Cover Page

Manuscript:

Diverging Ideas of Health?

Explanatory Factors for Self-Rated Health Across Gender, Age-Groups, and European Countries

Authors: Patrick Lazarevič[ab]

(Patrick.Lazarevic@TU-Dortmund.de)

& Martina Brandt[a]

(Martina.Brandt@TU-Dortmund.de)

[a]TU Dortmund University
Institute for Sociology
Emil-Figge-Str. 50
44227 Dortmund
Germany

[b]Austrian Academy of Sciences
Vienna Institute of Demography
Welthandelsplatz 2
Vienna, 1020
Austria

Corresponding author: Patrick Lazarevič

(Patrick.Lazarevic@oeaw.ac.at)

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 725187) and the Mercator Research Center Ruhr (MERCUR).

Abstract

Background: Self-rated health (SRH) is arguably the most widely used generic health measurement in survey research. However, SRH remains a black box for researchers. In our paper, we want to gain a better understanding of SRH by identifying its determinants, quantifying the contribution of different health domains to explain SRH, and by exploring the role of gender, age-groups, and the country of residence.

Method: Using data from 61,027 participants of the fifth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE) living in fifteen European countries, we explain SRH via linear regression models. The independent variables are grouped into five health domains: functioning, diseases, pain, depression, and behavior. Via dominance analysis, we focus on their individual contribution to explaining SRH and compare these contributions across gender, three age-groups, and fifteen European countries.

Results: Our model explains SRH rather well ($R^2 = 0.50$ for females/0.45 for males) with diseases contributing most to the appraisal (.16/.18). Functioning is the second most relevant factor (.16/.14) followed by pain (.09/.07) and depression (.06/.05). Health behavior (.02/.01) is less relevant for health ratings. This ranking holds true for almost all countries with only little variance overall. A comparison of age-groups, however, indicates that the contribution of diseases and behavior to SRH decreases over the life-course while the contribution of functioning to R^2 increases.

Conclusion: Our paper demonstrates that SRH is largely based on diverse health information with functioning and diseases being most important. However, there is still room for idiosyncrasies or even bias.

Research Highlights

- Self rated health (SRH) is explained via five dimensions of health information.
- Amount of explained variance by type is compared by gender, age-group, and country.
- Strongest predictors of SRH are functioning and diseases.
- Only minor differences can be found by gender and country.
- Contribution to R^2 by diseases/behavior decreases with age, opposite for functioning.

Keywords

Self-Rated Health, Survey of Health, Ageing and Retirement in Europe (SHARE), Cross-National Comparison, Epidemiology, Response Behavior, Measurement Invariance, Europe

Introduction

Self-rated health (SRH) is the most used generic health indicator in a wide array of scientific disciplines. It is usually collected via a single question asking for the respondents' health rating on a four or five-point scale (Jylhä 2009). In many studies, SRH is seen and treated as a valid generic health measurement and is used both as a resource (e.g., as a restriction or prerequisite for social participation) and as an outcome (e.g., when researching preservation or improvement of health). Its usage is most commonly (if at all) justified with its repeatedly demonstrated relation to mortality and it can be seen as an inclusive, dynamic, health behavior affecting, and resource reflecting measure of health status (Idler & Benyamini 1997; Benyamini 2011).

However, due to the vagueness of the question, survey respondents are relatively free to decide what to base their health rating on. The consequence is that researchers cannot be sure as to what exactly is measured by SRH (Jylhä 2009; Garbarski 2016). Yet, studies examining the health-related determinants of SRH found that it is also strongly and consistently correlated to a wide array of other common health indicators like symptoms or diagnoses of diseases and especially pain (e.g., Tornstam 1975), (consequences of) risky health-behaviors like being under-/overweight/obese (e.g., Manderbacka et al. 1999) or smoking (e.g., Wang & Arah 2015), mental health issues and depression (e.g., Kivinen et al. 1998), or health-related restrictions of the functional status or the daily life (e.g., Suchman et al. 1958).

Even if there is a vast amount of studies on determinants of subjective health-ratings, this should not hide the fact this research oftentimes lacks an underlying theoretical model, and does rarely touch up on the relative importance or weight of the identified determinants for the rating and/or the potential issue of measurement equivalence for different subgroups of the population, e.g., gender, age-groups, as well as for different countries. Yet, all these aspects are highly relevant for using SRH in empirical research.

Firstly, research on and with SRH without an underlying theoretical model of the

response process suitable to guide the analysis risks resulting in fragmented and isolated findings that do not contribute to scientific progress or purposeful evidence-based policy. Secondly, empirically determining to which extent respondents base their health-ratings on which health domains is important for survey research not least in order to know whether SRH is a suitable indicator of the intended concept of health in a given analysis. Thirdly, if there are group-differences in the health concept used to judge one's health status, any comparisons of SRH across these groups are called into question. This problem of measurement equivalence or differential item functioning, which is known as 'response shift' in the context of age-differences (Sprangers & Schwartz 1999), country-specific response styles in the context of international differences (Jürges 2007), or simply as sex/gender differences when referring to differences between male and female respondents (Schulz et al. 1994), can either produce health-differences when there are none or it can obscure actual group-differences. The purpose of this present paper is to (1) provide a cognitive model of the response process for SRH to guide concrete research for a better and more systematic understanding of SRH and (2) to contribute to research regarding SRH's underlying determinants, their relative importance, and possible group differences in order to lead to a better understanding of this generic health indicator in comparative settings.

We will firstly develop and describe a general cognitive model of the response process for SRH based on existing models on the cognitive process of answering survey questions, models of the process of health ratings and relevant empirical research. From this general model we will then derive a more specific analytical model for our analysis of the relative contributions of different health domains to SRH and potential group-differences. In this context, we are the first to systematize the determinants of SRH into different health domains.[1] This model will be followed by a short review of the literature concerning these strands of research. We then describe the data and methods used for our empirical analysis and the results from a linear regression explaining SRH by five types of health

[1]This theoretical discussion can also be found in greater detail elsewhere (Lazarevič 2018).

indicators already mentioned above: functioning, diseases, pain, depression, and behavior. The results will be separately by gender in order to compare women and men. We will then analyze the relative contributions of these five aspects to Rš for three age-groups and 15 European countries.

Theoretical Models: The Cognitive Process and Our Analytical Model

The Cognitive Process of Health-Ratings

[*Figure 1 here, portrait, one column*]

As a theoretical background for our analysis we have synthesized a model (Figure 1) by combining the general cognitive model of the response process for survey questions by Tourangeau (1984) in its extended form by Strack & Martin (1987) with the response process of SRH as proposed by Knäuper & Turner (2003) and Jylhä (2009). The model comprises, in accordance to the response process described by Strack & Martin (1987), four major steps and every step potentially influences the subsequent step.

