.41	Household: unmarried without children	0.014
12	Province: Overijssel (-)	0.21	0.08	0.41	Marital status: divorced	0.010
13	Province: Friesland (-)	0.14	0.04	0.31	Household: unmarried with children	0.008
14	Household: one-parent (+)	0.13	0.05	0.27	Ethnicity: no migration background	0.007
15	Province: Noord-Brabant (-)	0.12	0.03	0.29	Ethnicity: non-western migration background	0.006
16	Province: Drenthe (-)	0.09	0.01	0.24	Ethnicity: western migration background	0.004
17	Household: married without children (+)	0.09	0.07	0.17	Household: One-parent	0.004
18	Household: other or unknown (+)	0.08	0.01	0.29	Multiple jobs	0.003
19	Education level (-)	0.07	0.01	0.23	Independent contractor	0.003
20	Province: Utrecht (+)	0.07	0.01	0.22	Marital status: widowed	0.002
21	Province: Zuid-Holland (-)	0.07	0.01	0.21	Province: Utrecht	0.002
22	Marital status: widowed (-)	0.07	0.02	021	Household: other or unknown	0.000
23	Province: Flevoland (+)	0.06	0.01	0.22	Province: Friesland	0.000
24	Province: Limburg (-)	0.06	0.02	0.18	Province Drenthe	0.000
25	Province: Zeeland (-)	0.06	0.00	0.20	Province: Limburg	0.000
26	Household: unmarried with children (+)	0.06	0.04	0.13	Province: Flevoland	0.000
27	Province: Gelderland (-)	0.05	0.01	0.15	Province: Noord-Brabant	0.000
28	Independent contractor (+)	0.02	0.00	0.13	Province: Overijssel	-0.001
29	Province: Groningen (-)	0.02	0.00	0.10	Province: Gelderland	-0.002
30	-	-	-	-	Province: Groningen	-0.003
31	-	-	-	-	Province: Zeeland	-0.002
32	-	-	-	-	Province: Zuid-Holland	-0.0032
33	-	-	-	-	Province: Noord-Holland	-0.003
OP training set	R^2 = 15.03				R^2 = 12.68	
PBP training set	R^2 = 13.67				R^2 = 16.64	
PBP test set	R^2 = 14.28				R^2 = 17.13	

Note. OP = Overall performance; PBP = Performance Best Predictors; % = Percentage (i.e. percentage of total variance explained); * = Identified as best predictor; Number of bootstrap = 1000; Confidence interval = 95%; LLCI = Lower Level Confidence Interval; ULCI = Higher Level Confidence Interval; (+) implies a positive effect, (-) implies a negative effect, R^2 = Total variance explained (in percentages); Reference category 'Household' dummy-variable = Married with children; Reference category 'Marital status' dummy-variable = Married; Reference category 'Province' dummy-variable = Noord-Holland; Reference category 'Ethnicity' dummy-variable = No migration background; MSE = Mean Squared Error.

Appendix D: Cut-off point selection

Based on Figure D1a and Figure D1b, it was decided to use the fourth variable as cut-off point. Solely based on Figure D1b, one would rather select the third variable as a cut-off point, as it is the first variable where the difference between the importance metrics becomes much larger. However, considering the rather small difference between the third and fourth variable in Figure D1a, the large difference between fourth to the fifth variable in the relative weights analysis and still substantial difference between these two variables in the random regression forest ranking, the fourth variable was selected as cut-off point.

Table D1
Attitudes and behaviors – cut-off point

	Relative weights			Random regression forest		
Ranking	Variable	%	Difference	Variable	% Increase in MSE	Difference
1*	Employability	33.12	9.240	Satisfaction with working conditions	0.189	0.020
2*	Satisfaction with working conditions	23.88	2.250	Job satisfaction	0.169	0.033
3*	Job satisfaction	21.63	1.010	Employability	0.136	0.057
4*	Turnover intention	20.62	19.846	Turnover intention	0.079	0.060
5	Under-or overqualification	0.774	0.774	Under-or overqualification	0.019	0.019

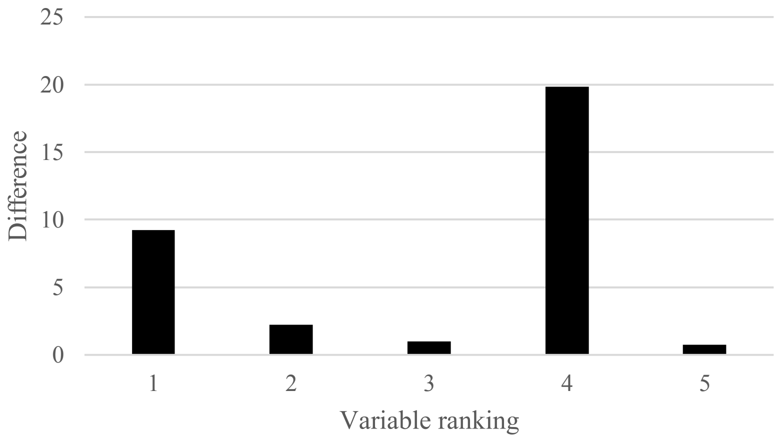

Figure D1a. Differences relative weights ranking.

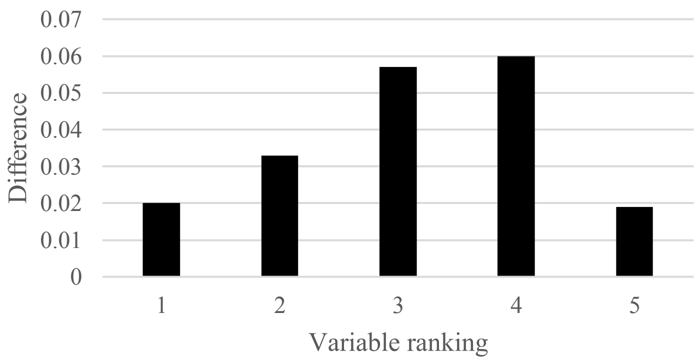

Figure D1b. Differences random regression forest ranking.

As visualized in Figure D2a and D2b, the most severe drop in importance metric both occurred after the first and second variable. Yet, because random regression forests do not work with only one variable, the first variable could not be used as a cut-off point. When comparing the two bar-charts, it is clear that especially the first five variables are characterized by rather large differences. Even though after the first five variables there are still some relatively steep drops in importance metrics, it was decided to use the fifth variable as cut-off point.

Table D2
Job characteristics – cut-off point

	Relative weights			Random regression forest		
Ranking	Variable	%	Difference	Variable	% Increase in MSE	Difference
1*	Workload	40.43	27.27	Workload	0.255	0.121
2*	Job insecurity	13.16	3.53	Job control	0.134	0.023
3*	Job control	9.63	0.27	Screen work	0.111	0.020
4*	Cognitive job demands	9.36	3.93	PD – Unusual body positions	0.091	0.010
5*	PD – Unusual body positions	5.43	1.72	Working hours	0.081	0.015
6	Screen work	3.71	0.11	Job insecurity	0.066	0.006
7	Working overtime	3.60	0.8	Cognitive demands	0.060	0.010
8	Working hours	2.80	1.16	Working at home	0.050	0.000
9	Demotion	1.64	0.06	PD – Environmental	0.050	0.009
10	PD – environmental	1.58	0.04	PD – Heavy loads	0.041	0.007
11	Dangerous work	1.54	0.11	Working 19:00 to 00:00	0.034	0.001
12	Working at home	1.43	0.18	Working Sundays	0.033	0.001
13	Employment contract	1.25	0.31	Working Saturdays	0.032	0.005
14	PD – Heavy loads	0.94	0.24	Dangerous work	0.027	0.007
15	Job enlargement	0.70	0.16	Shift work	0.020	0.001
16	Working Saturdays	0.54	0.09	Managerial position	0.019	0.000
17	Managerial position	0.45	0.04	Working overtime	0.019	0.003
18	Job change	0.41	0.02	Working 00:00 to 06:00	0.016	0.002
19	Promotion	0.39	0.1	Employment contract	0.014	0.006
20	Working 00:00 to 06:00	0.29	0.05	Job change	0.008	0.001
21	Shift work	0.24	0.04	On-call wok	0.007	0.001
22	Working 19:00 to 00:00	0.20	0.04	Job enrichment	0.006	0.001
22	On-call work	0.16	0.02	Promotion	0.005	0.001
24	Working Sundays	0.14	0.14	Demotion	0.004	0.004

