GLOBAL SELF-ESTEEM ACROSS THE LIFE SPAN:
A CROSS-SECTIONAL COMPARISON BETWEEN
REPRESENTATIVE AND SELF-SELECTED
INTERNET SAMPLES

Helle Pullmann
Jüri Allik
Anu Realo

Department of Psychology, University of Tartu, Tartu, Estonia; and The
Estonian Centre of Behavioural and Health Sciences, Tartu, Estonia

The cross-sectional trajectory of global self-esteem across the life span was examined administering the Rosenberg and Single-Item Self-Esteem scales to Estonians (N = 29,463) who were either randomly selected from the National Census to represent the population or self-recruited through the Internet. The results (a) challenge the recent conclusion of a universal age trajectory of self-esteem, (b) demonstrate that self-recruited Internet data collection method is biased compared to random sampling, and (c) present that different self-esteem items have dissimilar trajectories. A variance component analysis confirmed that age differences in self-esteem are relatively small compared to interindividual differences and measurement error.

Based on a huge data set collected from 326,641 individuals over the Internet, Robins and his colleagues proposed that there was a
universal trajectory of self-esteem across the life span (Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002; Robins & Trzesniewski, 2005). On average, self-esteem estimated by a single-item scale (“I see myself as someone who has high self-esteem”) is relatively high in childhood, drops during adolescence (particularly for girls), rises gradually throughout adulthood, and declines sharply in old age. This trajectory presumably holds across gender, socioeconomic status, ethnicity, and nationality (Robins et al., 2002). However, all these normative changes in the mean level of self-esteem are relatively small and do not exceed one third of standard deviation.

These findings, along with some earlier studies (e.g., Demo, 1992; McCarthy & Hoge, 1982; Robins, Norem, & Cheek, 1999), challenge the standard view according to which there are no systematic age differences in global self-esteem (Wylie, 1979), the latter construct defined as the individual’s generalized positive or negative attitude toward the self as a human being (Rosenberg, 1965). The considerable overlap with the Big Five personality dimensions may explain the normative trajectory of self-esteem. Indeed, it is well documented that measures of self-esteem correlate strongly with two Big Five personality dimensions, Neuroticism and Extraversion (Costa, McCrae, & Dye, 1991; Judge, Erez, Bono, & Thoresen, 2002; Kwan, Bond, & Singelis, 1997; Pullmann & Allik, 2000; Robins, Tracy, Trzesniewski, Gosling, & Potter, 2001) and that these relationships remain invariant across at least 53 cultures (Schmitt & Allik, 2005). Based on similar observations, Judge and colleagues (2002) proposed that self-esteem represents a general Neuroticism factor and contains only a small amount of unique information that goes beyond it (Judge, et al., 2002). This may indicate that self-esteem originates from personality dispositions, which, according to the Five-Factor Theory of personality (McCrae & Costa, 1999; Allik & McCrae, 2002), are deeply rooted in biology and remain relatively stable throughout the whole life course. Indeed, self-esteem appears to be a heritable trait (about 30% to 50%), although being also subject to considerable environmental influences that in most cases are probably not shared by members of a twin pair (Kendler, Gardner, & Prescott, 1998; Roy, Neale, & Kendler, 1995; for a review see Neiss, Sedikides, & Stevenson, 2002). Based on cross-sectional and longitudinal studies, Costa and McCrae (2002) identified a specific pattern of changes for each main personality trait; the changes are quite small compared with the general stability of personality traits. In most cultures, neuroticism, extraversion, and openness appear to decrease with age after late adolescence, whereas agreeableness and conscientiousness increase (Costa & McCrae, 2002; McCrae et al., 2000, 2004; Roberts,
Walton, & Viechtbauer, 2006). To the extent the inverse value of self-esteem is an indicator of neuroticism, the reported gradual rise of self-worth through adulthood is compatible with the normative trajectory of the decrease in neuroticism.

Albeit impressive, single-item self-esteem data collected through the Internet (Robins et al., 2002) probably do not give an adequate picture about the normative trajectory of self-esteem across the life span. Firstly, Internet samples are likely to be selective, which places limits on external validity. For instance, not all people have access to the Internet (Walsh, Kiesler, Sproull, & Hesse, 1992) and there is still a large number of people who have only a limited access to a computer despite exponential growth of the Internet (Barbeite & Weiss, 2004; Birnbaum, 2004b; Crawford, Couper, & Lamias, 2001). Further, those who have better access may not be representative of the population in general (Smith, 2001; Birnbaum, 2004b). Moreover, participants in the Internet-based research may have a different motivation for taking a test (Buchanan & Smith, 1999) and as was argued by Reips (2000), the typical Internet test-taker is more motivated than a typical laboratory participant. Until the problem of self-selection has been resolved, the results collected through the Internet cannot be generalized to the general population (Birnbaum, 2004a). Although Robins and his colleagues (2002) provided several strong arguments about the validity of the Internet method (e.g., similarity of the results across gender and nationality), the data collected through the Internet, as Robins, Tracy, et al. (2001) admit, retain in the status of preliminary findings until they are replicated on randomly selected, representative samples of individuals. Therefore, future research should examine age differences in self-esteem using representative samples of individuals, which would greatly strengthen the generalizability of these conclusions and alleviate concerns about the Internet sample (Robins et al., 2002, p. 432).

