

Embedding as a pitfall for survey-based welfare indicators: Evidence from an Experiment

Clemens Hetschko
Louisa von Reumont
Ronnie Schöb

School of Business & Economics

Discussion Paper

Economics

2017/8

Embedding as a pitfall for survey-based welfare indicators: Evidence from an Experiment

Clemens Hetschko

Freie Universität Berlin, Institute for Employment Research (IAB), CESifo

Louisa von Reumont*

DIW Econ, Freie Universität Berlin

Ronnie Schöb

Freie Universität Berlin, ifo Dresden, CESifo

This version: March 2017

Abstract

How can we assess the welfare of a society, its evolution over time and predict its change due to particular policy interventions? One way is to use survey-based welfare indicators such as the OECD Better Life Index. It invites people to weight a variety of quality of life indicators according to their individual preferences. 11 broad dimensions aggregate these indicators. Our experiment shows that people do not provide consistent ratings across differently labelled dimensions that embed the same indicators. They also do not adjust the rating of equally named dimensions changing sets of indicators. These results show that survey-based measures might suffer from strong embedding effects and, as a result, may fail to measure citizens' true preferences for the indicators.

JEL Classification Codes: B41, C43, C83, I31

Keywords: Beyond GDP, welfare measurement, survey-based welfare indicators, OECD Better Life Index, embedding effect

**Corresponding author.* Address. Louisa von Reumont, DIW Econ, Mohrenstraße 58, 10117 Berlin, Germany, e-mail: lreumont@diw-econ.de. Disclaimer: Louisa von Reumont has worked on this study in her personal capacity. The opinions expressed in this article are the authors' own and do not reflect the view of DIW Econ.

Acknowledgements. We are deeply indebted to Stephen Lumenta who programmed and implemented the experimental web-application. We thank Malte Preuss, Tobias Wolf, Katarina Bjelajac Witzheller, Daniel Nachtigall, Andreas Knabe, Marcel Thum, Robert Schwager, Nadine Riedel, Robert Fenge, Kerstin Schneider and Alfons Weichenrieder for their support in recruiting participants for the experiment. For very helpful comments on an earlier draft, we are grateful to Michael Ahlheim, Adrian Chadi, Jeremy Edwards, Andreas Knabe, Kai Ludwigs, Milena Nikolova and Marcel Thum as well as participants of workshops at the Max Planck Institute for Tax Law and Public Finance (Munich, 2016), the Institute for Labor Law and Industrial Relations in the EU (Trier, 2016) and the Research Seminar in Economics at the Freie Universität Berlin (2016).

1. Introduction

The welfare of nations is difficult to assess. For decades, gross domestic product (GDP) has been the preferred measure to compare and evaluate national wealth. Yet right from the beginning, GDP has been acknowledged to be an inadequate measure of well-being (cf. Coyle 2014). Several commissions of experts have thus thought about alternatives to GDP. The most famous one is the Stiglitz-Sen-Fitoussi Commission (Stiglitz, Sen, and Fitoussi 2009), which comprised a group of renowned scientists, including five Nobel laureates in economics. The commission suggests a broadly conceived multidimensional statistical system. Measures “centred on people’s well-being” (Stiglitz et al. 2009, p.12) as well as measures reflecting environmental sustainability are to supplement GDP. Only a coherent and explicit multidimensional database, including both objective and subjective indicators of well-being, could allow policy-makers to identify, assess and improve what matters to their citizens.

The long history of proposals and their critics (e.g. Fleurbaey and Blanchet 2013) shows how difficult it is to agree upon a comprehensive set of conceptually distinct, i.e. non-overlapping indicators (Benjamin et al. 2017). Achieving this ultimate list solves only half of the problem. The set itself does not allow for any inference about the welfare of a single person or country unless we know how citizens trade the different indicators against each other (i.e. the weighting issue). In the final step, these preferences are to be aggregated in some normatively justified way to result in a national index of welfare (i.e. the aggregation issue).

Traditional alternatives to GDP such as the Human Development Index implicitly resolve both issues in a top-down approach as they predetermine a formula aggregating all of the considered indicators (e.g. UNDP 2015). In line with the suggestions of the Stiglitz-Sen-Fitoussi Commission, more recent approaches employ surveys in order to allow people to state the extent to which the different indicators affect their welfare (e.g. Benjamin et al. 2013, 2014). One proposal in this regard originates from the OECD’s Better Life Initiative (OECD 2011). The OECD considers 11 *dimensions* of life to be universal and thus essential and relevant to all societies, namely housing, income, jobs, community, education, environment, civic engagement, health, life satisfaction, safety and work-life balance (e.g. OECD 2015). The dimensions summarize a set of 24 more detailed and measurable indicators, both objective and subjective by nature. Although the OECD does not recommend a politically controversial transformation of the dimensions to one index, it offers a “way of responding to the demand for a concise picture of overall well-being across countries” (OECD 2011, p. 26), i.e. a way of resolving the weighting issue. This is the *Better Life Index*. The web-based survey allows people

all over the world to create individual indices of well-being by weighting the 11 dimensions. They rate each dimension on a scale from 0 (not important) to 5 (very important) and thus indirectly the underlying indicators, which are equally weighted within one dimension. A dimension's rating is divided by the sum of all of the dimension ratings to be translated into a relative weight. Finally resolving the aggregation issue, the OECD simply reports the averages of these relative ratings for a wide range of countries and demographic subgroups (e.g. OECD 2011, 2014, 2015).¹ Based on the relative weights and the objectively measured indicators, the OECD Better Life Index facilitates the computation of national indices of well-being.

The results of the survey as well as the tool itself have gained enormous media attention. They also constitute an influential contribution to the ongoing scientific debate on welfare measurement (e.g. Boarini and D'Ercole 2013, Delhey and Kroll 2013, Mizobuchi 2014, Durand 2015, Decancq 2016, Marković et al. 2016). Meanwhile, the Austrian government employed a domestic version of the Better Life Index to assess its citizens' policy priorities (WIFO 2012). This growing influence of the Better Life Index raises the fundamental question whether the tool facilitates a valid measuring of individual policy priorities. In this study, we examine whether the fact that the indicators are not rated directly, but as part of overarching dimensions, thwarts meaningful measurement.

Research in the assessment of environmental goods shows that individuals' valuation of one particular environmental improvement varies substantially depending on whether it is valued on its own or as part of a larger, more inclusive environmental program. Kahneman and Knetsch (1992a) label this phenomenon as the *embedding effect*. Likewise, the valuation of the Better Life Index' indicators may respond to the particular embedding in well-being dimensions. Such aggregated dimensions can only serve as summarizing labels if they do not affect the relative weights of the embedded indicators. Consider the example of four indicators that are weighted α each according to people's true preferences. If they are embedded in the same dimension, this dimension should receive a relative weight of 4α ; if only three of them belong to one dimension, the dimension weight should be 3α ; if two dimensions cover two indicators each, the two dimensions should be weighted 2α each. If, however, people associate only some particular aspect with the dimension label, they will rate depending on this individual notion of the label and thereby largely neglect the underlying indicators. In consequence, the two dimensions with two indicators from before might receive a weight greater than 4α in sum simply because they gain double attention and trigger two different associations.

¹ Sometimes, the OECD corrects "for biases in the age and gender composition" (OECD 2015, p. 37).

To investigate whether survey-based welfare indicators such as the Better Life Index are potentially prone to such embedding effects, we conducted an experiment, which assigned randomly selected participants different versions of a replication of the OECD's Better Life Index. In particular, we considered two distinct variations of the weighting tool. First, we look at whether allocating four indicators to either one or two dimensions has any impact on the weighting of the four underlying indicators in sum. We find that splitting the one dimension with four indicators into two dimensions with two indicators each leads to a much higher weight in sum. This suggests the presence of strong embedding effects. Second, we selectively withdraw particular indicators from one dimension and test whether the dimension's relative weight varies across the partially different sets of embedded indicators. It turns out that the relative weight of the same dimension is not affected at all by the composition of indicators embedded. Participants largely neglect the precise scope of the well-being dimension as described by the underlying indicators. We conclude that the Better Life Index is not suited to aggregate individual welfare over a range of well-being indicators.

We proceed as follows. Section 2 introduces the embedding phenomenon in more detail and describes its potential importance for the BLI. We describe the experiment, data and empirical methodology in Section 3. Section 4 presents the results and Section 5 concludes.