The *first step* is the comprehension and interpretation of the question. The respondent has to interpret what is meant by '(general/overall) health'. Respondents might be particularly influenced by preceding questions if they are also concerned with health aspects in the form of assimilation or contrast/subtraction effects (Garbarski 2016). In both cases, strong survey design effects are to be expected: In the first case this would mean that respondents more or less summarize the information they have already given (Garbarski et al. 2015) while the latter would mean that respondents evaluate their health aside from the health information they already provided (Tourangeau et al. 1991).

After interpreting the question, as a *second step*, respondents need to generate an opinion. For this, they can either recall an evaluation of their health from memory (e.g., in panel surveys or if they are frequently occupied with thinking about their health) or

they can generate a health evaluation on the spot. In the latter case they first have to choose which (health) information is relevant for their rating. They might consider factors like medical diagnoses, observations about the functional status, pain experiences, and body perceptions as indicated by Knäuper & Turner (2003) and Jylhä (2009). However, in line with the 1948 definition of health by the World Health Organization (2006), one might also consider depression or depressive symptoms as relevant factors (Kivinen et al. 1998; Schnittker 2005; Han & Jylhä 2006). Information regarding these factors has to be recalled from the respondents memory implying that more salient information, e.g., acute health problems or pain (Knäuper & Turner 2003), is more likely to be recalled and deemed relevant because it is more accessible. The respondents then incorporate the available information into one global evaluation of health by either weighting the recalled factors in some way in order to incorporate them (Anderson 1971) or by using simpler heuristics like focusing on the most available/salient information (Tversky & Kahneman 1973). Either way it can be assumed that the information at hand is incorporated somewhat systematically, nonetheless, the systematics can vary between different groups of respondents or even individually. As a last part of this step, respondents select a reference frame (e.g., age-peers or themselves at an earlier point in time) and compare their overall health to it (Strack & Martin 1987; Krause & Jay 1994; Knäuper & Turner 2003; Cheng et al. 2007).

Once respondents generated an overall evaluation of health, the *third step* of rating their health lies in choosing the most adequate response option. Obviously, this step is strongly influenced by questionnaire features (Schwarz 1999) like response options (Lee 2015) or their order (Garbarski et al. 2015), implying problems of comparability especially between surveys that implement SRH differently. Lastly, as the *fourth step*, respondents may choose to edit their answer due to factors like social desirability (Strack & Martin 1987), e.g., in order to not seem frail or to gain sympathy (Maddox 1962).

Of course, it is likely that the process and outcome of each step are modified by personal characteristics like belonging to certain demographic or socioeconomic groups. In

this paper, we explicitly focus on demographic factors (i.e., gender, age, and country/questionnaire language) and leave aside other possible group differences such as education or income. Gender might play a role because women and men might have different concepts of health or find different health domains more (or less) relevant than others, e.g., due to gender specific health-reporting norms (Undén & Elofsson 2006; Caroli & Weber-Baghdiguian 2016; Zajacova et al. 2017). The same presumably applies to age due to older respondents' greater experience or even (perceived) normativeness of (adapting to) chronic diseases, co-morbidity, health-related limitations, and general physical and cognitive decline – both individually and in age-peers. This perception potentially influences how older respondents interpret the meaning of health, which information is most salient, or how they incorporate the available information in manifold ways, e.g., changing health aspirations or standards, adaptation processes, or susceptibility to methodological context effects (Maddox 1962; Tornstam 1975; Idler 1993; Krause & Jay 1994; Sprangers & Schwartz 1999; Simon et al. 2005; Knäuper et al. 2007; Spuling et al. 2017). Further, the cultural background, country of origin, or language of the interview can influence what weight is placed on which health domains or which answer is chosen for reasons such as a varying access to (health) care, a different interpretation of the question or value labels used, or (culturally based) country-specific response styles (Bardage et al. 2005; Jürges 2007; Viruell-Fuentes et al. 2011).

Analytical Model

[*Figure 2 here, portrait, one column*]

To enable an empirical test of parts of the general cognitive model of the response process, we further developed an analytical model that mainly focuses on the second step, i.e., generating a health rating. A depiction of this model can be seen in Figure 2. This model states, in line with the model of the cognitive process, that respondents choose and recall knowledge pertinent to their health from memory and incorporate it into one overall rating. For a greater clarity and comparability, we assign all health information

to five general types or domains: functioning, diseases, pain, depression, and behavior.

(Physical) *functioning* represents how well the respondents function in their daily lives or how limited they are, respectively, as well as their general fitness (e.g., Suchman et al. 1958; Barsky et al. 1992; Schulz et al. 1994; Jylhä et al. 1998; Benyamini et al. 1999; Leinonen et al. 1999; Quinn et al. 1999; Pinquart 2001; Leinonen 2002; Simon et al. 2005; Liang et al. 2007; Shooshtari et al. 2007; Nakano 2014). This can, e.g., be measured via self-reports of limitations or via performance tests. The aspect *diseases* serves as a general category for all (chronic) diseases and health conditions that can be diagnosed and are known to the respondent (e.g., Tornstam 1975; Segovia et al. 1989; Fylkesnes & Førde 1991, 1992; Schulz et al. 1994; Jylhä et al. 1998; Kivinen et al. 1998; Cott et al. 1999; Leinonen et al. 1999; Quinn et al. 1999; Goldberg et al. 2001; Leinonen 2002; Mellner & Lundberg 2003; Simon et al. 2005; Singh-Manoux et al. 2006; Shooshtari et al. 2007; Nakano 2014). *Pain* and its intensity is classified here separately since it is especially salient to the respondents and can not necessarily be attributed to a specific health problem (e.g., Tornstam 1975; Idler 1993; Cott et al. 1999; Shooshtari et al. 2007). The category of *depression* comprises all issues connected to mental health, especially depressive symptoms, diagnosed depression, or intake of medication against depression or anxiety as an additional objective signal for the respondents (e.g., Kivinen et al. 1998; Leinonen et al. 1999; Quinn et al. 1999; Pinquart 2001; Goldberg et al. 2001; Leinonen 2002; Schnittker 2005; Han & Jylhä 2006; Nakano 2014; Spuling et al. 2015). Lastly, *behavior* is an additional category representing risk behaviors that are known to the respondents to have adverse health effects such as smoking or being overweight (e.g., Manderbacka et al. 1999; Månsson & Merlo 2001; Imai et al. 2008; Cotter & Lachman 2010; Zajacova & Burgard 2010; Wang & Arah 2015; Noh et al. 2017; Tang et al. 2017). These might be taken as an indicator of one's health status even if they do not affect their health yet.

The incorporation of information on these five domains is, however, likely modified by aspects like gender, age, and country. Reasons can, for example, lay in group-specific

health-reporting norms, health aspirations, reference frames, or their culture (e.g., how to view or talk about health and illness) or simply the language. This aspect represents both the potential differential choosing of reference frames and for group-specific ways to evaluate one's health and is, in our analysis, accounted for by using separate models for each subgroup.