Secondly, the conclusions from studies based on single-item measures might not replicate because there may be idiosyncrasies about the single item, or interactions of item with cohort or sample. These are real dangers because the large sample size makes trivial effects significant, and artifacts may be replicable across subsamples. For example, a national probability sample of 16,000 respondents in Canada, who answered an abridged six-item Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), showed no gradual increase of self-esteem throughout adulthood. On the contrary, self-esteem gradually declined with increase of age for both sexes and all income groups (McMullin & Cairney, 2004). This discrepancy may be due to sampling bias of the Internet sample, but it may be also that the
sum-score of six items behaves differently from the single-item measure of self-esteem. There is no evidence that the normative trajectories of single items of the same scale are identical. Although correlation between the RSES and single-item measure of self-esteem is relatively high in all age groups (Robins, Hendin, & Trzesniewski, 2001), this does not preclude a possibility that respondents of different age groups interpret the single item slightly differently, which could lead to selective elevation or drop of the mean score of this item in different age groups.

The main aim of the present study was to examine age differences in self-esteem across the entire life span, including representative samples. Except for a few studies (e.g., McMullin & Cainer, 2004), most of the literature describing self-esteem across the life span relies on nonrepresentative samples of particular age cohorts. The second goal of the study was to examine the validity of the Internet data collection method and of the single-item measure of self-esteem. More specifically, we examined normative age trajectory of global self-esteem among three nationally representative Estonian samples and compared these results with the data of self-recruited cross-sectional sample collected via the Internet. We focus on normative changes in global self-esteem (henceforth a cross-sectional trajectory across the life span), which is assessed by mean-level differences in self-regard across age groups over time. Mean-level change refers to changes in the average self-esteem level of a population and indicates whether the sample as a whole is increasing or decreasing on this trait. Although cross-sectional studies are unable to isolate age effects from period and cohort effects, they can estimate the existence of the mean level changes in self-esteem across age.

METHODS

Measuring Instruments

Rosenberg Self-Esteem Scale (RSES)
The Estonian version of the Rosenberg Self-Esteem Scale (Pullmann & Allik, 2000) was administered to the participants to measure their global self-esteem. Rosenberg (1965) defined self-esteem as a feeling of self-acceptance, self-respect, and positive self-evaluation that is conceptualized as a relatively enduring characteristic rather than something that shifts abruptly from one situation to another. Items of the RSES (see Table 2) were answered on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). All results are reported in terms of the mean value across all 10 items.
An Estonian version of the Single-Item Self-Esteem Scale (“Mul on kõrge enesehinnang”) was used in this study as a literal translation of the original SISE item “I have high self-esteem” (Robins, Hendin, et al., 2001). Items of the SISE were answered on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). For mnemonic purposes, samples in which the SISE was used instead of the 10-item RSES scale are marked with an asterisk (*).

**Samples**

Five different samples were used in this study. The first three samples were nationally representative samples of adult Estonians randomly selected from the National Census (Rahvastikuregister), the fourth was a geographically representative sample of adolescents, and the last sample self-recruited through the Internet. According to the Statistical Office of Estonia (2000), there were about 690,000 (53% females and 47% males) native Estonian-speaking individuals in age from 15 to 74 years living in Estonia in 2000. From those, about 42% had completed a basic education, 46% had finished a secondary or vocational school, and 13% had a university degree. In this study, representative samples were stratified to match National Census data on the variables of age, gender, educational attainment, and the approximate proportion between urban-rural residencies.

**Sample R1***

The first sample came from a wider study of social capital, cultural value dimensions, and identity in Estonia in 2002. The project principal investigators were the second and third authors of the article, and Aune Valk from the University of Tartu. The questionnaire consisted of several parts from which only the self-esteem measure was relevant for this study. Global self-esteem was measured by the SISE among 1395 native Estonians (56% females and 44% males) ranging from 15 to 74 years old, with the mean age of 43.5 years ($SD = 17.6$). The sample was randomly selected from the National Census and was representative of the Estonian-speaking population in Estonia regarding residence geography, age, gender, and educational level. About 28% of the respondents had completed a basic education, 55% secondary school level, 14% had a university degree, and about 3% declined to declare their educational level. About 8% were the residents of the capital city, 30% were living in different towns, and 62% in rural areas. The survey was carried out by the TNS Emor, the major service marketing research and consulting company in Estonia.
Sample R2*
Data from the second sample was derived from a larger study that aimed to study changing values, attitudes, attributes, and behaviour patterns within European polities. The European Social Survey (ESS) is a biennial multicountry survey covering over 20 nations. The project is directed by a Central Coordinating Team led by Roger Jowell at the Centre for Comparative Social Surveys, City University, London. Estonia participated in the second round of the ESS, which took place from December 2004 to January 2005. The survey was representative of all persons aged 16 and over (no upper age limit) resident within private households in Estonia, regardless of their nationality, citizenship, or language. The sample was selected by strict random probability methods at every stage and respondents were interviewed face-to-face. The survey was carried out by the Statistical Office of Estonia. It should be emphasized that the SISE was not included in the general ESS questionnaire (applied in all participating countries) but only in the Estonian survey. The Estonian version of the ESS questionnaire was administrated to 1469 native Estonians (59% females and 41% males) ranging from 16 to 98 years old, with the mean age of 47.9 years ($SD = 19.8$). Among these participants, about 44% had completed a basic education, 37% had finished secondary or a vocational school, and 19% had a university degree. The complete data for the SISE was available from 1457 participants (99% of the total sample).