2. The embedding effect

Societies are confronted with complex assessments of environmental amenities. How much worth has to be assigned to a certain landmark? How much should we value an old-growth forest? The contingent valuation method is frequently employed for the purpose of environmental valuation. To determine the value of the intangible good, survey respondents state either their willingness-to-pay (WTP) or their willingness-to-accept (Hanemann 1994, Carson and Hanemann 2005, Kling, Phaneuf, and Zhao 2012). The validity of the method has been heavily disputed. According to Hausman (2012), people often do not respond out of stable well-defined preferences, but invent "their answers on the fly" (p.43). The answers thus do not allow for serious welfare statements since they do not provide credible information about individuals' preferences.

For our purpose of evaluating the Better Life Index (BLI), we expect one line of critique of the contingent valuation method (CVM) to be of particular importance, the embedding phenomenon. It implies that the valuation of an item varies substantially "... depending on whether the good is assessed on its own or embedded as part of a more inclusive package"

(Kahneman and Knetsch 1992a, p. 58). This effect can appear in a weak and a strong form, respectively called regular and perfect embedding (Loomis, Lockwood, and DeLacy 1993, Svedsäter 2000, McDaniels et al. 2003). *Regular embedding* occurs if the value assigned to a single good is higher when evaluated on its own than when it is embedded within a larger, more inclusive good (Kahneman and Knetsch 1992b, Svedsäter 2000). *Perfect embedding* occurs if the valuation of a single good is almost identical to the value assigned to the more inclusive good, containing the item in question. Evidently, the stronger notion of *perfect embedding* implies *regular embedding* (Kahneman and Knetsch 1992b).

Kahneman and Knetsch (1992a) detect the embedding phenomenon in a survey collecting subjects' willingness-to-pay for changes in the provision of environmental services in the Greater Vancouver region. A first group was asked about their WTPs for the overarching category 'improvement of environmental services', the subcategory 'improvement of disaster preparedness' and the concrete measure 'improving rescue equipment and personnel'. A second group indicated the WTPs only for the latter two items. A third group valued solely the last item. The average WTP of this most specific item varied substantially depending on whether it was evaluated on its own (\$122.64), jointly with preparedness for disasters (\$74.65) or with both preparedness for disasters and environmental services (\$14.12). As the level of embedding alters the values assigned to the same item, it becomes completely unclear which embedding structure reveals respondents' actual preferences Kahneman and Knetsch (1992a, 1992b). Diamond and Hausman (1994) therefore conclude that because of the existence of the embedding effect, contingent valuation should not be used for cost-benefit analysis.²

The apparent analogy between BLI and CVM motivates our analysis. Both BLI and CVM employ surveys to detect subjects' preferences by considering a trade-off. In the assessment of an environmental good, subjects have to indicate a particular monetary value given their limited budget. As societies need to allocate scarce resources to different policy goals, the values BLI users assign to each dimension translate into relative weights reflecting marginal rates of substitution. Assigning one dimension a higher weight reduces the weights of all other dimensions. Whether or not a policy that shifts resources from one dimension to another dimension enhances social welfare thus crucially hinges on the accurate and consistent measurement of the relative weights.

² Embedding also alters private decisions. Investors apply a "1/n strategy" allocating money evenly across funds offered by a savings plan. Embedding effects are present if the share of money invested in risky funds relative to safe assets depends on the number of risky funds offered. This is what Benartzi and Thaler (2001) find.

To test in how far the BLI also suffers from embedding effects we apply the adding-up test as described by Diamond and Hausman (1994). With respect to the BLI being a valid measure of the individuals preference ranking of the dimensions of life, the adding-up property requires that the joint weight given for two items X and Y should be the same as the sum of the single weights of X and Y . Furthermore, the adding-up property implies that the weight given to one dimension becomes lower in at least one case when one or more underlying indicators are removed from the same overarching dimension.³

3. Experiment and hypotheses

In order to investigate the possible existence of the embedding effect, we set up a web application very similar to the BLI web application. Using this web application, we conducted an experiment at universities throughout Germany.

3.1 Recruitment of participants

From the 18th of January to the 5th of February 2016, we distributed information flyers in introductory undergraduate lectures at the universities of Bochum, Dresden, Frankfurt, Göttingen, Magdeburg, Rostock, Wuppertal as well as at FU Berlin and TU Berlin.⁴ The flyers stated that, for research purposes, it is of interest to know ‘what makes life worth living’ (see Appendix 1). Each flyer shows a randomly created six-digit identification number (in the following: ‘ticket’) and a web address. The ticket ensured anonymous participation, prevented multiple use and randomly assigned participants to one of the treatment and control groups (see below). 2,730 tickets were distributed, 538 students participated in the experiment (response rate of 19.7%). The web address leads participants to the landing page of our replication of the BLI website, which we name ‘Replicated BLI (R-BLI)’ in the following. It was accessible from the 18th of January to the 12th of February 2016. The landing page introduced the BLI and its usage (see Appendix 2 to compare the OECD-BLI website and the R-BLI landing page).⁵ At first, participants entered their ticket number.⁶ Then, they provided information about their year of birth, gender, the size of their home town (less than 20k, 20k-100k, 100k-500k, 500k-1M, more than 1M) and their major subject of study (economics, business administration,

³ We conduct an adding-up test without income adjustments which can yield small deviations from exact adding-up considering income effects (Diamond and Hausman 1994, footnote 11).

⁴ Our sample therefore matches the age group that uses the OECD weighting tool the most as the vast majority of worldwide and also of German users are between 15 and 35 years old (OECD 2014).

⁵ We used a German version of the website only and are happy to provide screenshots on demand.

⁶ Consequently, the ticket number became invalid to avoid multiple use.

mathematics, engineering, natural sciences, law, medicine, languages, social sciences, applied arts, teaching, non-enrolled). If they finally agreed to participate, they were redirected to the R-BLI weighting tool.

3.2 Weighting of well-being dimensions

Our weighting tool invites participants to rate the dimensions of the BLI on a scale from 0 to 5. As with the original BLI the valuation bar is set to an equal valuation of all dimensions at 1 at the beginning, but can be moved over the whole scale. The ratings assigned to each dimension translate into relative weights, which, identical to the BLI weighting mechanism, add up to 100% (i.e. the relative weight is the rating of one dimension divided by the sum of the ratings assigned to all of the dimensions). To make participants aware of the trade-off resulting from the conversion into relative weights, a bar chart is included in the R-BLI weighting tool. Each bar reflects the relative weight assigned to the respective dimension and adjusts immediately in response to changes of the weights (see Appendix 2). The visualization of the trade-off resulting from the relative weighting is important in order to replicate the original BLI, which visualizes the trade-off through the dynamic adjustment of country-specific petals.

The information provided about the R-BLI is in principle the same as that provided on the BLI website. Most importantly, participants in either case can easily see which indicators are embedded in the respective dimension. Informative mouse-over pop-ups, summarizing the indicators of the dimensions to be rated, show up once participants move their cursor over the title of each dimension. In contrast to the BLI website, the pop-up also appears on the R-BLI page when moving the cursor over the weighting bars. This renders the underlying indicators even more visible using our weighting tool than when using the original tool. In addition, each dimension title on the BLI and the R-BLI website provides a link to documents with detailed information on the underlying indicators (see Supplementary Material). We will check later whether or not accessing this extra information affects the prevalence of embedding effects. When the rating is completed, respondents finally indicate whether they are familiar with the original BLI. We also measured the time subjects spent to complete the ratings. Respondents who took less than 45 seconds are excluded from the sample as it is rather unlikely that they seriously reflected upon the importance of all the dimensions. We obtain the same qualitative results without this restriction.

3.3 Main treatment

Participants are assigned five different versions of the R-BLI weighting tool, which only varies with respect to the *Jobs* dimension. Figure 1 illustrates the weighting tool for these versions. The original BLI embeds four indicators in the *Jobs* dimension: long-term unemployment rate, employment rate, personal earnings and job security. *Jobs* is thus the dimension with the highest number of indicators, providing us with many opportunities to redefine dimensions by regrouping indicators.

A first version of the R-BLI weighting tool replicates the original BLI weighting tool including eleven dimensions in total. *Jobs* covers the four indicators listed above. Note, however, that the OECD's mouse-over pop-up differs slightly from the underlying indicators. It contains the two indicators personal earnings and job security but subsumes the other two indicators under the heading 'unemployment'. People assigned to this version of our weighting tool constitute a first control group (C1). A first treatment group (T1) consists of participants who rate twelve dimensions as the *Jobs* dimension is split up into two distinct dimensions. Personal earnings and job security are individual job-specific indicators, reflecting an important aspect of the job quality employed people enjoy. We subsume these two indicators in the new dimension *Job Quality (JQ)*. The long-term unemployment rate and the employment rate, by contrast, refer to the overall labor market situation. They constitute the new dimension *Labor Market (LM)*. The information in the mouse-over pop-ups changes accordingly. The *JQ* dimension prompts the pop-up information 'personal earnings' and 'job security', whereas *Labor Market* prompts 'unemployment'.

