Accuracy, Consensus, In-Group Bias, and Cultural Frame Shifting in the Context of National Character Stereotypes

Jan-Erik Lönnqvist a, Kenn Konstabel b, Nellystiina Lönnqvist a & Markku Verkasalo a

a University of Helsinki
b University of Tartu

Accepted author version posted online: 20 Sep 2013. Published online: 13 Dec 2013.

To cite this article: Jan-Erik Lönnqvist, Kenn Konstabel, Nellystiina Lönnqvist & Markku Verkasalo (2014) Accuracy, Consensus, In-Group Bias, and Cultural Frame Shifting in the Context of National Character Stereotypes, The Journal of Social Psychology, 154:1, 40-58, DOI: 10.1080/00224545.2013.843500

To link to this article: http://dx.doi.org/10.1080/00224545.2013.843500

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
Accuracy, Consensus, In-Group Bias, and Cultural Frame Shifting in the Context of National Character Stereotypes

JAN-ERIK LÖNNQVIST
University of Helsinki

KENN KONSTABEL
University of Tartu

NELLYSTIINA LÖNNQVIST
MARKKU VERKASALO
University of Helsinki

ABSTRACT. We examined Finns’ and bilingual Swedish-Finns’ stereotypes regarding personality differences between Finns and Swedish-Finns and compared them with their respective self-ratings. Stereotype ratings by both groups converged on depicting Swedish-Finns as having a more desirable personality. In-group bias also influenced stereotypes. Contrary to predictions based on the Stereotype Content Model, out-group stereotypes were not compensatory. Consistent with the kernel of truth hypothesis of national stereotypes, Swedish-Finns’ aggregate self-ratings resembled their stereotype of personality differences between the two groups, and their personality self-ratings were more desirable than Finns’ self-ratings. Tentatively suggesting the occurrence of cultural frame shifting, the resemblance between Swedish-Finns’ self-ratings and their stereotype of Swedish-Finns was, although only marginally statistically significantly, somewhat stronger when the self-ratings were provided in Swedish.

Keywords: cultural-frame shifting, intergroup perceptions, national character stereotypes, personality, social content model, social desirability

RECENT STUDIES ON INTERGROUP PERCEPTION have fallen into two broad and somewhat overlapping categories: (a) those dealing with social perception as a function of group membership and intergroup relations, with a particular interest in perceptions of in-group and out-group warmth and competence (e.g., Cuddy, Fiske, & Glick, 2008; Fiske, Cuddy, Glick, & Xu, 2002; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005; Yzerbyt, Provost, & Corneille, 2005), and (b) those examining the possible accuracy of rather stable ethnic or national character stereotypes in terms of a wider range of personality traits, with less emphasis on intergroup relations as determinants of stereotypes (e.g., Abate, & Berrien, 1967; Peabody, 1985; Realo...
et al., 2009; Terracciano et al., 2005). These two lines of research have emphasized the involvement of somewhat different mechanisms in stereotype formation.

Perhaps the most prominent model to have emerged from research on stereotypes as a function of intergroup relations is the Stereotype Content Model (SCM; e.g., Fiske et al., 2002; Cuddy et al., 2008). This model defines two fundamental dimensions of social perception, warmth and competence. These two dimensions are predicted, respectively, by intergroup competition and status differences. The model posits that in-groups and out-groups are judged differently, with only in-groups (e.g., students) and the most prototypical groups (e.g., middle class whites) stereotyped as high on both warmth and competence. Out-groups, on the other hand, will either receive more ambivalent stereotypes (positive on one dimension and negative on the other; e.g., elderly as high on warmth, low on competence; rich as low on warmth, high on competence), or be stereotyped as low on both dimensions (e.g., welfare recipients).

National character stereotypes are defined as shared beliefs about the features of a typical representative of a particular nation (e.g., Terracciano et al., 2005). Contrasting the above view that in-groups and out-groups are judged differently, research on national character stereotypes has emphasized that auto-stereotypes, referring to stereotypes of one’s in-group, and heterostereotypes, referring to stereotypes of an out-group, are highly similar to each other (e.g., judges from around the world agree with Americans’ in-group ratings on the American personality profile; Terraciano & McCrae, 2008) and also relatively stable (evidence of stability over time is perhaps best provided by a series of studies known as the Princeton trilogy and its extensions; see Katz & Braly, 1933; Gilbert, 1951; Karlins, Coffman, & Walters, 1969; Madon et al., 2001). Furthermore, some research on national stereotypes suggests that such stereotypes may have a “kernel of truth” (Allport, 1978/1954; Brigham, 1971; Realo et al., 2009), although the empirical evidence is far from uniform (Church, & Katigbak, 2002; Terracciano et al., 2005).