Previous Studies: Relative Importance and Group-Differences

The Relative Importance of Health Indicators

Even though there are a lot of studies investigating the influence of health indicators on SRH in one way or another, systematic approaches to examine the relative importance of indicators or health domains are fairly scarce. Yet, evidence from this line of research is vital to working with SRH in order to know what it actually measures. The first study examining this subject matter was conducted by Tornstam (1975). In his paper he found that aches and serious diseases were the most relevant aspects when explaining SRH. Barsky et al. (1992) however, studying hospital patients, found only restrictions of the functional status as a significant health-related determinant of SRH with somatization and hypochondriasis being much more important. Quinn et al. (1999) found in their article that physical health, comprising both functioning, disease, is more important for SRH than mental health while Ratner et al. (1998) found only physical health to be relevant for SRH. Shooshtari et al. (2007) found that while all aspects considered in the present paper (i.e., functioning, diseases, pain, depression, and behavior) are relevant for rating one's health, functioning and diseases were the most important.

One of the few studies that explicitly sought to explore the relative importance of health indicators on SRH was conducted by Singh-Manoux et al. (2006). They used data from the Whitehall II and Gazel cohort study in order to quantify the contribution

of different aspects of health towards the explanation of SRH. Their conclusion was that health indicators such as symptoms, longstanding illnesses, health problems, and mobility accounted for 35–41 percent of SRH's variance, depending on the data set, while other aspects were less relevant. Using rather broad categories of variables, they also showed that physical health was the most important health domain, followed by mental health, and health behavior.

Differences in How Groups of Respondents Rate Their Health

There are not many studies on group differences in health ratings yet, even though systematic differences might be expected, e.g. in terms of gender, age, and country with gender being the most prominently researched in the pertinent literature. For example, men might put greater weight on physical functioning while women signify the importance of the absence of illnesses (Peersman et al. 2012) or only women might be influenced by depressive symptoms (Leinonen et al. 1999). However, some, usually European, studies did not find any notable differences in the rating behavior by men and women (Jylhä et al. 1998, Undén & Elofsson 2006; Zajacova et al. 2017).

A second aspect that might be relevant for how respondents rate their health might be their age since the aspirational level of health decreases with age, meaning that older respondents potentially have lower expectations or are more tolerant of health problems than younger respondents (Tornstam 1975) or have different frames or reference, e.g., specific health problems, physical functioning, or health behaviors (Krause & Jay 1994). If, e.g., older persons do not take medical conditions or diseases into account when evaluating their health in a survey interview (while younger respondents do), their responses are not directly comparable regarding health. The same applies to other groups like men and women or respondents from different countries. If they would base their appraisal on different sets of indicators, weight them differently, or generally show different response behaviors, researchers could not directly compare their health measurements. The already cited study by Tornstam (1975), for example, found overall weaker negative effects of

adverse health due to lower health aspirations in older age. Other studies found differences in the relevance of various health domains, such as: symptoms and mental health being more relevant for young-old (51–55) while old-old (71–75) lay more weight on chronic diseases (Jylhä et al. 1986); mental health being more relevant for older respondents while for younger respondents physical health/functioning and chronic diseases were more important (Pinquart 2001; Schnittker 2005; Jylhä et al. 2001); younger people in general using more diverse aspects in rating their health (Shooshtari et al. 2007); the importance of mental health being stable while medical conditions and functioning losing in importance with age (French et al. 2012); chronic conditions being stable and mental health being more relevant in younger cohorts (Spuling et al. 2015); and behavior being less relevant for SRH in older age (Manderbacka et al. 1999).

The third aspect possibly responsible for different rating behavior explored in this paper is the country of residence or language of the interview. The pertinent literature can be broadly classified in two groups: studies that find substantial country- or language-differences in how health indicators are related to SRH (Viruell-Fuentes et al. 2011) and those that do not (Jylhä et al. (1998); Bardage et al. (2005); Verropoulou 2009). Interestingly, the latter studies were conducted with European data while the first one paper utilized US-data.

This short overview illustrates the dire need for more studies aimed at systematically quantifying and comparing the relative importance of different health domains commonly associated with SRH across different groups. While we can expect functioning and diseases, as classical determinants of subjective health, to be of great importance for SRH, the role of pain, mental health/depression, and behavior is rather unclear. The same, maybe to an even greater extent, is true for group-differences in rating behavior. The state of research for all three aspects (i.e., gender, age, and country/language) is inconclusive regarding the type of effect to be expected (age) or whether to expect any meaningful differences at all (gender and country). This ambiguity further demonstrates the necessity of research on this subject.

Data and Method

Data

For our analysis we use data of the 5th wave of the Survey of Health, Ageing and Retirement in Europe (SHARE) from 2015 comprising a wide array of health information for more than 64,000 respondents from 15 European countries (Austria, Belgium, Czech Republic, Denmark, Estonia, France, Germany, Italy, Israel, Luxembourg, Netherlands, Slovenia, Spain, Sweden, Switzerland) aged 50 years or older. The multivariate analyses comprises information from 61,027 respondents (33,576 women and 27,451 men).

Analysis and Measurement

In order to implement the analysis according to our analytical model, we used linear regression models with SRH as the dependent and all aforementioned health-related variables as independent variables. The independent variables were blocked according to the proposed five types of health information: functioning, diseases, pain, depression, and behavior:

Self-rated health As a measure for SRH, we used the question 'Would you say your health ...?' with the response options 'Excellent, Very good, Good, Fair, Poor'. We treated SRH quasi-metric which enables a linear regression analysis (Manderbacka et al. 1998; Leinonen 2002). In SHARE, this question is not preceded by any other health-specific question, implying a free interpretation of the meaning of health by the respondents.

Functioning In order to operationalize (physical) functioning, we used four different aspects comprising both self-reports and physical performance tests. As for the latter, we used both a measure of grip strength and the chair stand (Cooper et al. 2011) to explain SRH. To this end, we generated two dummy-variables that represent being in the lowest performance quartile of one's own gender (weakest for grip strength, slowest for chair stand) and not having a measurement taken. Item-nonresponse for these variables can be

seen as informative nonresponse since it can be assumed that it means respondents were (deemed or feeling) unfit to participate in the measurement (Herzog & Rodgers 1992). Since this would mean that the item-nonresponse is missing not at random (MNAR), a simple exclusion would bias the results (Gardette et al. 2007). As for the self-reports, we used count variables for the number of restrictions in (I)ADL (13 items, e.g., *dressing, including putting on shoes and socks* or *shopping for groceries*) and mobility (10 items, e.g., *walking 100 metres* or *stooping, kneeling, or crouching*). In order to account for the nonlinear association of SRH and these count variables, we transformed them utilizing an inverse hyperbolic sine transformation ($log(x_i + (x_i^2 + 1)^{.5})$). This transformation is similar to a logarithmic transformation but allows the transformation of zero-values (Burbidge et al. 1988; Zhang et al. 2000) which are common for these (I)ADL and mobility restrictions.

Diseases Diseases were operationalized via a count-variable of different conditions and diagnoses (17 items like *high blood cholesterol* or *cancer* and including *other conditions, not yet mentioned*) and a general question whether or not the respondent suffered from a chronic or long-term health problem. The count variable was also transformed via an inverse hyperbolic sine transformation.