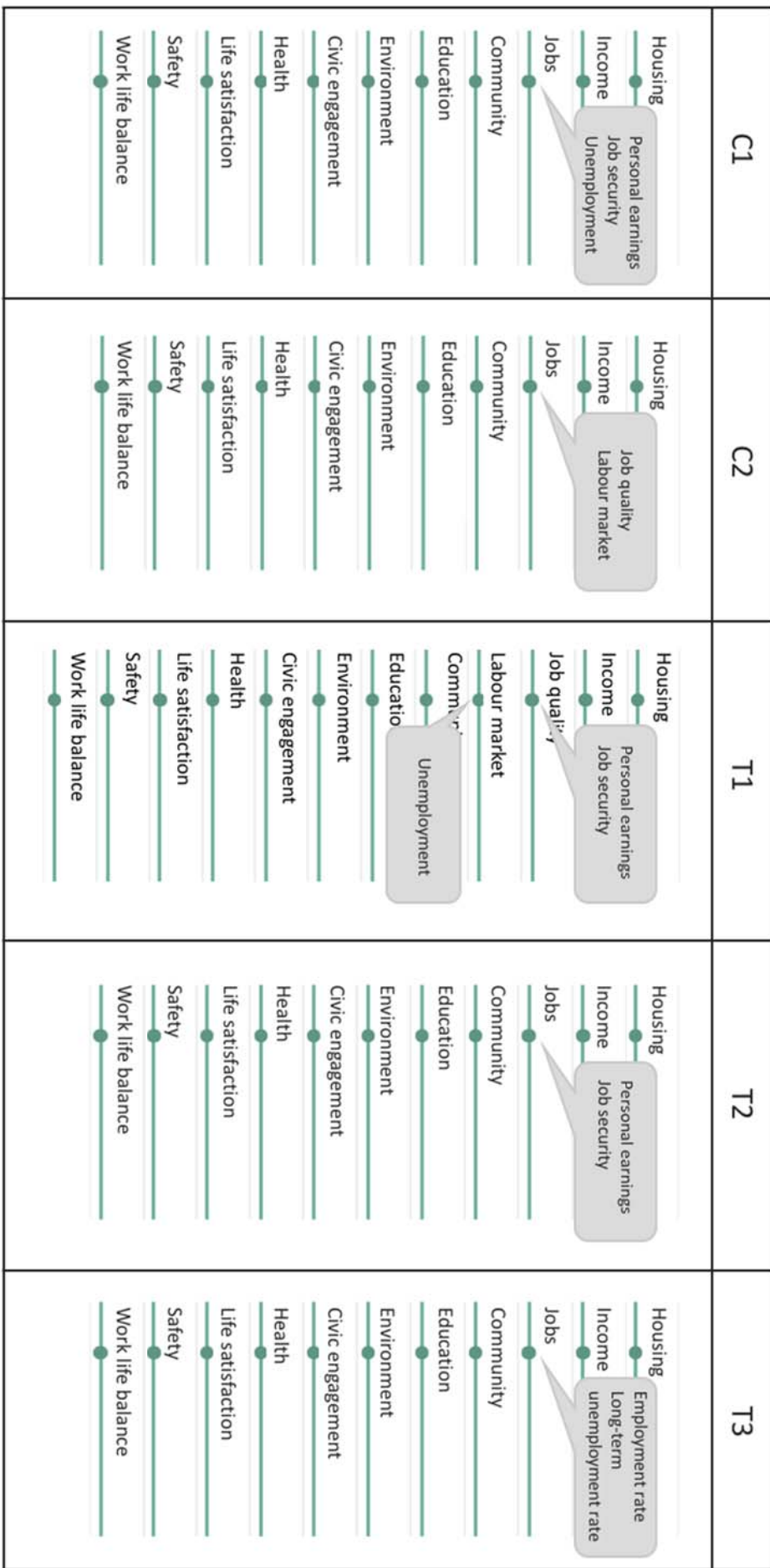
Ideally, people would be fully aware of the weights they implicitly attribute to the four embedded indicators. Adding up the relative weights (*RW*) the treated assign to the two new dimensions should therefore not differ from the controls' relative weighting of the whole *Jobs* dimension. In this case the adding-up property (cf. Diamond and Hausman 1994) would hold. In the reverse case, the sum of the relative weights in T1 should exceed the relative weighting of *Jobs* in C1. As we *expect such* an embedding effect, we postulate

$$\text{Hypothesis 1a: } RW_{C1}(Jobs) < RW_{T1}(JQ) + RW_{T1}(LM).$$

Hypothesis 1a assumes regular embedding. The stronger notion of perfect embedding would imply that the maximum of the relative weights T1 participants assign to *JQ* or to *LM* exactly equals the control groups' relative weight of *Jobs*. Accordingly, we formulate

$$\text{Hypothesis 2: } RW_{C1}(Jobs) = \max(RW_{T1}(JQ); RW_{T1}(LM)).$$

Figure 1: Treatment and control versions of the weighting tool



We could have split the dimension *Jobs* into four dimensions representing one indicator each. This, however, would not fit into the concept of dimensions as overarching categories (e.g. environment, education), thus leading to a discontinuity within the weighting tool. Participants' attention might be directed especially to these single-indicator dimensions, which we wanted to rule out as alternative explanation for embedding effects.

3.4 Embedding or framing?

The first treatment varies from the first control group with respect to the dimension titles describing working life. Hence, the treatment frames the same information differently. It is therefore in principle possible that a confirmation of the first hypothesis reflects a framing effect, as described by Tversky and Kahneman (1981). For instance, if the term 'job quality' were perceived more positively than 'job security' and 'personal earnings', the *Job Quality* dimension could receive more attention and would thus be given a higher weight.

To investigate framing effects, we introduced a second control group C2. Subjects in this group rated the same 11 dimensions as C1 (including *Jobs*), but the mouse-over pop-ups were framed differently. C2 participants saw the pop-ups 'job quality' and 'labor market', whereas C1 participants saw 'unemployment', 'job security' and 'personal earnings'. The underlying indicators of *Job Quality* were personal earnings and job security, while the underlying indicators of *Labor Market* were long-term unemployment rate and employment rate. If framing plays no role in our context, the relative weights assigned to *Jobs* should not differ between the two control groups, i.e.

$$\text{Hypothesis 3: } RW_{C1}(Jobs) = RW_{C2}(Jobs).$$

holds. Compared to C2, the first treatment T1 changes the embedding as *Job Quality* and *Labor Market* become dimensions, but it does not change the labelling of aspects of working life. Hence, we expect regular embedding also shows up when we compare C2 and T1:

$$\text{Hypothesis 1b: } RW_{C2}(Jobs) < RW_{T1}(JQ) + RW_{T1}(LM).$$

If our analysis supports either Hypothesis 1a or Hypothesis 1b but not both, it will indicate the existence of a framing effect.

3.5 The role of preconceived individual notions

People may have preconceived individual notions of a dimension title based on previous experience and pre-determined attitudes. This might influence the perception of the dimension and thus its rating. If these notions are very strong, the ratings will be immune to the information provided by the OECD about what is actually supposed to be rated, i.e. the underlying indicators. Any difference in the *Jobs* weight(s) from C1 and C2 and T1 will indirectly point in this direction. Once new dimension titles are used, the ratings change irrespectively of the fact that the underlying indicators remain the same. Individual notions could thus induce embedding effects. To shed more light on this conjecture, we test whether dimensions that are labelled the same way, and should thus recall the same individual notion, get the same relative weights although they carry distinct sets of indicators. We therefore introduced two further treatment groups, T2 and T3. As in the two control groups, both rated *Jobs* as one dimension of the original list of eleven dimensions, but the pop-ups showing up once people moved the mouse over *Jobs* as well as the underlying information were different. T2 only covered the job-specific aspects, namely personal earnings and job security, whereas T3 only got to see the labor market indicators long-term unemployment rate and employment rate. Subjects who take the information provided into account probably value the *Jobs* dimension differently, depending on the importance they assign to the embedded items. If people only translate their preconceived notions of the term ‘jobs’ into the relative weights and do not care about any information provided, the following hypothesis holds:

$$\text{Hypothesis 4: } RW_{T_2}(\text{Jobs}) = RW_{T_3}(\text{Jobs}) = RW_{C_1}(\text{Jobs}).$$

3.6 Summary of experimental groups and descriptive statistics

Table 1 sums up our modifications of the R-BLI weighting tool as well as information on the five experimental groups gathered by the landing page. The randomization process produced fairly similar groups given that the group sizes are not tremendous. The groups only differ on a statistically significant level in a very few characteristics. Comparing many characteristics across five groups inflates the likelihood of type I errors. It is very likely that at least one significant difference will be found even if by pure coincidence. We apply the Bonferroni correction to tackle this so-called multiple-comparison problem.⁷ The only significant

⁷ The basic idea is that the type I errors of the number of dependent tests (e.g. the tests per group comparison) add up at the maximum. Thus, the 10% level of significance must be divided by the number of tests to obtain the test specific equivalent level of 10% significance, i.e. in our case the p -value of each test must be lower than 0.1/15.

difference emerging is that ‘economics as major’ is more likely in group T3 than in group T2. To be on the safe side, we consider in our regression analyses the characteristics described in Table 1 as control variables.