In the present research, we examined Finns’ and bilingual Swedish-Finns’ stereotypes regarding personality differences between Finns and Swedish-Finns. Based on SCM, Finns’ and Swedish-Finns’ stereotypes could be expected to differ, with members of both groups rating their in-group as possessing more desirable characteristics than the other group. But the opposite prediction—that the groups will reveal highly similar stereotypes—can be derived from research on national character stereotypes.

Besides comparing the stereotypes of Finns and Swedish-Finns, we examined the hypothesis that national character stereotypes reflect actual differences in self-rated personality traits (Allport, 1978/1954; Realo et al., 2009; Terracciano et al., 2005). Moreover, we investigated whether cultural mindset can influence personality self-ratings (e.g., Oyserman & Lee, 2008; Ramírez-Esparza, Gosling, Benet-Martínez, Potter, & Pennebaker, 2006). Such an effect could contribute to explaining the sometimes contradictory results pertaining to the accuracy of national character stereotypes. We therefore examined whether personality changes as a function of test administration language, and whether the changes are predictable from in-group stereotypes (see Chen & Bond, 2010).

Predictions Based on the Stereotype Content Model

The SCM (Cuddy et al., 2008) defines two fundamental dimensions of social perception, warmth and competence. A basic tenet of SCM is that in-groups and out-groups are evaluated differently. More specifically, in-groups tend to be rated more favorably in terms of being rated as high on
both warmth and competence. Based on SCM, one would thus expect both Finns and Swedish-Finns to rate their in-group as high on warmth and competence. This prediction is also consistent with the more general and extensive literature on bias in favor of in-groups (Hewstone, Rubin, & Willis, 2002).

Regarding out-group stereotypes, SCM suggests that perceptions of out-groups’ competence and warmth are predicted, respectively, by perceived intergroup differences in status and by competition between the groups. There are clear status differences between Finns and Swedish-Finns. The latter are the most important minority group in Finland, comprising around 6% of the 5.4 million population. Their history goes back to the time when Finland was part of the Swedish kingdom (around 1150 to 1809), a time during which almost all of the noble families in Finland were Swedish-speaking. Traces of this period are to be found in the economic power that Swedish-Finns still exhibit (e.g., Roos & Roos, 1984). Swedish-Finns are generally regarded as one of the least conflict-ridden and most privileged minorities in the world (e.g., McRae, 1999). They enjoy all the same rights as does the majority, and do very well on indices that represent health and quality of life. For instance, Swedish-Finns, as compared to the Finnish majority, divorce less frequently (Finnäs, 1997), are more employable (Saarela & Finnäs, 2003), and have more social capital (Hyyppä & Mäki, 2001), which leads to lower morbidity, disability and mortality (Hyyppä & Mäki, 2003). Importantly, Swedish-Finns are also stereotypically perceived to be both the economic and cultural elite. For instance, in a recent study, almost 90% of both Swedish-Finns and Finns perceived Swedish-Finns as having higher status (Liebkind, Henning-Lindblom, & Solheim, 2008; see also Kivistö & Mäkelä, 1967; Peltonen, 1998). Such status differences could, based on SCM, be expected to influence both Finns’ and Swedish-Finns’ stereotypes. The Finns, comparing themselves with a higher status out-group, should rate the Swedish-Finns as highly competent. By contrast, the Swedish-Finns, comparing themselves with a lower status out-group, should rate the Finns as low in competence.