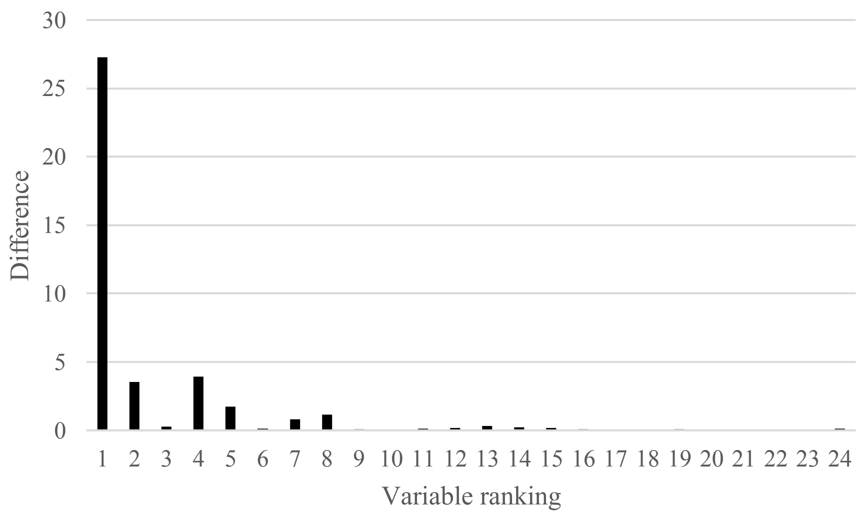

Figure D2a. Differences relative weights ranking.

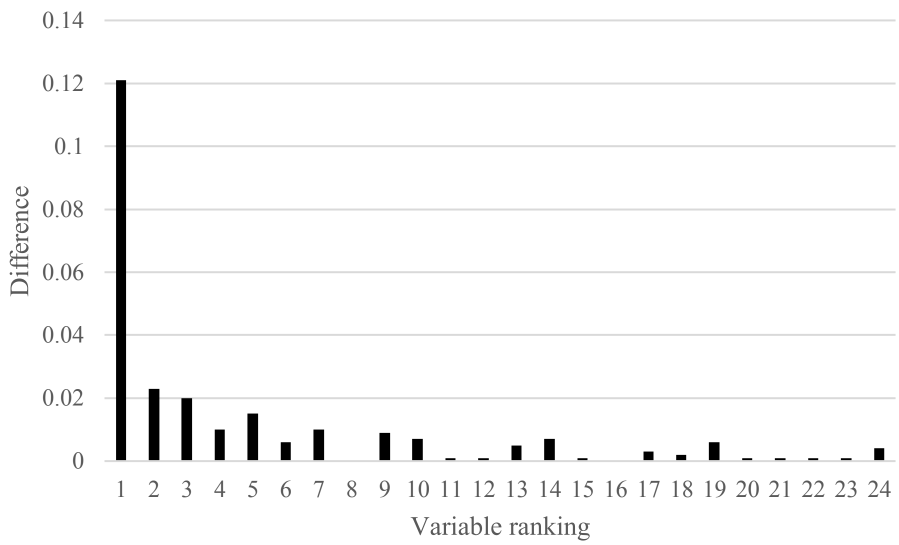

Figure D2b. Differences random regression forest ranking.

The social characteristics ranking is characterized by a steep drop in importance metric after the first variable. Since random forests do not work with only one explanatory variable, the first variable could not be used as a cut-off point. As visualized in Figure D3a, there is a very steep drop after the fourth variable. Although less steep in Figure D3b, it was assumed that the fourth variable was the most suitable cut-off point for model evaluation.

Table D3
Social characteristics – cut-off point

	Relative weights			Random regression forest		
Ranking	Variable	%	Difference	Variable	% Increase in MSE	Difference
1*	Supervisor support	31.76	9.04	Supervisor support	0.1378	9.04
2*	Intra-personal conflict	22.72	2.51	Intra-personal conflict	0.0648	2.51
3*	Intimidation	20.21	1.16	Intimidation	0.0643	1.16
4*	Bullying	19.05	14.16	Bullying	0.0501	14.16
5	Colleague support	4.89	3.52	Colleague support	0.0290	3.52
6	Physical violence	1.37	1.37	Physical violence	0.0000	1.37

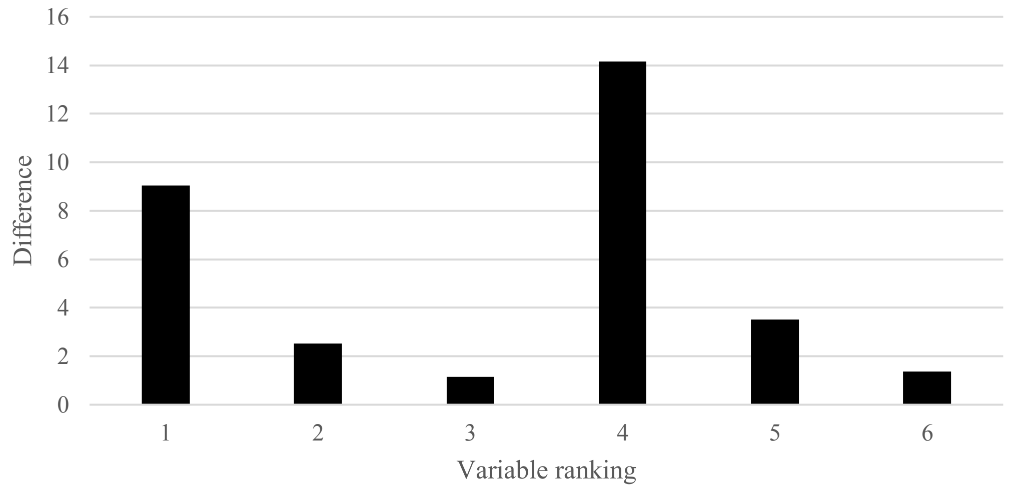

Figure D3a. Differences relative weights ranking.

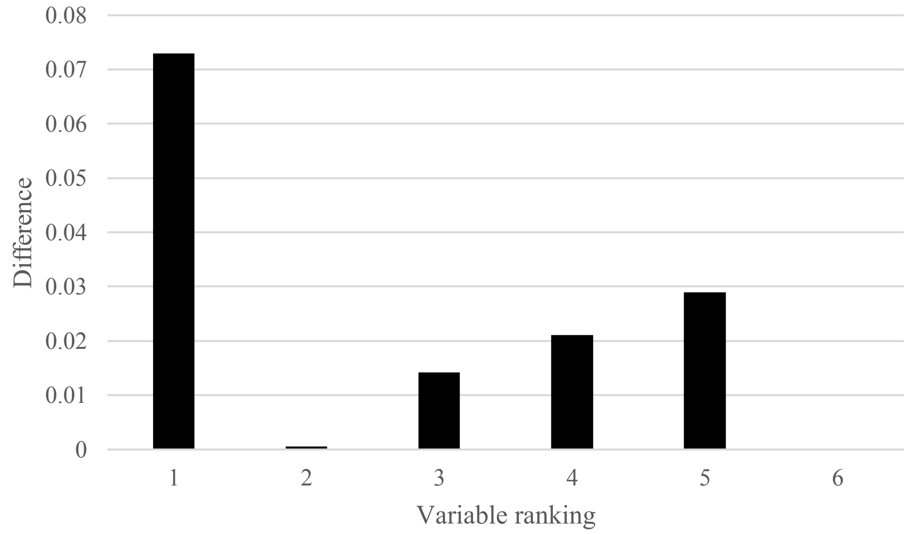

Figure D3b. Differences random regression forest ranking.

While the drop in importance metric was very severe after the first variable, OHS could not be used as cut-off point, because random forests do not work with only one independent variable. So, when looking for the first steep drop in performance, Figure D4b showed that the third variable would be a suitable cut-off point. As the relative weights method only had negligible differences after the first variable (see Figure D4a), it seems logical to select the cut-off point based on the bar-chart that actually shows other differences, i.e. Figure D4b). Hence, the third variable was selected as cut-off point.