Sample R3
The third representative sample of 655 adult individuals (52% females and 48% males) completed the Estonian version of the RSES in 2000 as a part of a wider study of the integration processes in Estonia. This project was coordinated by Jüri Kruusvall and Raivo Vetik from Tallinn University. The participants ranged from 15 to 74 years old, with mean age 42.0 years ($SD = 17.3$), and this sample was representative of the Estonian-speaking population in Estonia. In this sample, 36% had completed a basic education, 47% had finished secondary or a vocational school, 16% had a university degree, and about 1% declined to declare their educational level. The internal reliability coefficient (Cronbach $\alpha$) of the RSES was $\alpha = .81$. The survey was carried out by the Saar Poll Ltd., the second largest social and market research company in Estonia.

Sample R4
The representative sample of Estonian adolescents ($N = 2708$; mean age = 14.9 years, range 11 to 18, $SD = 2.04$) was drawn from 27 socially and geographically representative schools from all of the 15...
Estonian counties, including the capital city of Tallinn, smaller towns (e.g., Tartu) and rural areas in 2001. Similarly to general population at this age, the current sample contained approximately equal numbers of boys (47%) and girls (53%). The project was coordinated by the first and second authors of this article. The internal reliability coefficient was \( \alpha = .81 \).

**Sample IS**

This self-recruited Internet sample consisted of 23,248 Estonian-speaking individuals (63% females and 37% males; mean age = 29.3 29.3 years, range 9 to 89, \( SD = 8.73 \)) who visited a noncommercial, advertisement-free Web site and completed an online version of the RSES during 2000 to 2001. After filling in the entire questionnaire, the participants received the online feedback sheet with a comparison of their results with others’. Although the participants were asked to complete gender and age information, they did not provide any personal identification, therefore, complete anonymity was assured. In total, about 10% of all Estonian Internet users participated in this study, whereas about a third of the Estonian population from 15 to 74 years old had access to the Internet during the data collection according to the regular e-track survey conducted by the TNS Emor (see www.emor.ee, for details). In this sample, the internal reliability coefficient of the RSES was \( \alpha = .82 \).

**Table 1. Number of participants and descriptive statistics of the global self-esteem scales by age groups for the samples**

<table>
<thead>
<tr>
<th>Age groups</th>
<th>The Single-Item Self-Esteem Scale (SISE)</th>
<th>The Rosenberg Self-Esteem Scale (RSES)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R1</td>
<td>R2</td>
</tr>
<tr>
<td>9–12</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>13–17</td>
<td>94</td>
<td>3.06</td>
</tr>
<tr>
<td>18–22</td>
<td>118</td>
<td>3.03</td>
</tr>
<tr>
<td>23–29</td>
<td>152</td>
<td>3.22</td>
</tr>
<tr>
<td>30–39</td>
<td>257</td>
<td>3.28</td>
</tr>
<tr>
<td>40–49</td>
<td>227</td>
<td>3.10</td>
</tr>
<tr>
<td>50–59</td>
<td>196</td>
<td>3.06</td>
</tr>
<tr>
<td>60–69</td>
<td>253</td>
<td>3.13</td>
</tr>
<tr>
<td>70–90</td>
<td>98</td>
<td>3.19</td>
</tr>
<tr>
<td>Total</td>
<td>1395</td>
<td>3.14</td>
</tr>
</tbody>
</table>

*Note.* N = number of participants; M = mean; SD = standard deviation; \( N^{R3} \) = number of participants in the sample R3; \( N^{R4} \) = number of participants in the sample R4.
In this study, all the samples were divided into 9 age groups using intervals identical with that used by Robins and his colleagues (2002): 9–12, 13–17, 18–22, 23–29, 30–39, 40–49, 50–59, 60–69, and 70–90 years. Two representatives samples tested by the RSES—adults (R3) and adolescents (R4)—were combined for subsequent analyses to reflect the entire life span from 11 to 74 years, with a partial overlap in two age groups: 13–17 and 18–22 years (3% versus 97% and 12% versus 88% for the R3 and R4, respectively). The numbers of participants across age groups for the samples are reported in Table 1.