Table 1: Overview of the experimental groups

	Control group C1	Control group C2	Treatment group T1	Treatment group T2	Treatment group T3
Number of dimensions	11	11	12	11	11
Label of <i>Jobs</i>	<i>Jobs</i>	<i>Jobs</i>	<i>Job Quality; Labor Market</i>	<i>Jobs</i>	<i>Jobs</i>
Mouse-over pop-up*	Personal earnings, Job security; Unemployment	Job quality; Labor market	Personal earnings; Job security; Unemployment	Personal earnings; Job security	Long-term unemployment rate, Employment rate
Number of observations	105	119	101	96	101
Female (share)	44%	39%	56%	45%	46%
Age (in years)	22.18 (4.64)	21.71 (2.85)	22.27 (5.51)	22.23 (3.50)	22.34 (3.02)
Knowledge of the BLI (share)	21%	16%	27%	25%	20%
Time spent weighting (in minutes, median)	1:42	1:39	1:46	1:43	1:44
Accessed information (share)	26%	26%	19%	27%	27%
<i>Size of home town (shares)</i>					
20,000 or less	30%	29%	29%	18%	25%
20,000 – 100,000	22%	17%	20%	23%	17%
100,000 – 500,000	14%	20%	19%	19%	23%
500,000 – 1,000,000	10%	9%	8%	9%	13%
1,000,000 or more	25%	24%	25%	31%	23%
<i>Major (shares)</i>					
Economics	17%	18%	24%	8%	23%
Business	33%	34%	28%	39%	33%
Administration					
Mathematics	13%	13%	17%	16%	15%
Languages	8%	12%	7%	6%	11%
Arts	10%	6%	7%	14%	5%
Other	18%	17%	18%	18%	14%

*Note: Standard deviation in parentheses. *Most mouse-over pop ups display indicators. Unemployment and labor market each represent the two indicators long-term unemployment rate and employment rate. Job Quality consists of the two indicators personal earnings and job security.*

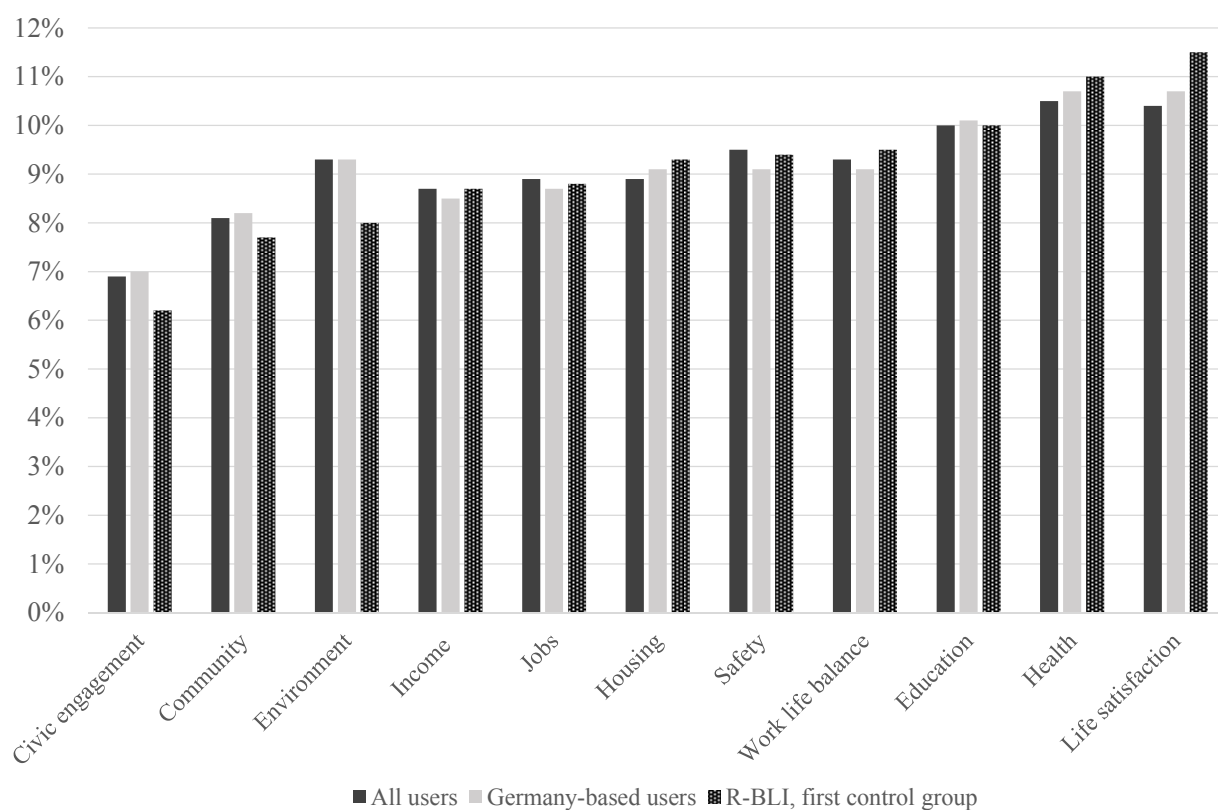
4. Experimental results

4.1 Mean analysis

We present the results of our replication of the original BLI in Figure 2, i.e. the ratings of C1. This allows us to compare the results obtained from our weighting tool to those published by the OECD. The relative weights of the 11 dimensions range from 6.2% (civic engagement) to

11.5% (life satisfaction). This matches the data from the worldwide ranking and the country specific ranking of Germany. There the lowest rankings are somewhat above 6% and the highest ranking somewhat above 10%. In addition, the order of these rankings is very similar to that of our first control group. In all three cases, civic engagement ranks last, whereas life satisfaction, health and education are at the top (OECD 2015). Thus, we find no evidence that participating in our experiment and using a replication of the BLI weighting tool affects the rating process in a way different from the original. We therefore consider the following results to generalize to the use of the original BLI.

Figure 2: Dimension ratings across different user groups



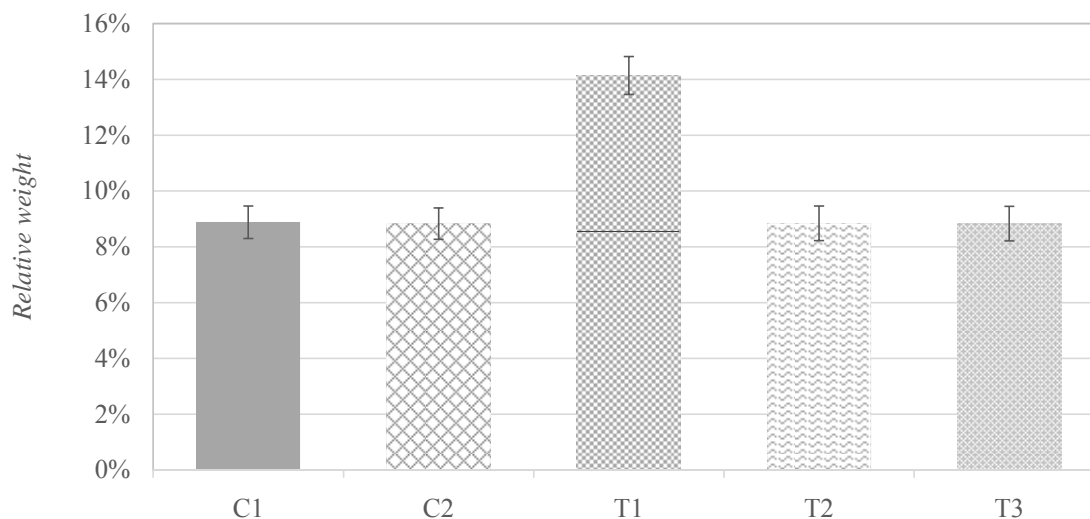
Note: The bars displays average relative weights of dimensions according to ratings of all users (left bars) and Germany-based users (central bars) of the original BLI (OECD 2015) as well as of control group 1 (right bars).