Pain In order to measure pain, we included a single general question whether or not the respondent was troubled with pain at the time of the interview. For respondents experiencing pain, this question was supplemented with information on whether they consider the pain to be mild, moderate, or severe.

Depression Depression was measured through the number of depressive symptoms on the Euro-D scale (Prince et al. 1999) which was also transformed like the other count variables. Additionally, we included a general question regarding taking medication against depression or anxiety in our model.

Behavior The measurement of behavior was twofold to depict two common types of risky health behavior or its consequence, respectively: smoking and non-normal weight. Smoking was captured with a question whether the respondent currently smokes while

the body-mass-index (BMI) was calculated as the self-reported weight (in kg) of the respondent divided by their squared self-reported height (in m). To account for the nonlinear relationship of BMI and health, we used dummy-variables for being underweight, overweight, and being adipose.[2]

We firstly applied the general model to men and women separately, while also quantifying the contribution to R^2 by each of the five health domains. A discussion of these results will be followed by a figure showing the contribution of the five types of health information by gender and age-group in order to examine age-differences in how European respondents rate their health. Lastly, we will compare the relative amount of explained variance for each of the 15 analyzed European countries to demonstrate the extent of country-specific health ratings and thus the comparability of self-rated health across countries.

In order to assess the health domains' contributions to R^2, we conducted dominance analyses with the Stata-module domin (Luchman 2013). This approach compares R^2 for all possible subsets of variables or variable-sets in order to determine the variance explained by them or, in other words, their contribution to overall R^2 (Budescu 1993; Luchman 2014, 2015). To compare these contributions, we estimated confidence intervals through bootstrapping (10,000 samples for each model).

Results

Table 1 shows the regression results separately by gender. The models, comprising extensive health information, explained 49 percent of SRH's variance for women and 45 percent for men, documenting that SRH is heavily reliant on the health information known by the respondent and potentially ascertainable in surveys. This supports its use as a simple and inclusive measure of generic health representing a host of health indicators.

[2]We explicitly did not include alcohol consumption and physical activity, which are typically seen as health-related behaviors, in this health domain. The reason for this is that both are not only (subjective) reasons for bad health but they are also strongly restricted by it (i.e., not being able to drink alcohol or exercise due to medication or functional limitations).

It can be seen that all measured health indicators significantly and negatively influenced SRH for women. The same was true for men with the exception that being underweight did not influence European men in their health ratings when controlling for other health-related factors. All coefficients except the behavioral variables were rather similar in size between genders. Further, a missing measurement on each performance test was significantly related to SRH, suggesting informative nonresponse.

[*Table 1 here, portrait, one columns*]

[*Figure 3 here, portrait, two columns*]

As can be seen from Table 3 the most relevant health domains, functioning and diseases, both accounted for around a third of the variance explained by the model for both men and women. These health domains were followed by pain with 17 percent and 15 percent respectively and then depression with approximately twelve percent for both gender on overall R^2. The least relevant health domain in our analyses were behavioral variables with four percent for women and two percent for men. There was a slight but significant gender difference in the share of explained variance by diseases which was greater for men than for women. Moreover, there was a significant difference between the share of functioning and diseases on R^2 in women but not in men. The opposite is true for pain and behavior since these health domains accounted for a greater share on R^2 in women than in men. Overall, however, we would argue that the gender differences, albeit significant, were rather small and that European women and men are remarkably similar in how they rate their health, replicating results already shown by Zajacova et al. (2017) with US data.

[*Figure 4 here, portrait, two columns*]

[*Table 2 here, portrait, one column*]

A comparison of the three age-groups separately by gender regarding the health domain's share on explained variance overall can be seen in Figure 4. While there were, with the exception of pain, no meaningful differences between genders, it shows that there were clear and consistent differences in how the relevance of functioning, diseases, and behavior

differs between age-groups. Even though the share of explained variance by functioning immensely increased over the three age-groups, this share decreased for the relevance of diseases and, to a lesser extent, behavior. Depression, however, was remarkably stable in its contribution to R^2, at least for women.

These results show differences as well as congruencies in how people from different age-groups rate their health. It is also worth noting that there were, with a minor exception for disease, no significant gender differences in the shares on R^2 in the youngest age-group and none at all for the oldest respondents. This corroborates the previous impression of similarities of rating behavior between genders overall. Generally, overall R^2, as can be seen from Table 2, did not vary too much between age-groups or genders.

[*Figure 5 here, portrait, two columns*]

[*Table 3 here, portrait, two columns*]

Figure 5 shows the share of explained variance by the five types of health information for each European country used in our analysis. Since gender differences were relatively low in the previous analyses (and were also in this case), only overall results are shown. The countries were sorted by the amount of variance explained by functioning. Overall, country differences appear rather small: The general ranking of importance, with functioning and diseases being most relevant, pain and then depression in the middle, and behavior being the least relevant, held true for every single country in our analysis. Although there was some variation in overall R^2 reported in Table 3, this variation did not seem to be systematic in any way.

Nevertheless, there were some signs of country-specific health rating behavior in terms of the relative contributions of different dimensions to overall R^2. While functioning was more important than diseases for Estonians, the opposite was true for Swedes, Germans, and Italians. Further, pain was more important for respondents from the Czech republic, Spain, and Germany, than depression with rather large differences for Czechs and Germans. Notably, in all other countries there were no significant differences in the importance of functioning and disease, or pain and depression respectively. This suggests

an overall quite similar rating-process of general health in these countries.

Conclusion and Discussion

The purpose of this paper was to identify, quantify, and compare the relevance of five different health domains: functioning, diseases, pain, depression, and risk-behavior. In order to do so, we analyzed data from the more than 61,000 respondents aged 50–90 living in fifteen European countries collected in the 5^{th} wave of the SHARE. The explanatory power of our models is relatively high since almost half of the variance of SRH can be explained with these health-related data. This finding corroborates early findings of "the centrality of objective health status in explaining self-assessments of health" (Maddox 1962: 183). Apparently (and unsurprisingly), SRH is based to a large extent on health information known to the respondent. Still, it should be noted that also half of SRH's variance is not related to SHARE's rather comprehensive health data, leaving much room for differences due to health knowledge, non-health related idiosyncrasies, and even bias. Especially the influence of non health-related aspects, such as respondent or survey or interviewer characteristics, potentially biasing SRH merits further investigation.

Interestingly, missing values for performance tests turned out to be negatively related to overall SRH. This can be explained by the fact that missing performance tests are related to health because the interviewer or respondent deems the respondent in too bad health to participate in the test, thus creating nonresponse. This indicates that missingness for these variables is indeed MNAR and therefore excluding persons without measurement would presumably bias results of health-related research.

One main result of this paper is that functioning an diseases are by far the two most relevant health domains when it comes to rating one's health. Ranking by contribution to explaining SRH, they are followed by pain and depression and then by risk-behavior which appears to be only of subordinate importance for SRH-scores. As can be seen from our subgroup analyses and consistent with earlier research, there were no marked or systematic

differences by gender (e.g., Jylhä et al. 1998; Undén & Elofsson 2006; Zajacova et al. 2017) or (European) country (e.g., Jylhä et al. 1998; Bardage et al. 2005; Verropoulou 2009).