RESULTS

Age Differences in Self-Esteem Measured by the Single-Item Self-Esteem Scale

Figure 1 shows cross-sectional trajectories of global self-esteem across the life span for all samples. The third curve from the top (crosses) is reproduced from Robins et al. (2002; Table 2, p. 429) and is based on the data collected from 326,641 individuals over the Internet who expressed their agreement with a single statement “I see myself as someone who has high self-esteem” on a 5-point scale. The lowest curve (squares) present the mean levels of self-esteem across the age groups based on answers to the SISE of 2852 Estonians who were randomly selected from the whole population (averaged data of samples R1* and R2*). These two normative trajectories of self-esteem are not very similar. According to data reported by Robins and his colleagues (2002), self-esteem gradually rises through adulthood and reaches its highest peak in the 60s. Unlike Robins’s data (2002), two representative Estonian samples (R1* + R2*) tested by the SISE scale reach the maximum level in the 30s and demonstrate after that a steady decline toward lower self-esteem level in accordance with the data reported by McMullin and Cairney (2004). Concerning the mean levels, the multinational Internet sample demonstrates significantly higher scores compared to the Estonian representative samples in all age groups. However, the mean level differences between these samples are hard to interpret because of differences in language (English versus Estonian, respectively). Because Pearson correlation is insensitive to the mean level differences, we computed congruence coefficients between the mean levels of self-esteem trajectories of a multinational Internet sample and two nationally representative Estonian samples. The results confirmed that even a small number of cases is enough to claim that the form of trajectories is not identical since the coefficient was negative ($r(8) = -0.36, p = 0.382$).
It is noteworthy that although two Estonian samples, R1 and R2, were both randomly drawn from the National Census and exactly the same question was used, the mean levels of the SISE differed slightly across the life span. The maximum difference reached 0.51 scale among the oldest age group (age 70 to 79 years, $d = .51$, $p < .001$) and the analysis of variance (ANOVA) revealed that the overall difference between these two samples was statistically significant, $F(1, 339) = 13.44$, $p < .001$. After adjustment for educational level and gender, the analysis of covariance (ANCOVA) revealed a significant effect for the samples across age groups, $F(7, 2821) = 2.32$, $p < .05$. The mean values of the SISE scale across age groups for these samples are reported in Table 1.

Figure 1. The cross-sectional life span trajectories of global self-esteem measured by the RSES and SISE. Vertical bars denote 0.95 confidence intervals. Data for the multinational Internet sample are adapted with permission of the first author from “Global self-esteem across the life span,” by R. W. Robins, K. H. Trzesniewski, J. L. Tracy, S. D. Gosling, & J. Potter (2002), Psychology and Aging, 17, 429. Copyright 2002 by the American Psychological Association.
Age Differences in Self-Esteem Measured by the Rosenberg Self-Esteem Scale

In Figure 1 (top) are presented two cross-sectional age trajectories measured by the Estonian version of the RSES for the two types of samples: nationally representative sample and the Estonian self-recruited Internet test-takers. The upper curve (circles) demonstrates mean-level changes of self-esteem across the age range based on data received from 23,248 individuals who volunteered to fill in the questionnaire at an Internet site (I5) and the curve marked with triangles is an aggregate of two nationally representative (R3 + R4) samples.

Table 2. A variance component analysis of the RSES Items and the total scores of 26,611 individuals

<table>
<thead>
<tr>
<th>Items of the RSES</th>
<th>Variance component analysis (%)</th>
<th>Correlation r</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Error</td>
<td>Sex</td>
</tr>
<tr>
<td>(1) I feel I’m a person of worth, at least on an equal plane with others</td>
<td>84.77</td>
<td>1.02</td>
</tr>
<tr>
<td>(2) I feel I have a number of good qualities</td>
<td>90.27</td>
<td>0.24</td>
</tr>
<tr>
<td>(3) All in all, I am inclined to feel that I am a failure (R)</td>
<td>92.73</td>
<td>0.08</td>
</tr>
<tr>
<td>(4) I am able to do things as well as most other people</td>
<td>88.12</td>
<td>0.02</td>
</tr>
<tr>
<td>(5) I feel I do not have much to be proud of (R)</td>
<td>95.39</td>
<td>1.61</td>
</tr>
<tr>
<td>(6) I take a positive attitude toward myself</td>
<td>98.74</td>
<td>0.00</td>
</tr>
<tr>
<td>(7) On the whole, I am satisfied with myself</td>
<td>98.87</td>
<td>0.00</td>
</tr>
<tr>
<td>(8) I wish I could have more respect for myself (R)</td>
<td>93.71</td>
<td>0.08</td>
</tr>
<tr>
<td>(9) I certainly feel useless at times (R)</td>
<td>99.22</td>
<td>0.00</td>
</tr>
<tr>
<td>(10) At times I think that I am no good at all (R)</td>
<td>86.30</td>
<td>0.05</td>
</tr>
<tr>
<td>Total Score</td>
<td>88.85</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note. The participants’ gender, age group, and the type of samples (nationally representative R3 + R4 versus self-recruited Internet I5) are treated as three random variables. (R) = Reversed scored item; interactions = Sex × Age, Sex × Sample, Age × Sample, and Sex × Age × Sample; r = Pearson correlation coefficient between age and the RSES items and the total score; *** p < .001.