Table D4
Organizational characteristics – cut-off point

	Relative weights			Random regression forest		
Ranking	Variable	%	Difference	Variable	% Increase in MSE	Difference
1*	OHS	83.07	80.50	OHS	0.481	0.407
2*	Sector: education	2.57	0.11	Change: no change	0.074	0.004
3*	Change: reorganization	2.46	0.45	Organizational size	0.070	0.024
4	Change: downsizing with forced layoffs	2.01	0.28	Change: reorganization	0.046	0.019
5	Change: downsizing without forced layoffs	1.73	0.01	Change: downsizing without forced layoffs	0.027	0.001

6	Change: outsourcing	1.72	0.57	Change: downsizing with forced layoffs	0.026	0.005
7	Change: automation	1.15	0.24	Sector: education	0.021	0.005
8	Organizational size	0.91	0.28	Sector: health-care	0.016	0.003
9	Sector: health-care	0.63	0.01	Sector: retail	0.013	0.003
10	Sector: retail	0.62	0.08	Sector: public governance	0.010	0.001
11	Change: acquisition own organization	0.54	0.18	Sector: information & communication	0.009	0.000
12	Sector: industry	0.36	0.01	Sector: industry	0.009	0.001
13	Sector: transport & storage	0.35	0.04	Change: outsourcing	0.008	0.001
14	Sector: information & communication	0.31	0.01	Change: automation	0.007	0.001
15	Change: merger	0.30	0.05	Change: merger	0.006	0.001
16	Sector: agriculture. fishery & forestry	0.25	0.02	Change: acquisition of other company	0.005	0.000
17	Change: off-shoring	0.23	0.05	Sector: financial services	0.005	0.004
18	Sector: hospitality	0.18	0.02	Change: acquisition of own company	0.001	0.000
19	Sector: construction	0.16	0.00	Sector: business services & real estate	0.001	-0.003
20	Sector: public governance	0.16	0.03	Sector: hospitality	0.004	0.001
21	Sector: financial services	0.13	0.01	Change: off-shoring	0.003	0.003
22	Change: acquisition other organization	0.12	0.08	Sector: transport & storage	0.000	0.000
23	Sector: culture. sport. recreation & other	0.04	0.04	Sector: construction	0.000	0.000
24	-	-		Sector: culture. sport. recreation & other	0.000	0.000
25	-	-	-	Sector: agriculture. fishery & forestry	0.000	0.000

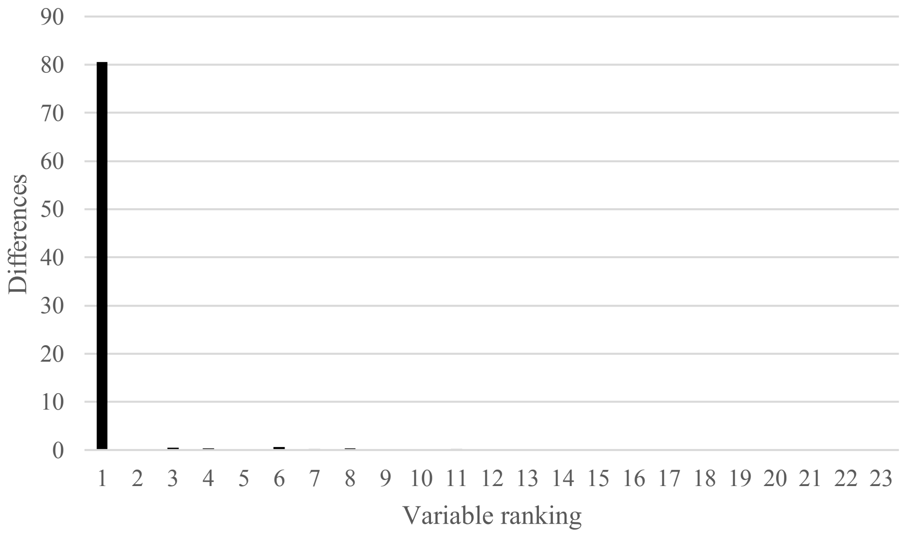

Figure D4a. Differences relative weights ranking.

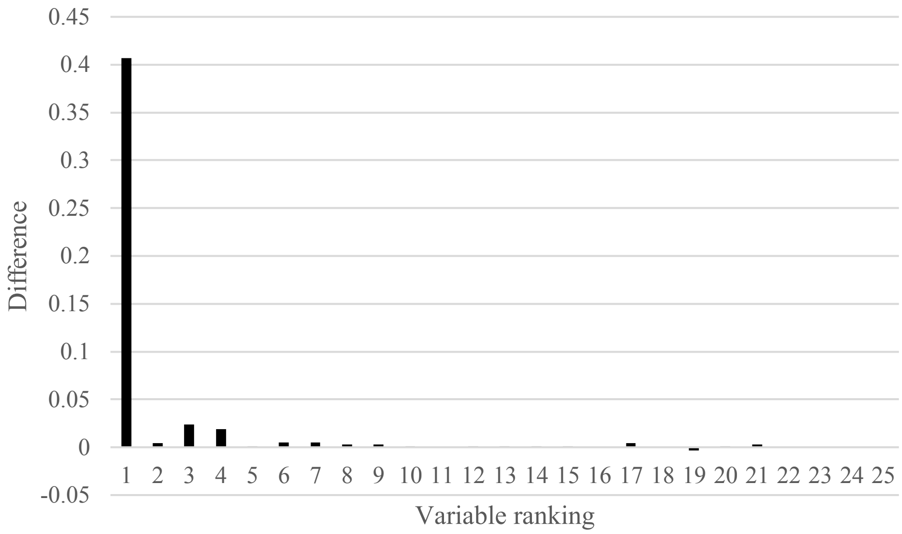

Figure D4b. Differences random regression forest ranking.

The personal characteristics ranking of the relative weights method does not show steep drops after the first variable (see Figure D5a), while the ranking by the random regression forest (see Figure D5b) shows some steep declines. Accordingly, the random regression forest was used to determine the cut-off point. As the difference between the second and third variable was both the first and the greatest (disregarding the difference between the first and second variable), the second variable was selected as cut-off point.

Table D5
Personal characteristics – cut-off point

	Relative weights			Random regression forest		
Ranking	Variable	%	Difference	Variable	% Increase in MSE	Difference
1*	General health	83.11	78.14	General health	0.407	0.210
2*	Occupational accident occurrence	4.97	1.28	Age	0.197	0.099
3	Household: one-person	3.69	1.66	Organizational tenure	0.098	0.008
4	Ethnicity: non-western migration background	2.03	0.23	Marital status: married	0.090	0.029
5	Household: unmarried without children	1.80	1.20	Marital status: never married	0.061	0.005
6	Age	0.60	0.08	Household: married with children	0.056	0.004
7	Organizational tenure	0.52	0.00	Educational level	0.052	0.025
8	Ethnicity: western migration background	0.52	0.03	Household: married without children	0.027	0.006
9	Marital status: divorced	0.49	0.03	Occupational accident	0.021	0.001
10	Multiple jobs	0.46	0.14	Household: one person	0.020	0.006
11	Marital status: never married	0.32	0.11	Household: unmarried without children	0.014	0.004
12	Province: Overijssel	0.21	0.07	Marital status: divorced	0.010	0.002
13	Province: Friesland	0.14	0.01	Household: unmarried with children	0.008	0.001
14	Household: one-parent	0.13	0.01	Ethnicity: no migration background	0.007	0.001
15	Province: Noord-Brabant	0.12	0.03	Ethnicity: non-western migration background	0.006	0.002
16	Province: Drenthe	0.09	0.00	Ethnicity: western migration background	0.004	0.000
17	Household: married without children	0.09	0.01	Household: One-parent	0.004	0.001
18	Household: other or unknown	0.08	0.01	Multiple jobs	0.003	0.000
19	Education level	0.07	0.00	Independent contractor	0.003	0.001

20	Province: Utrecht	0.07	0.00	Marital status: widowed	0.002	0.000
21	Province: Zuid-Holland	0.07	0.00	Province: Utrecht	0.002	0.002
22	Marital status: widowed	0.07	0.01	Household: other or unknown	0.000	0.000
23	Province: Flevoland	0.06	0.00	Province: Friesland	0.000	0.000
24	Province: Limburg	0.06	0.00	Province Drenthe	0.000	0.000
25	Province: Zeeland	0.06	0.00	Province: Limburg	0.000	0.000
26	Household: unmarried with children	0.06	0.01	Province: Flevoland	0.000	0.000
27	Province: Gelderland	0.05	0.03	Province: Noord-Brabant	0.000	0.001
28	Independent contractor	0.02	0.00	Province: Overijssel	-0.001	0.001
29	Province: Groningen	0.02	0.02	Province: Gelderland	-0.002	0.001
30	-	-	-	Province: Groningen	-0.003	-0.001
31	-	-	-	Province: Zeeland	-0.002	0.0012
32	-	-	-	Province: Zuid-Holland	-0.0032	-0.0002
33	-	-	-	Province: Noord-Holland	-0.003	-0.003