Yet, there were strong and systematic differences by age-group in that functioning explained more of SRH's variance in older age-groups while the opposite was true for both diseases and behavior. While younger cohorts of this 50+ population appear to base their health more strongly on diagnoses of diseases and health conditions, older respondents lay more weight on how well they function. Behavior, in general, appears to be only relevant for respondents younger than 80 which might also reflect selective mortality. This suggests that respondents of different ages indeed have a different understanding of what constitutes 'health' and how to rate it. This highlights the importance of taking age-specific response behavior into account when using and comparing SRH-scores across a diverse population. Nevertheless, our analysis was limited to European countries and respondents aged 50–90 and as such cannot be generalized to other populations or contexts. Further research with other populations are advisable to attain a better understanding of the mechanisms underlying SRH.

Bibliography

Anderson, Norman H., 1971: Integration Theory and Attitude Change. Psychological Review 78 (3): 171–206.

Bardage, Carola, Saskia M. F. Pluijm, Nancy L. Pedersen, Dorly J. H. Deeg, Marja Jylhä, Marianna Noale & Tzvia Blumstein Angel Otero, 2005: Self-Rated Health among Older Adults: A Cross-National Comparison. European Journal of Ageing 2 (2): 149–158.

Barsky, Arthur J., Paul D. Cleary & Gerald L. Klerman, 1992: Determinants of Perceived Health Status of Medical Outpatients. Social Science & Medicine 34 (10): 1147–1154.

Benyamini, Yael, 2011: Why Does Self-Rated Health Predict Mortality? An Update on Current Knowledge and a Research Agenda for Psychologists. Psychology & Health 26 (11): 1407–1413.

Benyamini, Yael, Elaine A. Leventhal & Howard Leventhal, 1999: Self-Assessments of Health: What Do People Know That Predicts Their Mortality? Research on Aging 21 (3): 477–500.

Budescu, David V., 1993: Dominance Analysis: A New Approach to the Problem of Relative Importance of Predictors in Multiple Regression. Psychological Bulletin 114 (3): 542–551.

Burbidge, John B., Lonnie Magee & A. Leslie Robb, 1988: Alternative Transformations to Handle Extreme Values of the Dependent Variable. Journal of the American Statistical Association 83 (401): 123–127.

Caroli, Eve & Lexane Weber-Baghdiguian, 2016: Self-Reported Health and Gender: The Role of Social Norms. Social Science & Medicine 153: 220–229.

Cheng, Sheung-Tak, Helene Fung & Alfred Chan, 2007: Maintaining Self-Rated Health

Through Social Comparison in Old Age. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences 62 (5): P277–P285.

Cooper, Rachel, Diana Kuh, Cyrus Cooper, Catharine R. Gale, Debbie A. Lawlor, Fiona Matthews, Rebecca Hardy & The Falcon and Halcyon Study Teams, 2011: Objective Measures of Physical Capability and Subsequent Health: A Systematic Review. Age and Ageing 40 (1): 14–23.

Cott, Cheryl A., Monique A. Gignac & Elizabeth M. Badley, 1999: Determinants of Self Rated Health for Canadians with Chronic Disease and Disability. Journal of Epidemiology and Community Health 53 (11): 731–736.

Cotter, Kelly A. & Margie E. Lachman, 2010: Psychosocial and Behavioural Contributors to Health: Age-Related Increases in Physical Disability Are Reduced by Physical Fitness. Psychology & Health 25 (7): 805–820.

French, Davina J., Kerry Sargent-Cox & Mary A. Luszcz, 2012: Correlates of Subjective Health Across the Aging Lifespan: Understanding Self-Rated Health in the Oldest Old. Journal of Aging and Health 24 (8): 1449–1469.

Fylkesnes, Knut & Olav Helge Førde, 1991: The Tromsø Study: Predictors of Self-Evaluated Health – Has Society Adopted the Expanded Health Concept? Social Science & Medicine 32 (2): 141–146.

Fylkesnes, Knut & Olav Helge Førde, 1992: Determinants and Dimensions Involved in Self-Evaluation of Health. Social Science & Medicine 35 (3): 271–279.

Garbarski, Dana, 2016: Research in and Prospects for the Measurement of Health Using Self-Rated Health. Public Opinion Quarterly 80 (4): 977–997.

Garbarski, Dana, Nora Cate Schaeffer & Jennifer Dykema, 2015: The Effects of Response Option Order and Question Order on Self-Rated Health. Quality of Life Research 24 (6): 1443–1453.

Gardette, Virginie, Nicola Coley, O. Toulza & S. Andrieu, 2007: Attrition in Geriatric Research: How Important Is It and How Should It Be Dealt With? The Journal of Nutrition Health and Aging 11 (3): 265–271.

Goldberg, P., Alice Guéguen, A. Schmaus, J.-P. Nakache & Marcel Goldberg, 2001: Longitudinal Study of Associations Between Perceived Health Status and Self Reported Diseases in the French Gazel Cohort. Journal of Epidemiology & Community Health 55 (4): 233–238.

Han, Beth & Marja Jylhä, 2006: Improvement in Depressive Symptoms and Changes in Self-Rated Health among Community-Dwelling Disabled Older Adults. Aging & Mental Health 10 (6): 599–605.

Herzog, A. Regula & Willard L. Rodgers, 1992: The Use of Survey Methods in Research on Older Americans. S. 60–90 in: Robert B. Wallace & Robert F. Woolson (Eds.), The Epidemiologic Study of the Elderly. New York: Oxford University Press.

Idler, Ellen L., 1993: Perceptions of Pain and Perceptions of Health. Motivation and Emotion 17 (3): 205–224.

Idler, Ellen L. & Yael Benyamini, 1997: Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. Journal of Health and Social Behavior 38 (1): 21–37.

Imai, Kumiko, Edward W. Gregg, Yiling J. Chen, Ping Zhang, Nathalie de Rekeneire & David F. Williamson, 2008: The Association of BMI With Functional Status and Self-Rated Health in US Adults. Obesity 16 (2): 402–408.

Jürges, Hendrik, 2007: True Health vs. Response Styles: Exploring Cross-Country Differences in Self-Reported Health. Health Economics 16 (2): 163–178.

Jylhä, Marja, 2009: What Is Self-Rated Health and Why Does It Predict Mortality? Towards a Unified Conceptual Model. Social Science & Medicine 69 (3): 307–316.

Jylhä, Marja, Esko Leskinen, Erkki Alanen, Anna-Liisa Leskinen & Eino Heikkinen, 1986: Self-Rated Health and Associated Factors Among Men of Different Ages. Journal of Gerontology 41 (6): 710–717.

Jylhä, Marja, Jack M. Guralnik, Luigi Ferrucci, Jukka Jokela & Eino Heikkinen, 1998: Is Self-Rated Health Comparable Across Cultures and Genders? Journal of Gerontology 53B (3): S144–S152.