Age Differences in Self-Esteem Measured by the Rosenberg Self-Esteem Scale

In Figure 1 (top) are presented two cross-sectional age trajectories measured by the Estonian version of the RSES for the two types of samples: nationally representative sample and the Estonian self-recruited Internet test-takers. The upper curve (circles) demonstrates mean-level changes of self-esteem across the age range based on data received from 23,248 individuals who volunteered to fill in the questionnaire at an Internet site (I5) and the curve marked with triangles is an aggregate of two nationally representative (R3 + R4) samples.
Generally, self-esteem level of adolescents in the self-recruited Internet sample (I5) rises gradually until it reaches a plateau in young adulthood and retains this level (4% ± 1.5% points) till the end of the age range (correlation with age was about zero). Thus, opposite to the multinational Internet sample (Robins et al., 2002), self-esteem is at its lowest level during childhood among the Estonian self-recruited Internet sample.

The mean values of global self-esteem scale are systematically higher in the Estonian self-recruited Internet sample compared to nationally representative samples. Namely, the mean score of the RSES of the representative sample is systematically and statistically significantly ($p < .05$) below the mean self-esteem level in the Estonian self-recruited Internet sample across the age groups from 13 to 59 years (Cohen’s $d = .18–.47$). In total, the RSES mean in the Internet sample is statistically significantly ($p < .0001$) higher compared to the nationally representative samples, $d = .42$ (see Table 1).

Therefore, the results of this study indicate that Estonians with a higher level of self-esteem are more likely to answer the self-esteem questionnaire posted on the Internet than those who are randomly or semi randomly selected from the general population. What could be a possible explanation for this discrepancy? One possibility is that the Estonian self-recruited Internet sample mostly consisted of individuals who were better educated and positioned at the higher level of the social stratification than average, as it is known that such individuals have generally higher global self-esteem (see Twenge & Campbell, 2002). As the respondents of the Internet sample did not provide information about their educational level, this hypothesis cannot be tested in this study. However, the results of ANCOVA revealed that in the adult representative sample (R3) more highly educated individuals tended to have a higher self-esteem after adjustment for gender, $F(2, 646) = 5.14$, $p < .01$. More specifically, participants with a university degree had statistically significantly ($p < .0001$) higher self-esteem measured by the RSES ($M = 4.02$, $SD = .55$) compared to adults who had completed only basic or secondary level of education ($M = 3.77$, $SD = .67$ and $M = 3.86$, $SD = .64$, respectively). Interestingly, there was no significant difference in mean scores of the RSES between individuals in the Internet sample and participants with a university degree in the representative adult sample ($d = .03$, ns). To conclude, the results of the current study confirm that Estonians with a higher educational level have higher self-esteem than persons with less educational attainment. As the mean level of the RSES in the self-recruited Internet sample is comparable to those who had a university degree in the nationally
representative sample, it is very likely that the Internet sample also consisted mostly of highly educated participants.

**Item Life Span Trajectories**

The mean scores of the RSES are considerably higher than the mean value of the SISE for all of the samples. This indicates that individuals are inclined to agree with the SISE statement “I have high self-esteem” more reluctantly than with most of the other 10 items included in the Estonian version of the RSES. Indeed, a more detailed inspection shows that the individual RSES items have not only different levels of endorsement but they also have quite different life span trajectories: some of the items are relatively stable with minimal variations in the mean levels from childhood to older age, the means of other items increase across age groups, and, finally, the third group of items reach the maximum value in early adulthood and start to decline afterwards. The cross-sectional developmental trajectories across different age groups for the all 10 RSES items for a representative (R3 + R4) and self-recruited Internet sample (I5) are presented in Figure 2.

Next we investigated what amount of variance in respondents’ self-esteem scores could be attributed to differences in age, sex, and sampling procedures as well as to the interactions of these effects. To answer this question, a variance component analysis with minimum variance quadratic unbiased estimation (MVQUE) was performed for the self-esteem measure scores. MVQUE is a variation of the restricted maximum likelihood estimation technique and it was chosen because there is no weighting of the random effects, so an iterative solution for estimating their variance components is not required. Like correlation, the basic goal of variance component estimation is to assess the population covariation between random factors and a dependent variable. The advantage of MVQUE is that it is not limited to linear relationships between random factors and dependent variables.

In our first analysis, the RSES scores of 26,611 individuals are treated as a dependent variable and the participants’ gender, age group, and the type of samples (nationally representative R3 + R4 versus self-recruited Internet I5) as three random variables. The results show that less than 11% of the total variance in the RSES scores is attributable to the random factors (i.e., age group, gender, and sample type), only 0.41% to the interaction between various random factors, 7.74% to differences in the type of sample, and, finally, only 3% to age differences (Table 2). Interestingly, there
are two items (item 1, “I feel I’m a person of worth, at least on an equal plane with others” and item 10, “At times I think that I am no good at all”) that demonstrate age dependence exceeding the level of 3%. The endorsement of these two items increases with age (Table 2, last column). Moreover, these two items also demonstrate the largest effect of the sample, 10% and 8%, respectively, indicating that the items that are most sensitive to age are simultaneously most vulnerable to the sampling bias. Five items, including both positively and negatively worded statements, show very modest (less than 1%) total variance caused by age.