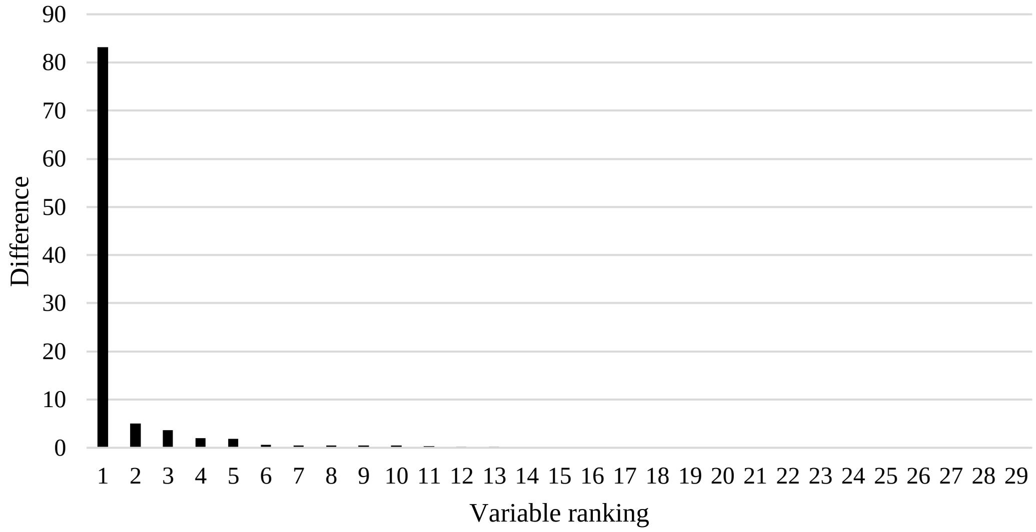

Figure D5a. Differences relative weights ranking.

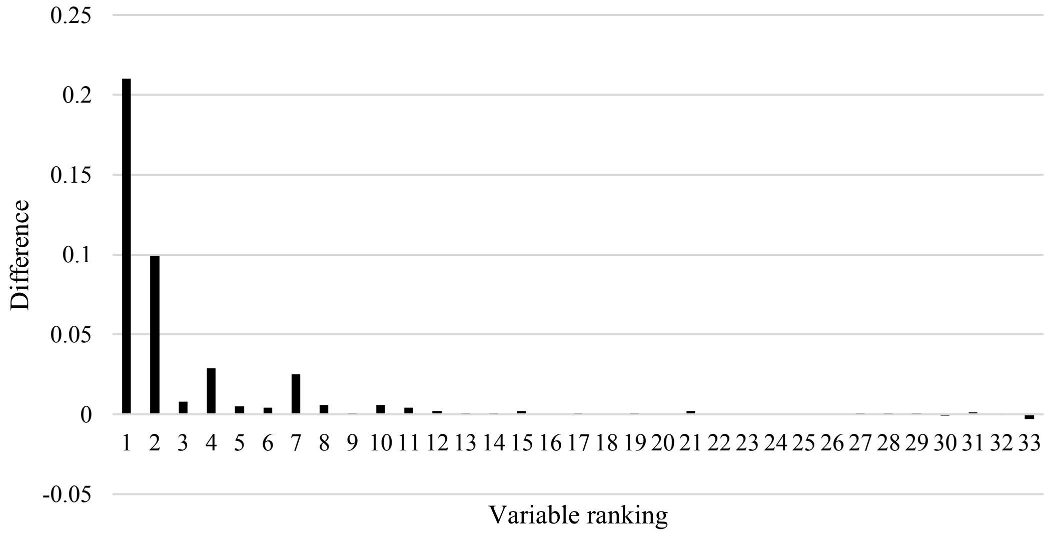

Figure D5b. Differences random regression forest ranking.

It proved rather difficult to select a cut-off point based on the two rankings (Figure D6a and D6b). Firstly, in the random regression forest the difference between the first and second variable was significantly larger than in the LMG ranking. Secondly, the LMG ranking showed a clear drop in importance metric after the second variable, whereas this difference in the random regression forest was negligible. Thirdly, now the other way around, the random regression forest displayed a relatively large difference between variable thee and four, while the LMG reported almost no difference. Still, a cut-off point had to be selected. On the one hand, the difference between the first and second variable in the random regression forest ranking was very high. One could argue that the differences between the other variables compared to this very high difference are trivial and hard to interpret. On the other hand, the difference between the first and second variable in the LMG ranking was much smaller. The difference between the second and third variable was even larger than the differences between the first and second. Now, mostly basing the decision on the LMG ranking and regarding the small difference between the third and fourth variable, it was decided to use the third variable as a cut-off point.

Table D6
Between categories – cut-off point

	LMG			Random regression forest		
Ranking	Variable	%	Difference	Variable	Importance metric	Difference
1*	Job characteristics	25.53	2.70	Attitudes and behaviors	0.055	0.043
2*	Attitudes and behaviors	22.83	3.89	Social characteristics	0.012	0.001
3*	Social characteristics	18.94	0.83	Job characteristics	0.011	0.006
4	Personal characteristics	18.11	3.51	Personal characteristics	0.005	0.001
5	Organizational characteristics	14.60	14.60	Organizational characteristics	0.004	0.004

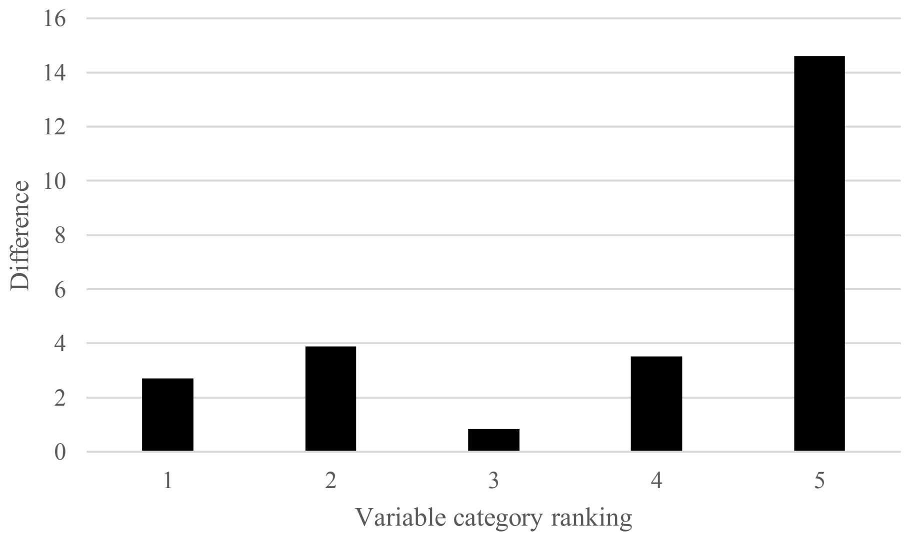

Figure D6a. Differences LMG ranking.