Jylhä, Marja, Jack M. Guralnik, Jennifer Balfour & Linda P. Fried, 2001: Walking Difficulty, Walking Speed, and Age as Predictors of Self-Rated Health: The Womens Health and Aging Study. The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences 56A (10): M609–M617.

Kivinen, Paula, Pirjo Halonen, Minna Eronen & Aulikki Nissinen, 1998: Self-Rated Health, Physician-Rated Health and Associated Factors Among Elderly Men: The Finnish Cohorts of The Seven Countries Study. Age and Ageing 27 (1): 41–47.

Knäuper, Bärbel & Patricia A. Turner, 2003: Measuring Health: Improving the Validity of Health Assessments. Quality of Life Research 12 (Suppl. 1): 81–89.

Knäuper, Bärbel, Norbert Schwarz, Denise Park & Andreas Fritsch, 2007: The Perils of Interpreting Age Differences in Attitude Reports: Question Order Effects Decrease with Age. Journal of Official Statistics 23 (4): 515–528.

Krause, Neal M. & Gina M. Jay, 1994: What Do Global Self-Rated Health ltems Measure? Medical Care 32 (9): 930–942.

Lazarevič, Patrick, 2018: Ausgezeichnet, sehr gut, gut, mittelmäßig, schlecht: Theoretische Konzepte und empirische Befunde zur Erhebung des allgemeinen Gesundheitszustandes bei älteren und alten Menschen in: Martina Brandt, Jennifer Fietz, Sarah Hampel, Judith Kaschowitz, Patrick Lazarevič, Monika Reichert & Veronique Wolter (Eds.), Methoden der empirischen Alter(n)sforschung. Weinheim: Beltz (forthcoming).

Lee, Sunghee, 2015: Self-Rated Health in Health Surveys. S. 193–216 in: Timothy P. Johnson (Eds.), Handbook of Health Survey Methods. Hoboken: Wiley.

Leinonen, Raija, 2002: Self-Rated Health in Old Age: A Follow-up Study of Changes and Determinants. Jyväskylä: University of Jyväskylä.

Leinonen, Raija, Eino Heikkinen & Marja Jylhä, 1999: A Path Analysis Model of Self-Rated Health Among Older People. Aging Clinical and Experimental Research 11 (4): 209–220.

Liang, Jersey, Benjamin A. Shaw, Joan M. Bennett, Neal Krause, Erika Kobayashi, Taro Fukaya & Yoko Sugihara, 2007: Intertwining Courses of Functional Status and Subjective Health Among Older Japanese. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences 62B (5): S340–348.

Luchman, Joseph N., 2013: DOMIN: Stata Module to Conduct Dominance Analysis. Retrieved from `http://ideas.repec.org/c/boc/bocode/s457629.html`.

Luchman, Joseph N., 2014: Relative Importance Analysis With Multicategory Dependent Variables: An Extension and Review of Best Practices. Organizational Research Methods 17 (4): 452–471.

Luchman, Joseph N., 2015: Determining Subgroup Difference Importance with Complex Survey Designs: An Application of Weighted Dominance Analysis. Survey Practice 8 (4): 1–10.

Maddox, George L., 1962: Some Correlates of Differences in Self-Assessment of Health Status Among the Elderly. Journal of Gerontology 17 (2): 180–185.

Manderbacka, Kristiina, Eero Lahelma & Pekka Martikainen, 1998: Examining the Continuity of Self-Rated Health. International Journal of Epidemiology 27 (2): 208–213.

Manderbacka, Kristiina, Olle Lundberg & Pekka Martikainen, 1999: Do Risk Factors and Health Behaviours Contribute to Self-Ratings of Health? Social Science & Medicine 48 (12): 1713–1720.

Månsson, Nils-Ove & Juan Merlo, 2001: The Relation between Self-Rated Health, Socioeconomic Status, Body Mass Index and Disability Pension among Middle-Aged Men. European Journal of Epidemiology 17 (1): 65–69.

Mellner, Christin & Ulf Lundberg, 2003: Self- and Physician-Rated General Health in Relation to Symptoms and Diseases Among Women. Psychology, Health & Medicine 8 (2): 123–134.

Nakano, Ai, 2014: The Relationship between Mental Health and Self-Rated Health in Older Adults. Graduate School of Economics Kobe University: Discussion Paper No. 1423.

Noh, Jin-Won, Jinseok Kim, Youngmi Yang, Jumin Park, Jooyoung Cheon & Young Dae Kwon, 2017: Body Mass Index and Self-Rated Health in East Asian Countries: Comparison among South Korea, China, Japan, and Taiwan. PLOS ONE 12 (8): 1–10.

Peersman, Wim, Dirk Cambier, Jan De Maeseneer & Sara Willems, 2012: Gender, Educational and Age Differences in Meanings That Underlie Global Self-Rated Health. International Journal of Public Health 57 (3): 513–523.

Pinquart, Martin, 2001: Correlates of Subjective Health in Older Adults: A Meta-Analysis. Psychology and Aging 16 (3): 414–426.

Prince, Martin J., Friedel Reischies, A.T.F. Beekman, Rebecca Fuhrer, C. Jonker, Sirkka-Liisa Kivelä, Brian A. Lawlor, Antonio Lobo, H. Magnusson, M.M. Fichter, Herman van Oyen, Marc Roelands, Ingmar Skoog, Cesare Turrina & John R.M. Copeland, 1999: Development of the EURO-D Scale: A European Union Initiative to Compare Symptoms of Depression in 14 European Centres. British Journal of Psychiatry 174 (4): 330–338.

Quinn, Mary Ellen, Mary Ann Johnson, Leonard W. Poon & Peter Martin, 1999: Psychosocial Correlates of Subjective Health in Sexagenarians, Octogenarians, and Centenarians. Issues in Mental Health Nursing 20 (2): 151–171.

Ratner, Pamela A., Joy L. Johnson & Bonnie Jeffery, 1998: Examining Emotional, Physical, Social, and Spiritual Health as Determinants of Self-Rated Health Status. American Journal of Health Promotion 12 (4): 275–282.

Schnittker, Jason, 2005: When Mental Health Becomes Health: Age and the Shifting Meaning of Self-Evaluations of General Health. The Milbank Quarterly 83 (3): 397–423.

Schulz, Richard, Maurice Mittelmark, Richard Kronmal, Joseph F. Polak, Calvin H. Hirsch, Pearl German & Jamila Bookwala, 1994: Predictors of Perceived Health Status in Elderly Men and Women. Journal of Aging and Health 6 (4): 419–447.

Schwarz, Norbert, 1999: Self-Reports: How Questions Shape the Answers. American Psychologist 54 (2): 93–105.

Segovia, Jorge, Roy F. Bartlett & Alison C. Edwards, 1989: An Empirical Analysis of the Dimensions of Health Status Measures. Social Science & Medicine 29 (6): 761–768.