We start testing our hypotheses by comparing the mean relative weights assigned to the *Jobs* dimension and its treatment 1 replacements *Labor Market* and *Job Quality*. To test for statistical significance, we always use the parametric t-test and the non-parametric Mann-Whitney-U test. The t-test is based on the assumption of normality, which is confirmed by the Shapiro-Wilk test.⁸ We also assess the variance of the relative weights across all of the groups. Only when

⁸ Since each group sample is considered to be a large sample ($n > 30$), one can rest on the central limit theorem which predicts a normal sampling distribution.

comparing T1 and C1 variance homogeneity is not supported. In this case we employ the Welch adjustment for the t-test.

Figure 3: Dimension ratings across different user groups



Note: For C1 and C2 as well as T2 and T3 the bars display average relative weights of the Jobs dimension. The treatment group 1 bar illustrates the sum of average relative weights of the dimensions Job Quality (below the black line) and Labor Market (above the black line).

C1 as well as T1 rate the same embedded indicators, made transparent by mouse-over pop ups. As Figure 3 illustrates, the treated assigned a substantially and significantly ($p \approx 0.0000$) higher relative weight to *Labor Market* and *Job Quality* in sum (0.141) than the controls attributed to *Jobs* (0.088). This result strongly supports Hypothesis 1a, implying at least regular embedding. As group T1's relative weight of *Job Quality* (0.086) does not differ significantly from group C1's relative weight of *Jobs*, perfect embedding cannot be ruled out in line with Hypothesis 2.

Our results would reflect a framing effect rather than an embedding effect if treated subjects perceived the labels 'job quality' and 'labor market' differently from 'personal earnings', 'job security' or 'unemployment' and therefore assigned higher weights to the new dimensions than the controls C1 assigned to *Jobs*. However, C2 subjects who could see 'job quality' and 'labor market' as mouse-over pop-ups did not rate *Jobs* higher than the first control group based on the originally framed mouse-over pop ups. Hence, the framing Hypothesis 3 is rejected. In addition, all of the effects emerging from the comparison of groups T1 and C1 also appear when comparing T1 and C2, supporting Hypothesis 1b. The same new labels of the indicators receive a higher relative weight when rated as dimensions than when embedded in *Jobs*, suggesting the expected embedding effect.

Figure 3 also allows us to investigate the role of preconceived notions of dimension titles as an explanation for our results. Subjects may update their perception of a dimension either

only partly or even not at all by the information they receive from the BLI's website about underlying indicators. T3 and T4 varied the set of information provided by our R-BLI website. While C1 subjects received the full set of information of the BLI, T2 and T3 subjects had to rate distinct subsets, which add up to the full set (see Subsection 3.5). As Figure 3 illustrates, the relative weights of the *Jobs* dimension do not vary across groups C1, T2 and T3. Hypothesis 4 cannot be rejected. The information provided by the website is largely neglected and subsets of indicators receive the same value as the complete set. This result is striking as nevertheless a fraction of 26%-28% of people in groups C1, T2 and T3 have accessed the detailed information about the underlying indicators (see Table 1).

4.2 Regression analysis

To consider demographic and further information in the identification of embedding effects (see again Table 1), we conduct ordinary least square (OLS) regression analyses. If potential differences in the composition of groups influence our results, the regression analysis will yield more accurate average treatment effects compared to the mean comparisons from before. The relative weight a participant i assigns to the *Jobs* dimension $RW_i(\text{Jobs})$ is our dependent variable. For T1, this variable adds up the relative weights of *Labor Market* and *Job Quality*. Binary independent variables indicate the three treatment groups and the second control group (T₁, T₂, T₃ / C₂), while C1 serves as reference category. We include various individual characteristics: gender (FEM), age (AGE), the size of hometown (vector $SIZE$ of binary variables indicating different classes) and people's major (vector $MAJOR$ of binary variables indicating different fields of study). In addition, variables related to the weighting exercise are controlled for: previous knowledge of the BLI ($KNOW$), time spent weighting (binary variable $LONG = 1$ if above-median length of 1:42 minutes) and having accessed extra information while weighting ($INFO$). Altogether, this yields the following econometric model with α as the relative weight of *Jobs* of the reference group and ε as individual error term:

$$\begin{aligned}
 RW_i(\text{Jobs}) = & \alpha + \beta_1 T_{1,i} + \beta_2 T_{2,i} + \beta_3 T_{3,i} + \gamma C_1 \\
 & + \delta FEM_i + \phi AGE_i + SIZE_i' \lambda + MAJOR_i' \mu_i \\
 & + \theta KNOWS_i + \rho TIME_i + \sigma INFO_i + \tau LONG_i + \varepsilon_i
 \end{aligned}$$

Table 2 summarizes the results of the regression analysis. The explanatory variables are added stepwise (I: group variables, II: individual characteristics, III: weighting variables).

Table 2: Results of multiple regression analyses (OLS)

	I	II	III.1	III.2	III.3
<i>Experimental groups (ref. Control group 1)</i>					
Control group 2	0.000 (0.004)	0.000 (0.004)	0.000 (0.004)	0.001 (0.004)	-0.000 (0.004)
Treatment group 1	0.053*** (0.005)	0.053*** (0.005)	0.054*** (0.005)	-0.002 (0.004)	-0.032*** (0.004)
Treatment group 2	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.004)
Treatment group 3	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)
Female		-0.002 (0.003)	-0.001 (0.003)	-0.003 (0.003)	-0.001 (0.003)
Age (ref. 21 years)		0.000 (0.000)	0.000 (0.000)	0.001** (0.000)	-0.000 (0.000)
<i>Size of home town (ref. 100,000 - 500,000)</i>					
below 20,000		-0.003 (0.004)	-0.002 (0.004)	-0.001 (0.004)	-0.005 (0.004)
20,000-100,000		-0.001 (0.005)	-0.001 (0.005)	0.001 (0.004)	-0.003 (0.004)
500,000-1,000,000		0.003 (0.005)	0.004 (0.005)	0.004 (0.004)	0.001 (0.005)
1,000,000+		-0.005 (0.004)	-0.005 (0.004)	-0.004 (0.004)	-0.006 (0.004)
<i>Major (ref. Economics)</i>					
Business Administration		0.002 (0.005)	0.002 (0.005)	0.002 (0.004)	0.001 (0.004)
Mathematics		-0.003 (0.006)	-0.004 (0.006)	-0.002 (0.005)	-0.005 (0.005)
Languages		-0.007 (0.005)	-0.007 (0.006)	-0.007 (0.005)	-0.009 (0.005)
Arts		0.003 (0.006)	0.003 (0.006)	0.000 (0.005)	0.002 (0.005)
Other		-0.000 (0.005)	-0.000 (0.006)	-0.001 (0.005)	-0.001 (0.005)
Knows the OECD Better Life Index			-0.000 (0.004)	-0.000 (0.003)	-0.002 (0.003)
Above-median time spent weighting			-0.005 (0.003)	-0.004 (0.003)	-0.004 (0.003)
Accessed information			0.004 (0.004)	0.004 (0.004)	0.006 (0.003)
Constant	0.088*** (0.003)	0.081*** (0.010)	0.081*** (0.010)	0.071*** (0.012)	0.099*** (0.011)
Observations	522	522	522	522	522
R ²	0.313	0.325	0.329	0.037	0.161

Note: The table presents OLS estimates of the relative weight of the Jobs dimension (TI: the sum of the relative weights of the Labor Market dimension and the Job Quality dimension). Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The regression analysis closely resembles the results of the previous mean analysis in Section 4.1. Once *Jobs* is split into *Labor Market* and *Job Quality*, the sum of the relative weights of these two dimensions exceeds the relative weight of *Jobs* for the first control group. This implies a regular embedding effect (columns I–III.1). In addition, the *Jobs* weights of groups C2 and C1 do not vary significantly. At the same time, a post-estimation test reveals the *Jobs* weight of C2 to be significantly lower than the sum of the relative weights of *Labor Market* and *Job Quality* of T1. This again confirms that the particular framing of the terms ‘labor market’ and ‘job quality’ does not drive the treatment effect.

Subjects who took some time to weight the dimensions and/or accessed the extra information about the underlying indicators might have been less prone to the regular embedding effect as they set their wits to the indicators. As Table 3 reveals, however, the effect of T1 shows up robustly across these participants as well as across a variety of further subsamples. It seems to vary neither in size nor in statistical significance. Post-estimation tests confirm that the regular embedding effect also emerges across subsamples when comparing T1 to C2 where a potential framing effect is ruled out.