Finally, we analyzed similarity of the life span trajectories by a correlational analysis. The cross-correlations between the item trajectories of the representative (R3 + R4) and self-recruited Internet sample (I5) are shown in Table 3. The Pearson correlation evaluates only similarity in shape being insensitive to the mean level and scatter. On the basis of this, it is possible to claim that the first five RSES

![Figure 2. The mean values of the RSES items across different age groups for the representative (left panel) and self-recruited Internet (right panel) samples. Item numbers correspond to those in the original RSES (Rosenberg, 1965) and in the Estonian version of the RSES (Pullmann & Allik, 2000); see Table 2.](image-url)
items in the Internet sample behave differently from the respective items in the representative sample (the main diagonal in Table 3). Only trajectories of three items, items 6 (“I take a positive attitude toward myself”), 9 (“I certainly feel useless at times”), and 10 (“At times I think that I am no good at all”), are sufficiently similar in the both samples. If one observes the trajectory of the item 6, self-esteem would appear to increase gradually from age 12 to 90 through the whole life span. This is a replicable finding: the correlation across the nine age groups for random versus Internet samples is $r = .96$
(p < .001). But if one chooses the item 9, the conclusion may be that self-esteem grows in the first half of life, afterwards at age 30 starts to decline, and again to increase after age 50 (except for a drop after age 70). Similarly, the latter pattern is also replicable across sample types, $r = .71$ ($p < .05$). However, these two patterns of age trends in general self-esteem across the life span are unrelated, $r = .31$ ($p = .41$) and $r = .39$ ($p = .29$), respectively. This is a demonstration that single items may provide replicable findings in large samples, but these are not necessarily generalizable to other, ostensibly similar items.

**DISCUSSION**

Taken together, this research makes three essential contributions to the debate whether there are systematic normative age differences in global self-esteem. Firstly, we presented some results that challenge generalization of the recently reported self-esteem trajectory (Robins et al., 2002; Robins & Trzesniewski, 2005; Trzesniewski, Robins, Roberts, & Caspi, 2004) across different cultures. Second, we provided evidence that the self-recruited Internet data collection method is essentially biased compared to a random sampling from the general population. Finally, we found that different items measuring global self-esteem have dissimilar trajectories across the life span and therefore, the single-item measures, like the SISE (Robins, Hendin, et al., 2001), may not be generalizable across samples and comparable in similar way concerning questionnaires consisting of several self-esteem items.

**Does Global Self-Esteem Have a Universal Trajectory across the Life Span?**

Surprisingly, none of the four normative self-esteem trajectories in this study resembled any other trajectory portrayed in Figure 1. Curiously, the most dissimilar were trajectories of two large Internet samples, the multinational English-speaking Internet sample (Robins et al., 2002) and the Estonian Internet sample (I5; this study). There are many possible reasons for this discrepancy. The first possibility, of course, is the difference between instruments used for measuring of global self-esteem. The multinational study was based on a single answer to the SISE item, whereas Estonian participants responded to the RSES containing 10 items. As we demonstrated, the normative trajectories of individual items through age were remarkably different indicating that the wording of self-esteem items is sensitive to age. This suggests that different age groups have a tendency to respond differently to the RSES items. For example, older respondents agreed
less frequently with the statement about feeling useless from time to time (item 9) compared to younger participants. At the same time, some other items (e.g., item 7: “On the whole, I am satisfied with myself”) did not exhibit remarkable differences between different age groups. There have been several previous attempts to differentiate various facets inside global self-esteem. For instance, it was proposed that global self-esteem consists of two subfactors, positive and negative self-esteem, whereas positively and negatively worded items in the RSES have a tendency to group into two separate factors (e.g., Kohn & Schooler, 1969; Owens, 1994). However, the division of the self-esteem trajectories into distinct types does not correspond to the distinction between positive and negative self-esteem. Anyhow, the results of this study support the opinion proposed by several researchers (e.g., Carmines & Zeller, 1979; Corwyn, 2000; Dunbar, Ford, Hunt, & Der, 2000; Greenberger, Chen, Dmitrieva, & Farruggia, 2003; Horan, DiStefano, & Motl, 2003; Marsh, 1996; Tomas & Oliver, 1999) that the wording of items has a systematic and enduring effect on the measurement of global self-esteem.

In addition to the differences in measurement instruments (i.e., SISE versus RSES), the discrepancy between the results of the multinational and Estonian Internet samples could be due to cultural and/or language differences, the pinpointing and explanation of which is beyond the scope of this article. Furthermore, the life span trajectories of two Estonian samples, R1* and R2*, randomly drawn from the National Census were rather dissimilar, although the identical single-item measure was used. This may mean that there is no universal cross-sectional self-esteem trajectory through the age range and the observed differences between different age groups remain well within the limits of the overall measurement accuracy. Thus, the results of this study do not support the idea of a single and invariable trajectory of global self-esteem across the human life span. If these data support any trajectory, it is a flat one with some random perturbations from the mean level.

**How Typical are Self-Recruited Internet Samples?**

Gosling and his colleagues (2004) argued that accumulating evidence indicates on the consistency between the Internet and traditional pencil-and-paper methods of gathering psychological data. Previous studies have shown, for instance, that the administration mode, traditional or computer-based, had little effect on the psychometric properties of the RSES (Vispoel, Boo, & Bleiler, 2001). This study supported this notion; the internal reliability of the self-esteem
measure was virtually identical in the Internet and traditional pencil-and-paper samples. Also, the congruence between factor structures of the RSES in these two samples was very high, consistent with near-perfect equivalence.