Shooshtari, Shahin, Verena Menec & Robert Tate, 2007: Comparing Predictors of Positive and Negative Self-Rated Health Between Younger (25-54) and Older (55+) Canadian Adults. Research on Aging 29 (6): 512–554.

Simon, Jeanette G., Josien B. De Boer, Inez M.A. Joung, Hans Bosma & Johan P. Mackenbach, 2005: How Is Your Health in General? A Qualitative Study on Self-Assessed Health. European Journal of Public Health 15 (2): 200–208.

Singh-Manoux, Archana, Pekka Martikainen, Jane Ferrie, Marie Zins, Michael Marmot & Marcel Goldberg, 2006: What Does Self Rated Health Measure? Results from the

British Whitehall II and French Gazel Cohort Studies. Journal of Epidemiology and Community Health 60 (4): 364–372.

Sprangers, Mirjam A.G. & Carolyn E. Schwartz, 1999: Integrating Response Shift into Health-Related Quality of Life Research: A Theoretical Model. Social Science & Medicine 48 (11): 1507–1515.

Spuling, Svenja M., Susanne Wurm, Clemens Tesch-Römer & Oliver Huxhold, 2015: Changing Predictors of Self-Rated Health: Disentangling Age and Cohort Effects. Psychology and Aging 30 (2): 462–474.

Spuling, Svenja M., Oliver Huxhold & Susanne Wurm, 2017: Predictors of Self-Rated Health: Does Education Play a Role Above and Beyond Age? Journals of Gerontology: Psychological Sciences 72 (3): 415–424.

Strack, Fritz & Leonard L. Martin, 1987: Thinking, Judging, Communicating: A Process Account of Context Effects in Attitude Surveys. S. 123–148 in: Hans-J. Hippler, Norbert Schwarz & Seymour Sudman (Eds.), Social Information Processing and Survey Methodology. New York: Springer.

Suchman, Edward A., Bernard S. Phillips & Gordon F. Streib, 1958: An Analysis of the Validity of Health Questionnaires. Social Forces 36 (3): 223–232.

Tang, Kun, Yingxi Zhao & Chunyan Li, 2017: The Association Between Self-Rated Health and Different Anthropometric and Body Composition Measures in the Chinese Population. BMC Public Health 17 (317): 1–9.

Tornstam, Lars, 1975: Health and Self-Perception: A Systems Theoretical Approach. The Gerontologist 15 (3): 264–270.

Tourangeau, Roger, 1984: Cognitive Sciences and Survey Methods. S. 73–100 in: Thomas B. Jabine, Miron L. Straf, Judith M. Tanur & Roger Tourangeau (Eds.), Cog-

nitive Aspects of Survey Methodology: Building a Bridge Between Disciplines. Washington: National Academy Press.

Tourangeau, Roger, Kenneth A. Rasinski & Norman Bradburn, 1991: Measuring Happiness in Surveys: A Test of the Subtraction Hypothesis. Public Opinion Quarterly 55 (2): 255–266.

Tversky, Amos & Daniel Kahneman, 1973: Availability: A Heuristic for Judging Frequency and Probability. Cognitive Psychology 5 (2): 207–232.

Undén, Anna-Lena & Stig Elofsson, 2006: Do Different Factors Explain Self-Rated Health in Men and Women? Gender Medicine 3 (4): 295–308.

Verropoulou, Georgia, 2009: Key Elements Composing Self-rated Health in Older Adults: A Comparative Study of 11 European Countries. European Journal of Ageing 6 (3): 213–226.

Viruell-Fuentes, Edna A., Jeffrey D. Morenoff, David R. Williams & James S. House, 2011: Language of Interview, Self-Rated Health, and the Other Latino Health Puzzle. Research and Practice 101 (7): 1306–1313.

Wang, Aolin & Onyebuchi A. Arah, 2015: Body Mass Index and Poor Self-Rated Health in 49 Low-Income and Middle-Income Countries, By Sex, 2002–2004. Preventing Chronic Disease 12 (E133): 1–4.

World Health Organization, 2006: Constitution of the World Health Organization.

Zajacova, Anna & Sarah A. Burgard, 2010: Body Weight and Health from Early to Mid-Adulthood: A Longitudinal Analysis. Journal of Health and Social Behavior 51 (1): 92–107.

Zajacova, Anna, Snehalata Huzurbazar & Megan Todd, 2017: Gender and the Structure of Self-Rated Health Across the Adult Life Span. Social Science & Medicine 187: 58–66.

Zhang, Mingliang, John C. Fortney, John M. Tilford & Kathryn M. Rost, 2000: An Application of the Inverse Hyperbolic Sine Transformation: A Note. Health Services & Outcomes Research Methodology 1 (2): 165–171.

Appendix

Tables

Table 1: Results from OLS-Regression Explaining Self-Rated Health (b-Coefficients)

	Women	Men
Functioning		
Grip strength (RC: middle 50%)		
No measurement	-0.10^{**}	-0.18^{***}
Stronger 25%	-0.10^{**}	-0.18^{***}
Weaker 25%	0.07^{**}	0.10^{**}
Chair stand (RC: middle 50%)		
No measurement	-0.27^{***}	-0.20^{***}
Faster 25%	0.13^{***}	0.09^{*}
Slower 25%	-0.11^{***}	-0.02
Number of restrictions in daily life[a]	-0.08^{***}	-0.03
Number of restrictions in mobility[a]	-0.15^{***}	-0.20^{***}
Diseases		
Chronic diseases (RC: none)	-0.33^{***}	-0.43^{***}
Number of diagnoses[a]	-0.28^{***}	-0.28^{***}
Pain (RC: none)		
Mild	-0.11^{**}	-0.08
Moderate	-0.16^{***}	-0.19^{***}
Severe	-0.29^{***}	-0.32^{***}
Depression		
Medication for depression (RC: no)	-0.16^{***}	-0.14^{**}
Number of depressive symptoms[a]	-0.20^{***}	-0.20^{***}
Behavior		
BMI (RC: normal ($18.5 \leq BMI \leq 25$))		
Underweight (BMI < 18.5)	-0.20^{***}	0.43
Overweight ($25 \leq$ BMI < 30)	-0.06^{**}	-0.04
Adipose (BMI ≥ 30)	-0.12^{***}	-0.10^{**}
Current Smoker (RC: no)	-0.07^{*}	-0.12^{***}
Adj. R^2	0.50	0.45
n	33,576	27,451

[a]Inverse hyperbolic sine transformation to account for nonlinear relationship
$^{+}p \leq .1$ $^{*}p \leq .05$ $^{**}p \leq 0.01$ $^{***}p \leq 0.001$

Table 2: Adjusted R^2 and Number of Cases for Separate Models by Gender and Age

	Women			Men		
	50–64	65–79	80–90	50–64	65–79	80–90
Adj. R^2	0.45	0.47	0.46	0.42	0.42	0.44
n	16,064	13,758	3,754	12,611	11,951	2,889