Our regression analyses presented in Table 2 also provide evidence for perfect embedding. For the estimation underlying column III.2, we have redefined the dependent variable for T1 as the relative weight assigned to *Job Quality* only. The coefficient does not differ significantly from zero and thus implies that *Job Quality* does not receive a weight by the treated that varies from the controls’ rating of *Jobs*. Again, a subcategory gets the same value as the whole dimension. We repeat this procedure for *Labor Market* (column III.3), which is assigned a significantly lower weight by group T1 compared to that the *Jobs* dimension receives from group C1. These results also hold across all of the subsamples.

At the close of this section, we return to the first three columns of Table 2 to shed light on the role of preconceived notions. The coefficients of the second and third treatment show that the two distinct subsets of the *Jobs* dimension receive the same weights as C1 attributes to the whole dimension. Again, we ascertain that people do not update their own perception of what the term *Jobs* means to them by the specific indicators to be weighted. At least in this case, one would expect lower weights for the subsets of indicators compared to the full set among subjects who accessed the extra information or took a relatively long time weighting. But Table 3 reveals a converse pattern, if any. Those who rated quickly / did not access extra information tended to assign slightly lower weights to the *Jobs* dimension if it embedded long-term unemployment rate and employment rate only (3rd treatment) than if it covered all of the indicators. The same applies to men and people coming from relatively large cities.

Table 3: Subgroup analyses, dependent on individual and weighting characteristics

	<i>initial sample</i>	<i>female</i>	<i>male</i>	<i>age below 21 years</i>	<i>age above 21 years</i>	<i>small town</i>	<i>large town</i>	<i>major econ./bus adm. or business</i>	<i>major not econ./bus. adm.</i>
<i>Experimental groups (ref. Control group 1)</i>									
Control group 2	0.000 (0.004)	0.003 (0.006)	-0.001 (0.006)	0.003 (0.005)	-0.003 (0.007)	-0.004 (0.007)	0.002 (0.005)	0.002 (0.006)	-0.004 (0.006)
Treatment group 1	0.054*** (0.005)	0.059*** (0.006)	0.047*** (0.007)	0.056*** (0.006)	0.049*** (0.007)	0.056*** (0.007)	0.051*** (0.006)	0.051*** (0.007)	0.054*** (0.007)
Treatment group 2	-0.005 (0.004)	-0.009 (0.006)	-0.001 (0.006)	-0.005 (0.006)	-0.007 (0.007)	0.002 (0.007)	-0.012** (0.006)	-0.005 (0.007)	-0.005 (0.006)
Treatment group 3	-0.004 (0.004)	0.007 (0.006)	-0.012* (0.006)	-0.004 (0.006)	-0.007 (0.007)	0.002 (0.007)	-0.010* (0.005)	-0.003 (0.006)	-0.009 (0.006)
Individual characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Weighting characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.081*** (0.010)	0.083*** (0.015)	0.080*** (0.013)	0.093*** (0.009)	0.094*** (0.008)	0.054*** (0.021)	0.089*** (0.010)	0.070*** (0.012)	0.097*** (0.012)
Observations	522	239	283	301	221	239	283	269	253
R ²	0.329	0.436	0.266	0.366	0.311	0.336	0.350	0.317	0.374

	<i>initial sample</i>	<i>knows BLI</i>	<i>does not know BLI</i>	<i>short time spent</i>	<i>long time spent</i>	<i>read extra info</i>	<i>did not read extra info</i>
<i>Experimental groups (ref. Control group 1)</i>							
Control group 2	0.000 (0.004)	0.009 (0.010)	-0.003 (0.005)	-0.006 (0.006)	0.003 (0.006)	0.013 (0.008)	-0.005 (0.005)
Treatment group 1	0.054*** (0.005)	0.064*** (0.010)	0.050*** (0.005)	0.050*** (0.007)	0.056*** (0.006)	0.049*** (0.011)	0.054*** (0.005)
Treatment group 2	-0.005 (0.004)	0.003 (0.008)	-0.009* (0.005)	-0.007 (0.006)	-0.006 (0.007)	-0.001 (0.010)	-0.006 (0.005)
Treatment group 3	-0.004 (0.004)	0.003 (0.011)	-0.006 (0.005)	-0.015** (0.007)	0.004 (0.006)	0.014 (0.010)	-0.011** (0.005)
Individual characteristics	yes	yes	yes	yes	yes	yes	yes
Weighting characteristics	yes	yes	yes	yes	yes	yes	yes
Constant	0.081*** (0.010)	0.096*** (0.016)	0.080*** (0.013)	0.082*** (0.012)	0.072*** (0.017)	0.062*** (0.023)	0.087*** (0.011)
Observations	522	112	410	257	265	130	392
R ²	0.329	0.505	0.299	0.385	0.331	0.246	0.382

Note: The table presents OLS estimates of the relative weight of the Jobs dimension (treatment 1: the sum of the relative weights of the Labor Market dimension and the Job Quality dimension) across various subgroups. Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5. Discussion

Our experimental results show strong embedding effects in the survey-based BLI and identify a likely reason for this. Individual preconceived notions of the dimensions seem to decisively affect subjects' ratings, whereas the underlying indicators are largely ignored. Further reasons might play a part in explaining our findings as well. The embedding effect could also originate from a tendency to assign average relative weights to all dimensions (in analogy to the '1/n strategy', see footnote 2), since people might not spend enough time and effort during the weighting process. This refers to two problems that generally serve as explanations for

embedding effects. People complete questionnaires too carelessly as they quickly invent answers ‘on the fly’ and they are not informed enough about the implications of their ratings. However, we consider these two reasons less important in our context. Subjects who spent a relatively long time weighting and those who accessed the additional information provided by the weighting tool show the same embedding effects as the whole sample. Furthermore, the first treatment group assigns the lowest weight to *Labor Market*, which is far from the average weight. Thus, at least a large group of subjects seems to assign average weights after taking some time to think their preferences through while weighting. Thus, we conjecture that preconceived notions that are recalled by the dimension titles are the main source of embedding effects in the BLI. Providing more information about the indicators will probably not solve the problem.

One might think of a simple way to avoid embedding effects in the BLI: people should directly weight the 24 indicators that currently underlie the dimensions. The embedding phenomenon, however, may persist. For instance, personal earnings as one of the indicators can include fixed wage rate plus performance bonus, end-of-the-year bonus, fringe benefits, and many more aspects – or not. Indicators, at least implicitly, always embed subcategories such that the weighting process remains prone to embedding effects. An extensive list of very detailed indicators might circumvent embedding effects, but would probably overwhelm participants. Even in this case, embedding effects cannot be ruled out unless all indicators that at least some participants consider as relevant are included.

Our results yield various implications. Most importantly, they cast serious doubts on the possibility of designing survey-based welfare indicators and, in particular, the BLI’s validity as a measure of citizens’ true preferences. As the *Jobs* dimension receives a much lower weight in the BLI than two similar dimension names containing the same information, we cannot infer the true relative weights of work-related aspects of life. This alone disproves the validity of the index since it needs to measure all of the relative weights in an exact manner. In addition, there is no reason why our results should only concern the *Jobs* dimension. In consequence, the Better Life Index does not provide us with credible individual well-being indices, which could be aggregated to a national index. Its results should hence not be interpreted in any respect. Any comparison of dimension weights, such as the statement ‘financial wealth matters more to people than a social support network’, cannot be inferred.

Often, measurement problems of surveys can be solved by only relying on variation over time. Though we cannot interpret the weight of *Jobs* or that of any other dimension, an increase of a certain weight from one year to the next might imply that this specific aspect got more


important to the participants. However, in the specific case of the BLI, the data provided are not informative unless we know what people actually associate with the specific dimension. Which policy will take account of a growing importance of the *Jobs* dimension? A policy that leads to higher wages at the cost of higher unemployment or just the opposite measure? A dimension title is only a label that respondents ultimately interpret for themselves. Thus, statisticians and policy makers are left with no better understanding of what exactly people consider worthwhile or which policy implications such preferences could have. The sad lesson from our experiment is therefore that the Better Life Index does not contribute to the ongoing search for a valid measure of social welfare. Beyond that, our study poses a crucial challenge to rule out embedding effects for all survey-based approaches.