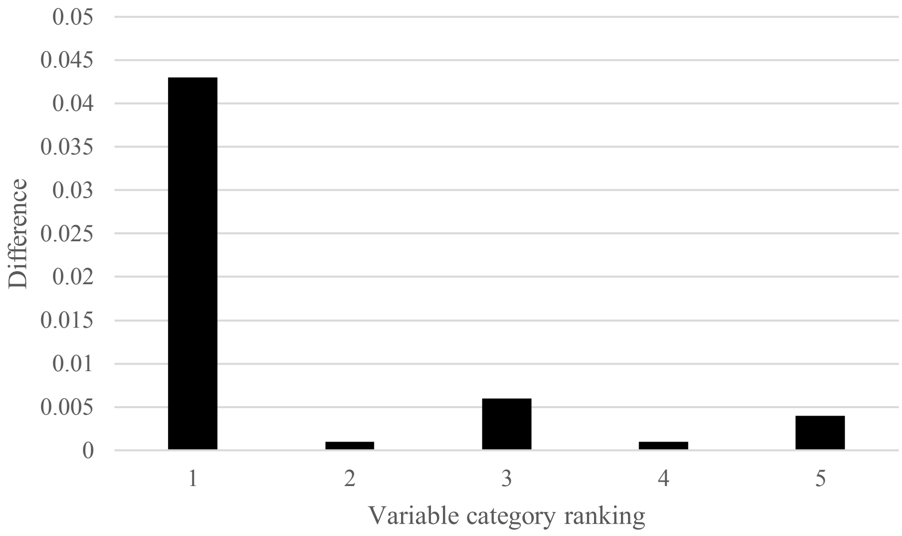

Figure D6b. Differences random regression forest ranking.

References

Adriaenssens, J., De Gucht, V., & Maes, S. (2015). Determinants and prevalence of burnout in emergency nurses: a systematic review of 25 years of research. *International journal of nursing studies, 52*(2), 649-661.

Ahola, K., Salminen, S., Toppinen-Tanner, S., Koskinen, A., & Väänänen, A. (2013). Occupational burnout and severe injuries: An eight-year prospective cohort study among Finnish forest industry workers. *Journal of Occupational Health, 55*(6), 450-457.

Alarcon, G. M. (2011). A meta-analysis of burnout with job demands, resources, and attitudes. *Journal of Vocational Behavior, 79*(2), 549-562.

Alarcon, G. M., Eschleman, K. J., & Bowling, N. A. (2009). Relationships between personality variables and burnout: A meta-analysis. *Work and Stress, 23*(3), 244-263. doi:10.1080/02678370903282600

Audigier, V., Husson, F., & Josse, J. (2016). A principal component method to impute missing values for mixed data. *Advances in Data Analysis and Classification, 10*(1), 5-26.

Aybas, M., Elmas, S., & Dundar, G. (2015). Job insecurity and burnout: The moderating role of employability. *European Journal of Business and Management, 7*, 195-202.

Bailey, D. E., & Kurland, N. B. (2002). A review of telework research: Findings, new directions, and lessons for the study of modern work. *Journal of organizational behavior, 23*(4), 383-400.

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of managerial psychology*, *22*(3), 309-328.

Bakker, A. B., Van Der Zee, K. I., Lewig, K. A., & Dollard, M. F. (2006). The relationship between the big five personality factors and burnout: A study among volunteer counselors. *The Journal of social psychology*, *146*(1), 31-50.

Bannai, A. (2014). The association between long working hours and health: a systematic review of epidemiological evidence. *Scandinavian journal of work, environment & health*, *40*(1), 5.

Beckers, D. G., van der Linden, D., Smulders, P. G., Kompier, M. A., van Veldhoven, M. J., & van Yperen, N. W. (2004). Working overtime hours: relations with fatigue, work motivation, and the quality of work. *Journal of Occupational and Environmental Medicine, 46*(12), 1282-1289.

Behson, S. J. (2012). Using relative weights to reanalyze research on the job characteristics model. *Journal of Organizational Psychology, 12*(2), 71.

Ben-Zur, H., & Yagil, D. (2005). The relationship between empowerment, aggressive behaviours of customers, coping, and burnout. *European journal of work and organizational psychology, 14*(1), 81-99.

Benach, J., Amable, M., Muntaner, C., & Benavides, F. G. (2002). The consequences of flexible work for health: are we looking at the right place? *Journal of epidemiology and community health, 56*(6), 405-406.

Berente, N., & Seidel, S. (2014). Big Data & Inductive Theory Development: Towards Computational Grounded Theory?

Berry, C. M., Lelchook, A. M., & Clark, M. A. (2012). A meta-analysis of the interrelationships between employee lateness, absenteeism, and turnover: Implications for models of withdrawal behavior. *Journal of organizational behavior, 33*(5), 678-699.

Beyene, J., Atenafu, E. G., Hamid, J. S., To, T., & Sung, L. (2009). Determining relative importance of variables in developing and validating predictive models. *BMC medical research methodology*, *9*(1), 64.

Bi, J. (2012). A review of statistical methods for determination of relative importance of correlated predictors and identification of drivers of consumer liking. *Journal of Sensory Studies, 27*(2), 87-101.
Black, C. D., & Frost, D. (2011). *Health at work-an independent review of sickness absence* (Vol. 8205): The Stationery Office.
Bonde, J. P. E. (2008). Psychosocial factors at work and risk of depression: a systematic review of the epidemiological evidence. *Occupational and environmental medicine, 65*(7), 438-445.
Borg, V., Kristensen, T. S., & Burr, H. (2000). Work environment and changes in self-rated health: a five year follow-up study. *Stress and Health, 16*(1), 37-47.
Brauchli, R., Bauer, G. F., & Hämmig, O. (2011). Relationship between time-based work-life conflict and burnout. *Swiss Journal of Psychology*.
Braun, M. T., & Oswald, F. L. (2011). Exploratory regression analysis: A tool for selecting models and determining predictor importance. *Behavior research methods, 43*(2), 331-339.
Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.
Brief, A. P. (1998). *Attitudes in and around organizations* (Vol. 9): Sage.
Brown, L. W., & Quick, J. C. (2013). Environmental Influences on Individual Burnout and a Preventive Approach for Organizations. *Journal of Applied Biobehavioral Research, 18*(2), 104-121.
Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological bulletin, 114*(3), 542.
Budescu, D. V., & Azen, R. (2004). Beyond global measures of relative importance: Some insights from dominance analysis. *Organizational Research Methods, 7*(3), 341-350.
Camerino, D., Sandri, M., Sartori, S., Conway, P. M., Campanini, P., & Costa, G. (2010). Shiftwork, work-family conflict among Italian nurses, and prevention efficacy. *Chronobiology international, 27*(5), 1105-1123.
Cardon, M. S., & Patel, P. C. (2015). Is Stress Worth it? Stress-Related Health and Wealth Trade-Offs for Entrepreneurs. *Applied Psychology, 64*(2), 379-420.
Carod-Artal, F. J., & Vázquez-Cabrera, C. (2013). Burnout syndrome in an international setting *Burnout for experts* (pp. 15-35): Springer.
CBS. (2016). Begrippen. Retrieved from https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen?tab=a - id=allochtoon
Ceja, L., & Navarro, J. (2012). 'Suddenly I get into the zone': Examining discontinuities and nonlinear changes in flow experiences at work. *Human relations, 65*(9), 1101-1127.
Chen, C.-F., & Kao, Y.-L. (2012). Moderating effects of work engagement and job tenure on burnout–performance among flight attendants. *Journal of Air Transport Management, 25*, 61-63.
Chiaburu, D. S., Oh, I. S., Wang, J., & Stoverink, A. C. (2016). A Bigger Piece of the Pie: The Relative Importance of Affiliative and Change-Oriented Citizenship and Task Performance in Predicting Overall Job Performance. *Human Resource Management Review (Forthcoming)*.
Cohen, G., Blake, R. S., & Goodman, D. (2016). Does turnover intention matter? Evaluating the usefulness of turnover intention rate as a predictor of actual turnover rate. *Review of Public Personnel Administration, 36*(3), 240-263.
Cooper, C., & Dewe, P. (2008). Well-being—absenteeism, presenteeism, costs and challenges. *Occupational medicine, 58*(8), 522-524.
Costa, G. (2010). Shift work and health: current problems and preventive actions. *Safety and health at Work, 1*(2), 112-123.