Table 3: Adjusted R^2 and Number of Cases for Separate Models by Country

	IL	EE	LU	AT	ES	DK	SE	FR	DE	BE	SI	IT	CZ	CH	NE
Adj. R^2	0.58	0.45	0.49	0.51	0.53	0.51	0.49	0.46	0.48	0.44	0.45	0.47	0.42	0.39	0.38
n	2,106	5,265	1,532	3,936	5,729	3,852	4,318	4,256	5,454	5,241	2,743	4,440	5,254	2,920	3,981

Figures

Figure 1: Cognitive Model to Explain the Process of Health-Ratings

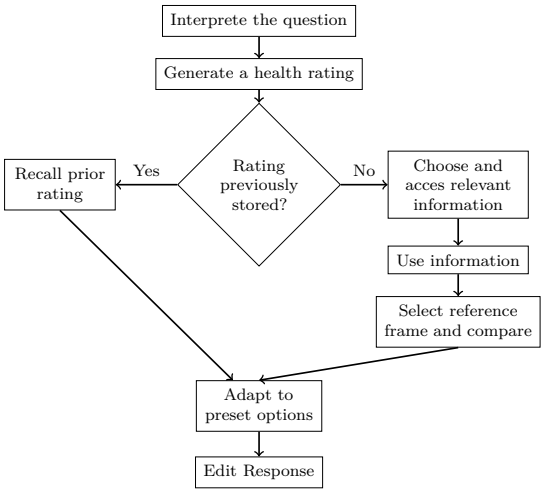

Figure 2: Analytical Model for Explaining SRH by health factors

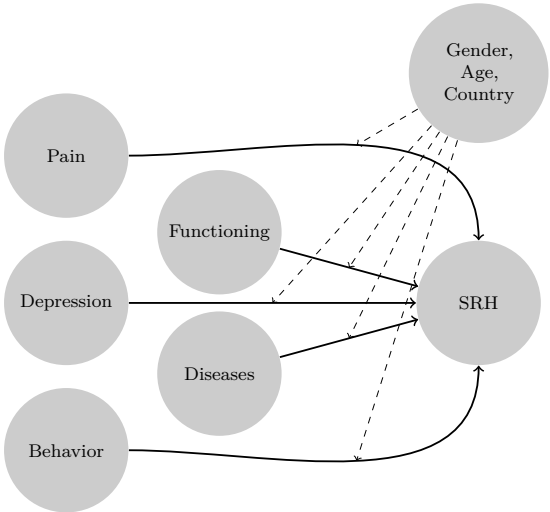

Figure 3: Amount of Explained Variance Accounted for by Health Domain by Gender (95%-Confidence Intervals)

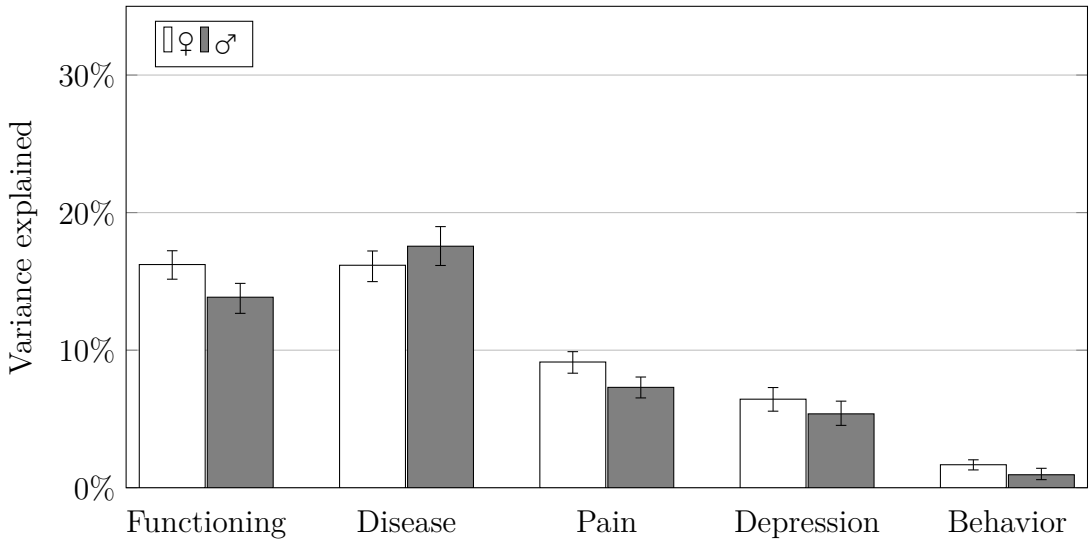

Figure 4: Amount of Explained Variance Accounted for by Health Domain by Gender and Age-Group (95%-Confidence Intervals)

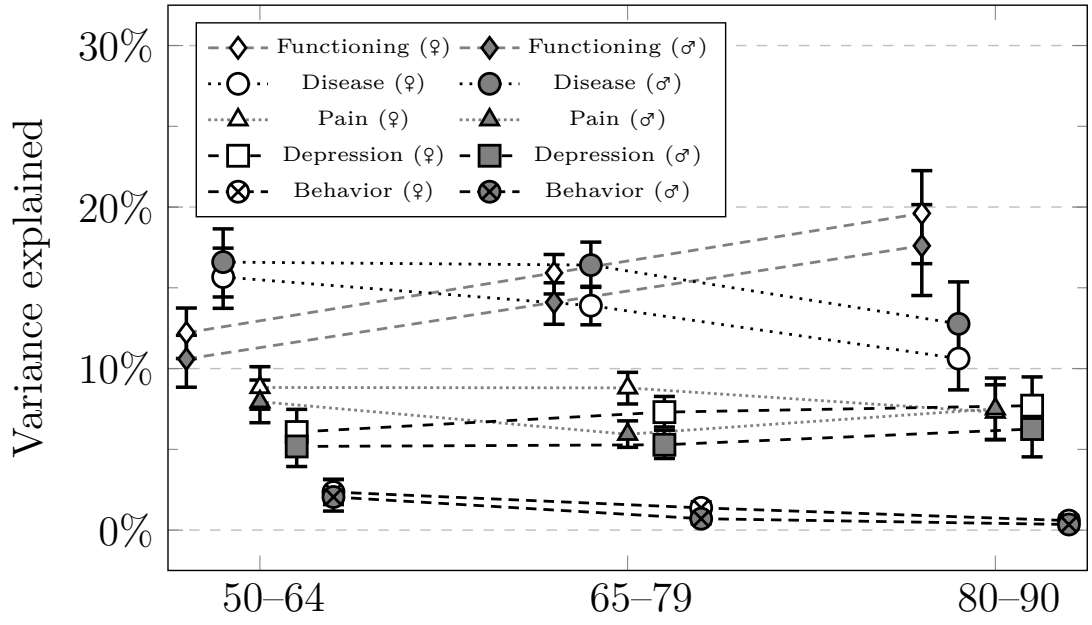

Figure 5: Amount of Explained Variance Accounted for by Health Domain by Country (95%-Confidence Intervals)