References

- Benartzi, Shlomo and Richard H. Thaler (2001): “Naive Diversification Strategies in Defined Contribution Saving Plans”, *The American Economic Review* 91(1), pp. 79–98.
- Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Nichole Szembrot (2013): “Aggregating local preferences to guide marginal policy adjustments”, *American Economic Review* 103(3), pp. 605–610.
- Benjamin, Daniel J, Kristen Cooper, Ori Heffetz, and Miles S Kimball (2017): *Challenges in Constructing a Survey-Based Well-Being Index*, NBER Working Paper Series No. 23111.
- Benjamin, Daniel J, Ori Heffetz, Miles Kimball, and Nichole Szembrot (2014): “Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference”, *American Economic Review* 104(9), pp. 2698–2735.
- Boarini, Romina and Marco Mira D’Ercole (2013): “Going beyond GDP: An OECD Perspective”, *Fiscal Studies* 34(3), pp. 289–314.
- Carson, Richard T. and W. Michael Hanemann (2005): “Chapter 17 Contingent Valuation”, *Handbook of Environmental Economics* 2(5), pp. 821–936.
- Coyle, Diane (2014): *GDP. A Brief but Affectionate History*, Princeton University Press: Princeton and Oxford.
- Decancq, Koen (2016): “Measuring Multidimensional Inequality in the OECD Member Countries with a Distribution-Sensitive Better”, *Social Indicators Research*, online first.
- Delhey, Jan and Christian Kroll (2013): “A ‘Happiness Test’ for the New Measures of National Well-Being: How Much Better than GDP are They?”, in: Hilke Brockmann and Jan Delhey (eds.), *Human Happiness and the Pursuit of Maximization. Is More Always Better?*, Springer Netherlands: Dordrecht, pp. 191–210.
- Diamond, Peter A. and Jerry A. Hausman (1994): “Contingent Valuation: Is Some Number Better than No Number?”, *Journal of Economic Perspectives* 8(4), pp. 45–64.
- Durand, Martine (2015): “The OECD better life initiative: How’s life? And the measurement of well-being”, *Review of Income and Wealth* 61(1), pp. 4–17.
- Fleurbaey, Marc and Didier Blanchet (2013): *Beyond GDP: Measuring welfare and assessing sustainability*, Oxford University Press: New York.
- Hanemann, W. Michael (1994): “Valuing the Environment Through Contingent Valuation”, *Journal of Economic Perspectives* 8(4), pp. 19–43.
- Hausman, Jerry (2012): “Contingent Valuation: From Dubious to Hopeless”, *Journal of Economic Perspectives* 26(4), pp. 43–56.
- Kahneman, Daniel and Jack L. Knetsch (1992a): “Valuing Public-Goods - the Purchase of Moral Satisfaction”, *Journal of Environmental Economics and Management* 22(1), pp. 57–70.
- Kahneman, Daniel and Jack L. Knetsch (1992b): “Contingent Valuation and the Value of Public Goods: Reply”, *Journal of Environmental Economics and Management* 22(1), pp. 90–94.
- Kling, Catherine L., Daniel J. Phaneuf, and Jinhua Zhao (2012): “From Exxon to BP: Has Some Number Become Better than No Number?”, *Journal of Economic Perspectives* 26(4), pp. 3–26.

-
- Loomis, John, Michael Lockwood, and Terry DeLacy (1993): “Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection”, *Journal of Environmental Economics and Management* 24(1), pp. 45–55.
- Marković, M., S. Zdravković, M. Mitrović, and A. Radojičić (2016): “An Iterative Multivariate Post Hoc I-Distance Approach in Evaluating OECD Better Life Index”, *Social Indicators Research* 126(1), pp. 1–19.
- McDaniels, Timothy L., Robin Gregory, Joseph Arvai, and Ratana Chuenpagdee (2003): “Decision structuring to alleviate embedding in environmental valuation”, *Ecological Economics* 46(1), pp. 33–46.
- Mizobuchi, Hideyuki (2014): “Measuring World Better Life Frontier: A Composite Indicator for OECD Better Life Index”, *Social Indicators Research* 118(3), pp. 987–1007.
- OECD (2011): *How’s Life? Measuring Well-Being*, OECD Publishing: Paris.
- OECD (2014): *OECD Better Life Index Country Reports*, OECD Publishing: Paris.
- OECD (2015): *How’s Life? 2015: Measuring Well-Being*, OECD Publishing: Paris.
- Stiglitz, Joseph E., Amartya K. Sen, and Jean-Paul Fitoussi (2009): *Report by the Commission on the Measurement of Economic Performance and Social Progress*, .
- Svedsäter, Henrik (2000): “Contingent valuation of global environmental resources: Test of perfect and regular embedding”, *Journal of Economic Psychology* 21(6), pp. 605–623.
- Tversky, Amos and Daniel Kahneman (1981): “The framing of decisions and the psychology of choice”, *Science* 211(4481), pp. 453–458.
- UNDP, United Nations Development Programme (2015): *Human Development Report 2015*, United Nations Development Programme: New York.
- WIFO (2012): *Mehr als Wachstum Messung von Wohlstand und Lebensqualität in ausgewählten OECD-Ländern Ausgangslage und Aufgabenstellung*, mimeo.

Appendix 1 Recruitment flyer


 Freie Universität Berlin

Was macht das Leben lebenswert?

Geld? Familie? Karriere? Umwelt?
 Im Rahmen einer sehr (!) kurzen Umfrage kannst Du festlegen, was für Dich im Leben am wichtigsten ist und gleichzeitig einen Beitrag zur Beantwortung aktueller Forschungsfragen leisten.

Besuche dazu die Internetseite:


<http://goo.gl/hAR8Nv>

und gib diese Ticket-Nummer an:


b6c9e2

Dieses Ticket ist nur einmal gültig. Deine Angaben sind völlig anonym. Die Umfrage ist Teil eines Forschungsprojekts des Lehrstuhl Schöb.

Vielen Dank für Deine Unterstützung!



Kontakt:
 Louisa.Reumont@fu-berlin.de
 Ls-Schoeb@wiwiss.fu-berlin.de


 Freie Universität Berlin

What makes life worth living?

Is it Money? Is it Family? A Career? The environment?
 This very short (!) survey allows you to determine what is most important for you in your life.

Please visit the website


<https://goo.gl/hAR8Nv>

and use the following number as identification:

"Ticket Number"

This ticket is valid only once. Your answers will remain completely anonymous. This survey is part of a research project of the Chair of International Public Economics.

Thank you very much for your support!



Contact:
 Louisa.Reumont@fu-berlin.de
 Chair of International Public Economics: LS-Schoeb@fu-berlin.de

Appendix 2: Webpages of OECD BLI and R-BLI

Landing Page of the OECD BLI

OECD Better Life Index

English Español Français Pycckий Deutsch Português Italiano Contact Us [en/eng](#)

Index Responses Countries Topics FAQ

Create Your Better Life Index

What is your recipe for a better life – a good education, clean air, nice home, money?
See how your country measures up on the topics important to you. [Help](#)

Start with all topics rated equally or set your own preferences [here](#).

How's life?

There is more to life than the cold numbers of GDP and economic statistics – This Index allows you to compare well-being across countries, based on 11 topics the OECD has identified as essential, in the areas of material living conditions and quality of life.

[Download executive summary](#)
[Download the index data](#)
[Learn more about the Better Life Initiative](#)

Mapping well-being

What matters most to people around the world?

More than 100,000 users of the Better Life Index around the world have shared their views on what makes for a better life. [Explore the map.](#)

Find out which topics are most important to people around the world.

Create Your Better Life Index

Rate the topics according to their importance to you:

- Housing
- Income
- Jobs
- Community
- Education
- Environment
- Civic Engagement
- Health
- Life Satisfaction
- Safety
- Work-Life Balance

Gender differences

Compare with others

Share your index

Source: www.oecdbetterlifindex.org

Landing Page of the R-BLI

Better Life Index

There is more to life than the cold numbers of GDP and economic statistics. This Index allows you to compare societal well-being based on various topics identified as essential.

Participate

Ticket N°

Year of birth

Please choose

Gender

Male

Female

Size of hometown

Please choose

Major of study

Please choose

R-BLI experimental page (default setting)

Create your Better Life Index!

Rate the topics according to their importance to you.

[Click here to obtain further information about the specific topics](#)

I am familiar with the OECD Better Life Index

Please choose

R-BLI experimental page with exemplified weighting

Create your Better Life Index!

Rate the topics according to their importance to you.

[Click here to obtain further information about the specific topics](#)

I am familiar with the OECD Better Life Index

Please choose

Supplementary Material

Information provided on the R-BLI web-application

The information provided below was extracted from the OECD BLI web-application and used in the experimental set-up for further informational purposes. Dimensions are in bold letters, underlying indicators are in italics.