Crawford, E. R., LePine, J. A., & Rich, B. L. (2010). Linking job demands and resources to employee engagement and burnout: a theoretical extension and meta-analytic test. *Journal of Applied Psychology, 95*(5), 834.
Cugnata, F., & Salini, S. (2017). Comparison of alternative imputation methods for ordinal data. *Communications in Statistics-Simulation and Computation, 46*(1), 315-330.
Danna, K., & Griffin, R. W. (1999). Health and well-being in the workplace: A review and synthesis of the literature. *Journal of Management, 25*(3), 357-384.
Datta, D. K., Guthrie, J. P., Basuil, D., & Pandey, A. (2010). Causes and effects of employee downsizing: A review and synthesis. *Journal of Management, 36*(1), 281-348.
Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application* (Vol. 1): Cambridge university press.
De Beer, L. T., Pienaar, J., & Rothmann Jr, S. (2013). Investigating the reversed causality of engagement and burnout in job demands-resources theory. *SA Journal of Industrial Psychology, 39*(1), 01-09.
De Croon, E. M., Sluiter, J. K., Blonk, R. W., Broersen, J. P., & Frings-Dresen, M. H. (2004). Stressful work, psychological job strain, and turnover: a 2-year prospective cohort study of truck drivers. *Journal of Applied Psychology, 89*(3), 442.
De Cuyper, N., Raeder, S., Van der Heijden, B. I., & Wittekind, A. (2012). The association between workers' employability and burnout in a reorganization context: longitudinal evidence building upon the conservation of resources theory. *Journal of occupational health psychology, 17*(2), 162.
de Heus, P., van der Leeden, R., & Gazendam, B. (1995). *Toegepaste data-analyse: technieken voor niet-experimenteel onderzoek in de sociale wetenschappen*: Lemma.
De Jonge, J., Reuvers, M. M., Houtman, I. L., Bongers, P. M., & Kompier, M. A. (2000). Linear and nonlinear relations between psychosocial job characteristics, subjective outcomes, and sickness absence: Baseline results from SMASH. *Journal of occupational health psychology, 5*(2), 256.
De Jonge, J., & Schaufeli, W. B. (1998). Job characteristics and employee well-being: A test of Warr's Vitamin Model in health care workers using structural equation modelling. *Journal of organizational behavior*, 387-407.
Demerouti, E., Bakker, A. B., & Bulters, A. J. (2004). The loss spiral of work pressure, work–home interference and exhaustion: Reciprocal relations in a three-wave study. *Journal of Vocational Behavior, 64*(1), 131-149.
Dohmen, T. J., Kriechel, B., & Pfann, G. A. (2004). Monkey bars and ladders: The importance of lateral and vertical job mobility in internal labor market careers. *Journal of Population Economics, 17*(2), 193-228.
Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field research. *Academy of management review, 32*(4), 1246-1264.
Edwards, D., Burnard, P., Coyle, D., Fothergill, A., & Hannigan, B. (2000). Stress and burnout in community mental health nursing: a review of the literature. *Journal of psychiatric and mental health nursing, 7*(1), 7-14.
Egger, M., Smith, G. D., & Sterne, J. A. (2001). Uses and abuses of meta-analysis. *Clinical Medicine, 1*(6), 478-484.
Einarsen, S. (2000). Harassment and bullying at work: A review of the Scandinavian approach. *Aggression and violent behavior, 5*(4), 379-401.
Einarsen, S., Hoel, H., Zapf, D., & Cooper, C. L. (2011). The concept of bullying and harassment at work: The European tradition. *Bullying and harassment in the workplace: Developments in theory, research, and practice, 2*, 3-40.

Eisenhardt, K. M., Graebner, M. E., & Sonenshein, S. (2016). Grand challenges and inductive methods: Rigor without rigor mortis. *Academy of Management Journal, 59*(4), 1113-1123.

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of statistical software, 48*(4), 1-18.

Erdogan, B., & Bauer, T. N. (2009). Perceived overqualification and its outcomes: the moderating role of empowerment. *Journal of Applied Psychology, 94*(2), 557.

Eurostat. (2016). Small and medium-sized enterprises (SMEs). Retrieved from http://ec.europa.eu/eurostat/web/structural-business-statistics/structural-business-statistics/sme?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_vxlB58HY09rg&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=4

Faragher, E. B., Cass, M., & Cooper, C. L. (2005). The relationship between job satisfaction and health: a meta-analysis. *Occupational and environmental medicine, 62*(2), 105-112.

Farsi, Z., Nayeri, N. D., & Sajadi, A. (2013). General health status and burnout of the hematopoietic stem cell transplantation nurses in Tehran. *International Journal of Community Based Nursing and Midwifery, 1*(1), 52-61.

Feldman, B. (1999). The proportional value of a cooperative game. *Manuscript: Scuder Kemper Investment*.

Ferris, G. R., Bowen, M. G., Treadway, D. C., Hochwarter, W. A., Hall, A. T., & Perrewé, P. L. (2006). The assumed linearity of organizational phenomena: Implications for occupational stress and well-being *Employee Health, Coping and Methodologies* (pp. 203-232): Emerald Group Publishing Limited.

Field, A. (2009). *Discovering statistics using SPSS*: Sage publications.

Fila, M. J., Purl, J., & Griffeth, R. W. (2017). Job demands, control and support: Meta-analyzing moderator effects of gender, nationality, and occupation. *Human Resource Management Review, 27*(1), 39-60.

Fragoso, Z. L., Holcombe, K. J., McCluney, C. L., Fisher, G. G., McGonagle, A. K., & Friebe, S. J. (2016). Burnout and Engagement Relative Importance of Predictors and Outcomes in Two Health Care Worker Samples. *Workplace Health & Safety*, 2165079916653414.

Frick, K., Jensen, P. L., Quinlan, M., & Wilthagen, T. (2000). *Systematic occupational health and safety management: perspectives on an international development*: Pergamon Press.

Gajendran, R. S., & Harrison, D. A. (2007). The good, the bad, and the unknown about telecommuting: meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology, 92*(6), 1524.

Garrosa, E., Moreno-Jimenez, B., Liang, Y., & González, J. L. (2008). The relationship between socio-demographic variables, job stressors, burnout, and hardy personality in nurses: An exploratory study. *International journal of nursing studies, 45*(3), 418-427.

Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research 1. *Educational researcher, 5*(10), 3-8.

Golino, H. F., & Gomes, C. M. (2016). Random forest as an imputation method for education and psychology research: its impact on item fit and difficulty of the Rasch model. *International Journal of Research & Method in Education, 39*(4), 401-421.

González-Morales, M. G., & Neves, P. (2015). When stressors make you work: Mechanisms linking challenge stressors to performance. *Work & stress, 29*(3), 213-229.

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2013). Correlation and variable importance in random forests. *Statistics and Computing*, 1-20.

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2015). Grouped variable importance with random forests and application to multiple functional data analysis. *Computational Statistics & Data Analysis, 90*, 15-35.
Grömping, U. (2006). Relative importance for linear regression in R: the package relaimpo. *Journal of statistical software, 17*(1), 1-27.
Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician, 61*(2), 139-147.
Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician, 63*(4), 308-319.
Grömping, U. (2015). Variable importance in regression models. *Wiley Interdisciplinary Reviews: Computational Statistics, 7*(2), 137-152.
Guyon, I., Gunn, S., Nikravesh, M., & Zadeh, L. A. (2008). *Feature extraction: foundations and applications* (Vol. 207): Springer.
Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning, 46*(1), 389-422.
Hakanen, J. J., Schaufeli, W. B., & Ahola, K. (2008). The Job Demands-Resources model: A three-year cross-lagged study of burnout, depression, commitment, and work engagement. *Work & stress, 22*(3), 224-241.
Halbesleben, J. R., & Buckley, M. R. (2004). Burnout in organizational life. *Journal of Management, 30*(6), 859-879.
Hancock, J. I., Allen, D. G., Bosco, F. A., McDaniel, K. R., & Pierce, C. A. (2013). Meta-analytic review of employee turnover as a predictor of firm performance. *Journal of Management, 39*(3), 573-603.
Hapfelmeier, A., & Ulm, K. (2013). A new variable selection approach using random forests. *Computational Statistics & Data Analysis, 60*, 50-69.
Hauge, L. J., Skogstad, A., & Einarsen, S. (2010). The relative impact of workplace bullying as a social stressor at work. *Scandinavian Journal of Psychology, 51*(5), 426-433.
Hawthorne, D. (2011). Predictor Importance for Selection: Best Practices and Latest Findings: A discussion of conventional wisdom surrounding predictor importance and selection systems.
Hooftman, W., Mars, G., Janssen, B., de Vroome, E., Michiels, J., & van den Bossche, S. (2016). Nationale Enquête Arbeidsomstandigheden 2015. Methodologie en globale resultaten.
Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics, 15*(3), 651-674.
Howitt, D., & Cramer, D. (2007). *Introduction to research methods in psychology*: Pearson Education.
Jebb, A. T., Parrigon, S., & Woo, S. (2016). Exploratory data analysis as a foundation of inductive research. *Human Resource Management Review*.
Jermier, J. M., Gaines, J., & McIntosh, N. J. (1989). Reactions to physically dangerous work: A conceptual and empirical analysis. *Journal of organizational behavior, 10*(1), 15-33.
Jin, J. (2014). *The introduction and application of recursive partitioning methods in organizational science.* University of Illinois at Urbana-Champaign.
Johns, G., & Saks, A. M. (2005). *Organizational behaviour: Understanding and managing life at work*: Toronto: Pearson Education.
Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate behavioral research, 35*(1), 1-19.
Johnson, J. W. (2004). Factors affecting relative weights: The influence of sampling and measurement error. *Organizational Research Methods, 7*(3), 283-299.

Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods, 7*(3), 238-257.
Josten, E., & Schalk, R. (2010). The effects of demotion on older and younger employees. *Personnel Review, 39*(2), 195-209.
Karanika-Murray, M., & Cox, T. (2010). The use of artificial neural networks and multiple linear regression in modelling work–health relationships: Translating theory into analytical practice. *European journal of work and organizational psychology, 19*(4), 461-486.
Kath, L. M., Stichler, J. F., Ehrhart, M. G., & Sievers, A. (2013). Predictors of nurse manager stress: a dominance analysis of potential work environment stressors. *International journal of nursing studies, 50*(11), 1474-1480.
Ketokivi, M., & Mantere, S. (2010). Two strategies for inductive reasoning in organizational research. *Academy of Management Review, 35*(2), 315-333.
Khamisa, N., Peltzer, K., & Oldenburg, B. (2013). Burnout in relation to specific contributing factors and health outcomes among nurses: a systematic review. *International journal of environmental research and public health, 10*(6), 2214-2240.
Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. *British Journal of Educational Psychology, 77*(1), 229-243.
Korpinen, L., Pääkkönen, R., & Gobba, F. (2013). Self-reported neck symptoms and use of personal computers, laptops and cell phones among Finns aged 18–65. *Ergonomics, 56*(7), 1134-1146.
Korpinen, L., Pääkkönen, R., & Gobba, F. (2016). *Self-Reported Ache, Pain, or Numbness in Feet and Use of Computers amongst Working-Age Finns.* Paper presented at the Healthcare.
Kruskal, W. (1987). Relative importance by averaging over orderings. *The American Statistician, 41*(1), 6-10.
Ladstätter, F., Garrosa, E., Moreno-Jiménez, B., Ponsoda, V., Reales Aviles, J. M., & Dai, J. (2016). Expanding the occupational health methodology: A concatenated artificial neural network approach to model the burnout process in Chinese nurses. *Ergonomics, 59*(2), 207-221.
Lance, C. E., Lance, C. E., & Vandenberg, R. J. (2010). *Statistical and methodological myths and urban legends: Doctrine, verity and fable in organizational and social sciences*: Routledge.
Landy, F. J., & Conte, J. M. (2016). *Work in the 21st Century, Binder Ready Version: An Introduction to Industrial and Organizational Psychology*: John Wiley & Sons.
Lasalvia, A., Bonetto, C., Bertani, M., Bissoli, S., Cristofalo, D., Marrella, G., . . . Lazzarotto, L. (2009). Influence of perceived organisational factors on job burnout: survey of community mental health staff. *The British Journal of Psychiatry, 195*(6), 537-544.
Lazarus, R. S. (2000). Toward better research on stress and coping.
LeBreton, J. M., Hargis, M. B., Griepentrog, B., Oswald, F. L., & Ployhart, R. E. (2007). A multidimensional approach for evaluating variables in organizational research and practice. *Personnel Psychology, 60*(2), 475-498.
Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news, 2*(3), 18-22.
Lim, S., Cortina, L. M., & Magley, V. J. (2008). Personal and workgroup incivility: impact on work and health outcomes. *Journal of Applied Psychology, 93*(1), 95.
Lindblom, K. M., Linton, S. J., Fedeli, C., & Bryngelsson, I.-L. (2006). Burnout in the working population: relations to psychosocial work factors. *International journal of behavioral medicine, 13*(1), 51-59.

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*: Scott, Foresman Glenview, IL.

Locke, E. A. (2007). The Case for Inductive Theory Building†. *Journal of Management, 33*(6), 867-890.

Luchman, J. N. (2014). Relative Importance Analysis With Multicategory Dependent Variables An Extension and Review of Best Practices. *Organizational Research Methods*, 1094428114544509.

Markoulli, M., Lee, C. I., Byington, E., & Felps, W. A. (2016). Mapping Human Resource Management: Reviewing the field and charting future directions. *Human Resource Management Review*.

Maslach, C. (2001). What have we learned about burnout and health? *Psychology & Health, 16*(5), 607-611.

Maslach, C. (2003). Job burnout: New directions in research and intervention. *Current directions in psychological science, 12*(5), 189-192.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual review of psychology, 52*(1), 397-422.

Matsuki, K., Kuperman, V., & Van Dyke, J. A. (2016). The Random Forests statistical technique: An examination of its value for the study of reading. *Scientific Studies of Reading, 20*(1), 20-33.

Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annu. Rev. Psychol., 59*, 537-563.

Morse, T., Dillon, C., Kenta-Bibi, E., Weber, J., Diva, U., Warren, N., & Grey, M. (2005). Trends in work-related musculoskeletal disorder reports by year, type, and industrial sector: A capture-recapture analysis. *American journal of industrial medicine, 48*(1), 40-49.

Mossholder, K. W., Settoon, R. P., & Henagan, S. C. (2005). A relational perspective on turnover: Examining structural, attitudinal, and behavioral predictors. *Academy of Management Journal, 48*(4), 607-618.

Murphy, K. R. (2015). What inferences can and cannot be made on the basis of meta-analysis?—SI. *Human Resource Management Review*.

Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. (2011). Safety at work: a meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology, 96*(1), 71.

Nauta, A., Vianen, A., Heijden, B., Dam, K., & Willemsen, M. (2009). Understanding the factors that promote employability orientation: the impact of employability culture, career satisfaction, and role breadth self-efficacy. *Journal of Occupational and Organizational psychology, 82*(2), 233-251.

Navarro, M. L. A., Mas, M. B., & Jiménez, A. M. L. (2010). Working conditions, burnout and stress symptoms in university professors: validating a structural model of the mediating effect of perceived personal competence. *The Spanish journal of psychology, 13*(01), 284-296.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods, 17*(4), 372-411.

Nicol, A.-M., & Botterill, J. S. (2004). On-call work and health: a review. *Environmental Health, 3*(1), 1.

Nimon, K. F., & Oswald, F. L. (2013). Understanding the results of multiple linear regression beyond standardized regression coefficients. *Organizational Research Methods, 16*(4), 650-674.

Niza, C., Silva, S., & Lima, M. L. (2008). Occupational accident experience: Association with workers' accident explanation and definition. *Safety science, 46*(6), 959-971.

Noblet, A. J., & Rodwell, J. J. (2009). Integrating job stress and social exchange theories to predict employee strain in reformed public sector contexts. *Journal of Public Administration Research and Theory, 19*(3), 555-578.

Norlund, S., Reuterwall, C., Höög, J., Lindahl, B., Janlert, U., & Birgander, L. S. (2010). Burnout, working conditions and gender-results from the northern Sweden MONICA Study. *BMC public health, 10*(1), 326.

Oakley, A. (1985). *Sex, gender and society.* (Vol. Revised edition). London: Temple Smith.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2016). Realizing the full potential of psychometric meta-analysis for a cumulative science and practice of human resource management. *Human Resource Management Review*.

Oshiro, T. M., Perez, P. S., & Baranauskas, J. A. (2012). *How many trees in a random forest?* Paper presented at the International Workshop on Machine Learning and Data Mining in Pattern Recognition.

Paris, M., & Hoge, M. A. (2010). Burnout in the mental health workforce: A review. *The journal of behavioral health services & research, 37*(4), 519-528.

Paterson, T. A., Luthans, F., & Jeung, W. (2014). Thriving at work: Impact of psychological capital and supervisor support. *Journal of organizational behavior, 35*(3), 434-446.

Pawlowski, S. D., Kaganer, E. A., & Cater III, J. J. (2007). Focusing the research agenda on burnout in IT: social representations of burnout in the profession. *European Journal of Information Systems, 16*(5), 612-627.