Nevertheless, the mean values of global self-esteem measures were systematically different in the self-recruited Internet sample and other samples representing more adequately the entire Estonian population. Most revealing was the contrast between the Estonian self-recruited Internet (I5) and other nationally representative (R3 + R4) samples. This finding demonstrated that those participants who, on their own initiative, discovered the questionnaire on the Internet and volunteered to answer scored generally higher on the RSES scale than those who were randomly selected from the National Census.

Why do Internet samples score systematically higher on self-esteem questionnaires compared to nationally representative samples? In 2000, when the Estonian Internet study was carried out, about a third of the Estonian population had an access to the Internet (compared with 52% in the United States and 23% on average in European Union weighed by population; Bauer, Berne, & Maitland, 2002). However, just 4 years later, already 52% of 16- to 74-year-old Estonians were using the Internet (TNS Emor, Spring 2004). A regular survey conducted by the TNS Emor revealed that, compared with the entire population, Estonian Internet users in 2000 had a better education (15 to 16 study years on average) and contained a disproportional number of managers, top officials, and executives (about a half of all Internet users).

This fact could explain the difference between the general population and Internet users because it is known that the individuals with higher income and better education have higher levels of self-esteem than persons with lower income and less educational attainment (McMullin & Cairney, 2004; Twenge & Campbell, 2002). Generally, both education and income have been found to be highly correlated with Internet use. Pettit (1999) examined whether respondents of an online version of a psychological questionnaire were an unusual set of volunteers and found that 90% of the respondents indicated that their standard of living was higher compared with that of others. Also, in terms of the highest level of education obtained, these respondents tended to be fairly well educated—almost 17% of the volunteers were postgraduates, more than 37% had obtained bachelors degrees, more than 18% had obtained community college education, and more than 25% had been educated at high school. Similarly, according to a report published by the U.S. National Telecommunications and Information Administration in
1999, although 61.6% of those with a college degree used the Internet, only 6.6% of those with elementary school education or less used the Internet (National Telecommunications and Information Administration, 1999). However, more research is needed on this topic as well as on the relationship between self-esteem and educational level throughout the life course.

The analyses of this study revealed that the self-esteem level of participants in the representative samples was similar to that of individuals, who on their own initiative had volunteered to answer the self-esteem questionnaire over the Internet. Therefore, excluding perhaps the youngest and the oldest age groups, we can conclude that the Estonian self-recruited Internet sample mostly consisted of individuals who were better educated and positioned at higher level of social stratification, which may account for their generally higher level of global self-esteem. The analysis provided an indication that this might very likely be the case: there were no significant differences between the two groups when the mean scores of the Internet sample were compared with the mean scores of those in the representative sample who had a university degree. Nevertheless, as educational level of self-recruited Internet respondents in the Internet sample was not provided, it was impossible to substantiate this hypothesis unconditionally in the current study.

**Normative Stability of Self-Esteem**

Research on general self-esteem development throughout human life has produced inconsistent results. For example, some studies have found a mean-level rise of self-esteem during adolescence whereas other studies registered no change or even a drop in self-esteem (cf. O'Malley & Bachman, 1983). The reason for these discrepancies may be in nonequivalent measuring instruments or variation in the tested samples (Robins et al., 2002; for cohort effects see Twenge & Campbell, 2001). This study, however, points in the direction of another interpretation, according to which the main cause of inconsistent results is the lack of a strong and systematic relationship between age and global self-esteem. More specifically, possible age effects are too small to be observed, even in large samples. Although individuals may have different developmental self-esteem trajectories—besides those who are consistent over time, there are others whose perception of self-worth is decreasing or increasing over time (cf. Zimmerman, Copeland, Shope, & Dielman, 1997)—the opposite tendencies seem to cancel each other out when the mean scores are studied on a cross-sectional level. It appears that the
magnitude of the age-related changes is considerably smaller than variations in the mean level of self-esteem produced by sampling or even small changes in the wording of items. For example, the greatest difference between the means of two age groups studied by Robins and his colleagues (2002) was 0.41 scale points, which is about 9% of the whole range of the 5-point SISE measure. Due to the fact that the standard deviation of the whole sample in the mentioned study was 1.31 scale units, the maximal age-related change they were able to observe was approximately one third of the standard deviation. The results of this study suggested that belonging to a certain age group accounted only for 3% of the variance in the mean levels of the RSES total scores while the influence of sampling (self-recruited versus random sampling) had a much greater (7.74%) effect on the mean values of the self-esteem.