1 Information on the dimension jobs, C1

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

2 Information on the dimension jobs, C2

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

'Job quality'

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing

earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

'Labour market'

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

3 Information on the dimensions *job quality* and *labour market*, T1

Job quality

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

Labour market

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

4 Information on the dimension jobs, T2

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

5 Information on the dimension jobs, T3

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

A3.6 Information on the other dimensions for all experimental groups

Housing

Living in satisfactory housing conditions is one of the most important aspects of people's lives. Housing is essential to meet basic needs, such as shelter, but it is not just a question of four walls and a roof. Housing should offer a place to sleep and rest where people feel safe and have privacy and personal space; somewhere they can raise a family. All of these elements help make a house a home. And of course there is the question whether people can afford adequate housing.

Housing expenditure

Housing costs take up a large share of the household budget and represent the largest single expenditure for many individuals and families, by the time you add up elements such as rent,

gas, electricity, water, furniture and repairs. In the OECD, households on average spend around 18% of their gross adjusted disposable income on keeping a roof over their heads.

Rooms per person and dwellings with basic facilities

When looking at housing, it is important to examine living conditions, such as the average number of rooms shared per person and whether dwellings have access to basic facilities.

The number of rooms in a dwelling, divided by the number of persons living there, indicates whether residents are living in crowded conditions. Overcrowded housing may have a negative impact on physical and mental health, relations with others, and children's development.

Income

While money may not buy happiness, it is an important means to achieving higher living standards and thus greater well-being. Higher economic wealth may also improve access to quality education, health care and housing.

Household net-adjusted disposable income

Household net-adjusted disposable income is the amount of money that a household earns, or gains, each year after taxes and transfers. It represents the money available to a household for spending on goods or services.

Household financial wealth

Household financial wealth is the total value of a household's financial worth, or the sum of their overall financial assets minus liabilities. Financial wealth takes into account: savings, monetary gold, currency and deposits, stocks, securities and loans.

Community

Humans are social creatures. The frequency of our contact with others and the quality of our personal relationships are thus crucial determinants of our well-being. Studies show that time spent with friends is associated with a higher average level of positive feelings and a lower average level of negative feelings than time spent in other ways.

Social support network

A strong social network, or community, can provide emotional support during both good and bad times as well as access to jobs, services and other material opportunities. Across the OECD, 88% of people believe that they know someone they could rely on in time of need.

Education

Education plays a key role in providing individuals with the knowledge, skills and competences needed to participate effectively in society and in the economy. In addition, education may improve people's lives in such areas as health, civic participation, political interest and happiness. Studies show that educated individuals live longer, participate more actively in politics and in the community where they live, commit fewer crimes and rely less on social assistance.

Years in education

In a fast-changing knowledge economy, education is about learning skills for life. But how many years of school, college, or training will future generations expect to have?

Educational attainment

Having a good education greatly improves the likelihood of finding a job and earning enough money. Highly-educated individuals are less affected by unemployment trends, typically because educational attainment makes an individual more attractive in the workforce. Lifetime earnings also increase with each level of education attained.

Environment

The quality of our local living environment has a direct impact on our health and well-being. An unspoiled environment is a source of satisfaction, improves mental well-being, allows people to recover from the stress of everyday life and to perform physical activity. Having access to green spaces for example, is an essential part of quality of life. Also, our economies rely not only on healthy and productive workers but also on natural resources such as water, timber, fisheries, plants and crops. Protecting our environment and natural resources therefore remains a long-term priority for both our generation and those to come.

Air pollution

Outdoor air pollution is one important environmental issue that directly affects the quality of people's lives. Despite national and international interventions and decreases in major pollutant emissions, globally the health impacts of urban air pollution continue to worsen, with air pollution set to become the top environmental cause of premature mortality by 2050.

Water quality

Access to clean water is fundamental to human well-being. Managing water to meet that need is a major – and growing – challenge in many parts of the world. Many people are suffering from inadequate quantity and quality of water. Despite significant progress in OECD countries in reducing water pollution, from fixed sources such as industrial and municipal wastewater treatment plants, diffuse pollution from agriculture and urban run-offs remains a challenge, and improvements in freshwater quality are not always easy to discern.

Civic engagement

Trust in government is essential for social cohesion and well-being. Today, more than ever, citizens demand greater transparency from their governments. Information on the who, why and how of decision making is essential to hold government to account, maintain confidence in public institutions and support a level playing field for business. Greater transparency is not only key to upholding integrity in the public sector; it also contributes to better governance. Indeed, openness and transparency can ultimately improve public services by minimising the risk of fraud, corruption and mismanagement of public funds.

Voter turnout

High voter turnout is a measure of citizens' participation in the political process. Voter turnout is defined as the percentage of the registered population that voted during an election. High voter turnout is desirable in a democracy because it increases the chance that the political system reflects the will of a large number of individuals, and that the government enjoys a high degree of legitimacy.

Consultation on rule-making

Transparency in communication and open access to regulations promotes government accountability, a business-friendly environment and public trust in government institutions.

Indeed, a country's regulations contain much information about how a society is organised, the rules of the game and the political decisions taken. If citizens can readily access and understand regulations, it is more likely that they will participate in the legislative process and comply with the rules.

Health

Good health is one of the most important things to people and also brings many other benefits, including enhanced access to education and the job market, an increase in productivity and wealth, reduced health care costs, good social relations, and of course, a longer life.

Life expectancy

Life expectancy is the most widely used measure of health, although it only takes into account the length of people's life and not their quality of life. There have been remarkable gains in life expectancy over the past 50 years in OECD countries.

Self-reported health

Most OECD countries conduct regular health surveys which allow respondents to report on different aspects of their health. The commonly-asked question, "How is your health?" is one way of collecting data on self-perceived health status. Despite the subjective nature of this question, the answers received have been found to be a good predictor of people's future health care use.

Life satisfaction

Measuring feelings can be very subjective, but is nonetheless a useful complement to more objective data when comparing quality of life across countries. Subjective data can provide a personal evaluation of an individual's health, education, income, personal fulfilment and social conditions. Surveys, in particular, are used to measure life satisfaction and happiness.

Life satisfaction

Life satisfaction measures how people evaluate their life as a whole rather than their current feelings.

Safety

Personal security is a core element for the well-being of individuals, and includes the risks of people being physically assaulted or falling victim to other types of crime. Crime may lead to loss of life and property, as well as physical pain, post-traumatic stress and anxiety. One of the biggest impacts of crime on people's well-being appears to be through the feeling of vulnerability that it causes.

Assault rate

According to recent data, 3.9% of people in OECD countries say they have been assaulted or mugged over the past 12 months.

Homicide rate

Homicide rates (the number of murders per 100,000 inhabitants) only represent the most extreme form of contact crime and thus do not provide information about more typical safety conditions. They are however a more reliable measure of a country's safety level because, unlike other crimes, murders are usually always reported to the police.

Work-life Balance

Finding a suitable balance between work and daily living is a challenge that all workers face. Families are particularly affected. The ability to successfully combine work, family commitments and personal life is important for the well-being of all members in a household. Governments can help to address the issue by encouraging supportive and flexible working practices, making it easier for parents to strike a better balance between work and home life.

Employees working long hours

An important aspect of work-life balance is the amount of time a person spends at work. Evidence suggests that long work hours may impair personal health, jeopardise safety and increase stress.

Time devoted to leisure and personal care

Furthermore, the more people work, the less time they have to spend on other activities, such as personal care or leisure. The amount and quality of leisure time is important for people's overall well-being, and can bring additional physical and mental health benefits.

Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin
Discussion Paper - School of Business and Economics - Freie Universität Berlin

2017 erschienen:

- 2017/1 ARONSSON, Thomas und Ronnie SCHÖB
Habit Formation and the Pareto-Efficient Provision of Public Goods
Economics
- 2017/2 VOGT, Charlotte; Martin GERSCH und Cordelia GERTZ
Governance in integrierten, IT-unterstützten Versorgungskonzepten im
Gesundheitswesen : eine Analyse aktueller sowie zukünftig möglicher
Governancestrukturen und -mechanismen
Wirtschaftsinformatik
- 2017/3 VOGT, Charlotte; Martin GERSCH und Hanni KOCH
Geschäftsmodelle und Wertschöpfungsarchitekturen intersektoraler,
IT-unterstützter Versorgungskonzepte im Gesundheitswesen
Wirtschaftsinformatik
- 2017/4 DOMBI, Akos und Theocharis GRIGORIADIS
Ancestry, Diversity & Finance : Evidence from Transition Economies
Economics
- 2017/5 SCHREIBER, Sven
Weather Adjustment of Economic Output
Economics
- 2017/6 NACHTIGALL, Daniel
Prices versus Quantities: The Impact of Fracking on the Choice of Climate
Policy Instruments in the Presence of OPEC
Economics
- 2017/7 STOCKHAUSEN, Maximilian
The Distribution of Economic Resources to Children in Germany
Economics