Peeters, M. C., De Jonge, J., & Taris, T. W. (2013). *An introduction to contemporary work psychology*: John Wiley & Sons.

Piha, K., Laaksonen, M., Martikainen, P., Rahkonen, O., & Lahelma, E. (2009). Interrelationships between education, occupational class, income and sickness absence. *The European Journal of Public Health*, ckp162.

Piotrowski, C. (2012). Occupational health psychology: Neglected areas of research. *Journal of Instructional Psychology, 39*(3), 189-191.

Porras, J. I., & Robertson, P. J. (1992). *Organizational development: Theory, practice, and research*: Consulting Psychologists Press.

Pretnar, N. (2015). Obesity, Food Insecurity, and Future Discounting: Comparisons of County-Level Empirical Analyses Using Random Forest and Linear Regression.

Purvanova, R. K., & Muros, J. P. (2010). Gender differences in burnout: A meta-analysis. *Journal of Vocational Behavior, 77*(2), 168-185.

Ramlall, S. (2004). A review of employee motivation theories and their implications for employee retention within organizations. *Journal of American Academy of Business, 5*(1/2), 52-63.

Roelen, C. A., Koopmans, P. C., Notenbomer, A., & Groothoff, J. W. (2008). Job satisfaction and sickness absence: a questionnaire survey. *Occupational medicine, 58*(8), 567-571.

Rogers, K.-A., & Kelloway, E. K. (1997). Violence at work: personal and organizational outcomes. *Journal of occupational health psychology, 2*(1), 63.

Rydstedt, L. W., Ferrie, J., & Head, J. (2006). Is there support for curvilinear relationships between psychosocial work characteristics and mental well-being? Cross-sectional and long-term data from the Whitehall II study. *Work & stress, 20*(1), 6-20.

Schaufeli, W. B., Bakker, A. B., & Van Rhenen, W. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of organizational behavior, 30*(7), 893-917.

Schaufeli, W. B., Leiter, M. P., & Maslach, C. (2009). Burnout: 35 years of research and practice. *Career development international, 14*(3), 204-220.

Schaufeli, W. B., & Taris, T. W. (2014). A critical review of the Job Demands-Resources Model: Implications for improving work and health *Bridging occupational, organizational and public health* (pp. 43-68): Springer.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*: Sage publications.

Siu, O. L., Cooper, C. L., & Phillips, D. R. (2014). Intervention studies on enhancing work well-being, reducing burnout, and improving recovery experiences among Hong Kong health care workers and teachers. *International Journal of Stress Management, 21*(1), 69.

Smith, P. F., Ganesh, S., & Liu, P. (2013). A comparison of random forest regression and multiple linear regression for prediction in neuroscience. *Journal of neuroscience methods, 220*(1), 85-91.

Smulders, P., Houtman, I., Rijssen, J. v., & Mol, M. (2013). Burnout: trends, internationale verschillen, determinanten en effecten. *Tijdschrift voor Arbeidsvraagstukken, 29*, 258-278.

Sorour, A. S., & El-Maksoud, M. M. A. (2012). Relationship between musculoskeletal disorders, job demands, and burnout among emergency nurses. *Advanced emergency nursing journal, 34*(3), 272-282.

Sparks, K., & Cooper, C. L. (2013). Occupational Differences in the Work-Strain Relationship: Towards the Use of Situation-Specific Models *From Stress to Wellbeing Volume 1* (pp. 315-326): Springer.

Spector, P. E. (2015). Induction, Deduction Abduction: Three Legitimate Approaches to Organizational Research.

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics, 28*(1), 112-118.

Stone, D. L., & Rosopa, P. J. (2016). The Advantages and Limitations of Using Meta-analysis in Human Resource Management Research. *Human Resource Management Review*.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC bioinformatics, 8*(1), 25.

Sverke, M., Hellgren, J., & Näswall, K. (2002). No security: a meta-analysis and review of job insecurity and its consequences. *Journal of occupational health psychology, 7*(3), 242.

Taris, T., Houtman, I., & Schaufeli, W. (2013). Burnout: de stand van zaken. *Tijdschrift voor Arbeidsvraagstukken, 29*(3), 241-257.

Taris, T. W., & Schreurs, P. J. (2009). Well-being and organizational performance: An organizational-level test of the happy-productive worker hypothesis. *Work & stress, 23*(2), 120-136.

Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: path analyses based on meta-analytic findings. *Personnel Psychology, 46*(2), 259-293.

Thomas, M., Kohli, V., & Choi, J. (2014). Correlates of job burnout among human services workers: Implications for workforce retention. *J. Soc. & Soc. Welfare, 41*, 69.

Tjosvold, D. (2008). The conflict-positive organization: It depends upon us. *Journal of organizational behavior, 29*(1), 19-28.

Tonidandel, S., King, E. B., & Cortina, J. M. (2016). Big Data Methods Leveraging Modern Data Analytic Techniques to Build Organizational Science. *Organizational Research Methods*, 1094428116677299.

Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology, 26*(1), 1-9.

Tonidandel, S., LeBreton, J. M., & Johnson, J. W. (2009). Determining the statistical significance of relative weights. *Psychological methods, 14*(4), 387.

Toppinen-Tanner, S. (2011). Process of burnout: Structure, antecedents, and consequences.

Tsigilis, N., Zachopoulou, E., & Grammatikopoulos, V. (2006). Job satisfaction and burnout among Greek early educators: A comparison between public and private sector employees. *Educational Research and Reviews, 1*(8), 256.

Turnipseed, D. L. (1994). An analysis of the influence of work environment variables and moderators on the burnout syndrome. *Journal of Applied Social Psychology, 24*(9), 782-800.

Van De Voorde, K., Paauwe, J., & Van Veldhoven, M. (2012). Employee well-being and the HRM–organizational performance relationship: a review of quantitative studies. *International Journal of Management Reviews, 14*(4), 391-407.

Van den Heuvel, S. G., Geuskens, G. A., Hooftman, W. E., Koppes, L. L., & Van den Bossche, S. N. (2010). Productivity loss at work; health-related and work-related factors. *Journal of occupational rehabilitation, 20*(3), 331-339.

Van den Tooren, M. (2011). *Job demands, job resources, and self-regulatory behavior: Exploring the issue of match.* PhD Thesis. Technische Universiteit Eindhoven.

Van Der Ploeg, E., & Kleber, R. J. (2003). Acute and chronic job stressors among ambulance personnel: predictors of health symptoms. *Occupational and environmental medicine, 60*(suppl 1), i40-i46.

Van Droogenbroeck, F., Spruyt, B., & Vanroelen, C. (2014). Burnout among senior teachers: Investigating the role of workload and interpersonal relationships at work. *Teaching and Teacher Education, 43*, 99-109.

Vargas, C., Cañadas, G. A., Aguayo, R., Fernández, R., & Emilia, I. (2014). Which occupational risk factors are associated with burnout in nursing? A meta-analytic study. *International Journal of Clinical and Health Psychology, 14*(1), 28-38.

Walker, E., Hernandez, A. V., & Kattan, M. W. (2008). Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine, 75*(6), 431.

Watts, J., & Robertson, N. (2011). Burnout in university teaching staff: a systematic literature review. *Educational Research, 53*(1), 33-50.

Preamble to the Constitution of the World Health Organization as adopted by the International Health

Conference., Official Records of the World Health Organization 100 (1946).

Woo, S. E., O'Boyle, E. H., & Spector, P. E. (2016). Best practices in developing, conducting, and evaluating inductive research. *Human Resource Management Review*.

Xie, J. L., & Johns, G. (1995). Job scope and stress: Can job scope be too high? *Academy of Management Journal, 38*(5), 1288-1309.

Ye, C., Yang, Y., & Yang, Y. (2016). Sparsity Oriented Importance Learning for High-dimensional Linear Regression. *arXiv preprint arXiv:1608.00629*.

Zapf, D. (2002). Emotion work and psychological well-being: A review of the literature and some conceptual considerations. *Human Resource Management Review, 12*(2), 237-268.

Zivnuska, S., Kiewitz, C., Hochwarter, W. A., Perrewé, P. L., & Zellars, K. L. (2002). What is too much or too little? The curvilinear effects of job tension on turnover intent, value attainment, and job satisfaction. *Journal of Applied Social Psychology, 32*(7), 1344-1360.