Most theories of self-esteem suggest that feeling of self-worth originates either from life experiences or interpersonal relationships. For example, it was assumed that self-esteem is related to achievement and mastery experience. In turn, failures in personal life contribute negatively to the self-esteem (Rosenberg, 1979). According to the sociometer theory, however, people have a fundamental motive to maintain connectedness with other people and for that purpose they routinely monitor how much they have been accepted or rejected by others (Baumeister & Leary, 1995; Leary, Haupt, Strausser, & Chokel, 1998). People’s feelings about themselves reflect, in a substantial extent, how they believe that they are perceived and evaluated by others (cf. Shrauger & Schoeneman, 1979). All these approaches assume that the explanation of individual variation of self-esteem can be found in environmental factors—socialization, life experience, or culture. Contrary to this perspective, another line of findings indicates that a substantial part of self-esteem can be understood as an enduring tendency to feel and think about oneself. From one third to one half of the variance of this tendency is controlled by genes (Kendler, et al., 1998; Roy, et al., 1995). Like most personality traits, shared environmental influences on self-esteem are small or negligible (Neiss, et al., 2002). The fact that self-esteem does not arise solely from environmental circumstances is also supported by observations that rank-order stability is similar to other personality traits (Trzesniewski et al., 2004). Most therapeutic interventions or school programs aiming to boost self-esteem have demonstrated a resistance to change and have produced only temporary or limited benefits (Baumeister, Campbell, Krueger, & Vohs, 2003). Although cross-cultural psychologists have argued that the way how selfhood is constructed differs fundamentally across regions of the world, a
recent large-scale comparative study has been more favorable for the theoretical position holding that self-esteem is a universal phenomenon that most likely stems from common human motivations: the internal reliability and factor structure of the RSES was generalizable across 53 nations and its scores correlated in an identical way with the key personality traits of neuroticism and extraversion (Schmitt & Allik, 2005).

Global self-esteem appears to be a relatively enduring individual difference that is strongly associated with core personality dimensions. Many studies have shown that scores on the RSES correlate significantly with two of the Big Five dimensions, Neuroticism and Extraversion (Costa et al., 1991; Judge et al., 2002; Kwan et al., 1997; Pullmann & Allik, 2000; Robins, Tracy, et al., 2001). The correlation between self-esteem and emotional stability is comparable to the average relationship between alternative measures of neuroticism. For instance, according to a recent meta-analysis of the published articles in 10 leading journals, the average correlation between self-esteem and Neuroticism was –.64 (Judge et al., 2002). In a large comparative study involving 53 nations, the mean uncorrected correlations of the RSES scores with the Neuroticism scale of the Big Five Inventory (Benet-Martinez & John, 1998) was –.41 across 53 studied nations (Schmitt & Allik, 2005). For a comparison, correlations between six facets of the NEO-PI-R that measure Neuroticism range from –.31 to –.64, with the mean value of –.48 (Costa & McCrae, 1992, Appendix F). A particularly strong negative correlation with the Neuroticism score suggests that the main function of positive self-evaluation is to buffer and protect its holder from frustration and anxiety (Pyszczynski, Greenberg, Solomon, Arndt, & Schimel, 2004; Rosenberg, Schooler, Schonbach, & Rosenberg, 1995). These observations seem to indicate that the uniqueness of self-esteem measures is overwhelmed by its commonality with other measures of emotional stability (Judge et al., 2002). Thus, measures of global self-esteem seem to represent reasonably well a general neuroticism factor as it is conceptualized in the Big Five with perhaps some contamination from extraversion.

Placed into a broader context of personality traits, the relative stability of global self-esteem across the life span is less surprising. The Five Factor Theory (McCrae & Costa, 1996, 1999) was proposed in order to explain an extraordinary stability of personality traits in time (McCrae & Costa, 2003), a powerful effect of genes and vanishingly small effect of the shared environment on personality traits (Plomin & Caspi, 1999), and their transcendence across boundaries of culture (Allik & McCrae, 2002). Based on cross-sectional and longitudinal
studies, Costa and McCrae (2002) proposed that on the background of a general temporal stability of personality traits, there are still specific age curves for each personality trait. Among all five personality dimensions, neuroticism is the most stable across the life span: in some countries, such as the United States (Costa & McCrae, 2002; McCrae et al., 2004), the mean level of self-reported neuroticism decreases with age and in some other countries, such as Estonia and Italy, it remains basically the same across the whole life span (Allik, Laidra, Realo, & Pullmann, 2004; Costa, McCrae, Martin et al., 2000) or have a curvilinear trend (McCrae et al., 2004). If the main function of positive self-evaluation is to buffer and protect its holder from frustration and anxiety, then it is expected that self-esteem, like other components of emotional stability, is also relatively stable across the life span. This provides a support to a standard view, according to which normative age differences in global self-esteem are small and most likely beyond our ability to measure them.

Finally, the results of this study warn about limitations of the self-recruited Internet studies. Firstly, such convenience-type studies might be better suited for testing theoretical ideas concerning relationship between various concepts rather than providing point estimates of parameters that are intended to apply to a population. Secondly, biases related to Internet samples seem to affect mean levels more than patterns of covariation (i.e., mean levels are more affected than, say, factor structures).

REFERENCES


Costa, P. T., Jr. & McCrae, R. R. (2002). Looking backward: Changes in the mean levels of personality traits from 80 to 12. In D. Cervone & W. Mischel (Eds.), *Advances in personality science* (pp. 219–237). New York: Guilford.