Perceptions of intergroup competition, on which perceptions of warmth would be expected to be based, are more difficult to ascertain than status differences. However, out-group stereotypes are, according to SCM, usually compensatory, meaning that such stereotypes typically reveal a negative relation between perceptions of warmth and competence (Yzerbyt et al., 2005). In the present context, regarding Finns’ out-group stereotypes, such a compensatory relation would suggest that Swedish-Finns, if perceived as high in competence due to intergroup status differences, should be perceived as low in warmth. Swedish-Finns would thus be viewed by Finns in the same way that groups such as the Asians, the rich, or Jews, are typically viewed according to SCM (see e.g., Cuddy et al., 2008). By contrast, Swedish-Finns, if perceiving Finns as low in competence due to inter-group status differences, should perceive Finns as high in warmth. Consequently, Finns would be viewed by Swedish-Finns in the same way that groups such as the disabled, traditional women, and elderly, are typically viewed according to SCM.

Predictions Based on Research on National Character Stereotypes

The SCM literature has typically not dealt with national stereotypes. These can be defined as shared beliefs about the features of a typical representative of a particular nation (e.g., Terracciano et al., 2005). To give some examples, Americans are widely believed to be sociable (e.g., Terraciano & McCrae, 2007), Japanese intelligent and industrious (Gilbert, 1951), and Germans thrifty and self-controlled (e.g., Peabody, 1985). In contrast to the in-group favoritism that one
would expect based on SCM, research on national stereotypes suggests high agreement between in-group and out-group ratings of national character. For instance, Abate and Berrien (1967) found very high agreement between Japanese and American raters on the national character of those two cultures, and similar results were reported by Church and Katigbak (2002) for American and Filipino raters. Peabody (1985), in one of the largest studies in this field of research, found that European raters from several countries generally agreed on the national character of European nations, the United States, and Russia. In another large-scale study, Terracciano and McCrae (2007) examined perceptions of the “typical American” in 49 cultures, and found strong consensus between how Americans rated themselves and were rated by others.

Besides consensus, some prior research on national character stereotypes has suggested that these stereotypes may, like some other stereotypes (e.g., gender stereotypes; see Costa, Terraciano, & McCrae, 2001), generally have a kernel of truth (e.g., Abate, & Berrien, 1967, Allport, 1978/1954; Brigham, 1971). However, some recent results have challenged this view. In a large scale cross-cultural survey conducted by Terraciano and colleagues (2005), respondents from 49 countries rated the personality of a typical member of their cultural group. These stereotype ratings were compared with the average self- or peer-rated aggregate personality scores of members of the same culture. The stereotype ratings were found to be uncorrelated with the aggregate personality profiles within almost all cultures.

A methodologically different approach, closer to the one that we adopt, was taken by Church and Katigbak (2002). They asked expert judges—i.e., people who had lived in both cultures under investigation (the United States and Philippines) to judge the personality traits of Filipinos and Americans. Although the judges agreed with each other on which traits characterized Filipinos and which Americans, their judgments did not correspond with mean personality trait differences in self-ratings between American and Filipino samples. Thus, the results of Terraciano and colleagues (2005) and Church and Katigbak (2002), although obtained with very different methodologies, converged on the conclusion that national stereotypes, although widely agreed upon, do not reflect mean personality ratings. However, even more recently, Realo and colleagues (2009) presented results from a seven country study that showed that if all assessments are made using the same measurement instrument, national character stereotypes are moderately similar to self-rated aggregate personality traits.

Cultural Frame Shifting

The result that national character stereotypes do not correspond with mean personality trait differences in self-ratings could be considered surprising in light of recent research that emphasizes the influence of culture on core psychological constructs—shifting the salience of certain aspects of culture (i.e., priming culture) has been shown to influence, for instance, values, self-concept, and cognitive processes (e.g., Oyserman & Lee, 2008). One could thus expect culture to also influence personality. Such cultural influences could be expected to result in some degree of correspondence between conceptions of what is typical of the culture and personality self-ratings.

The evidence suggesting cultural influences on psychological constructs comes primarily from studies on characteristics that have been argued to be more prone to be influenced by cultural factors, such as personal values, whereas personality traits may be more innate and stable (e.g., McCrae & Costa, 2008). However, in a sample of Spanish-English bi-linguals, Ramírez-Esparza et al. (2006) showed cultural frame shifting (CFS; Hong, Morris, Chiu, & Benet-Martínez, 2000)
in the domain of personality. Specifically, bilingual participants showed different personalities when responding in English and in Spanish, and these differences were consistent with differences between English and Spanish speaking cultures. This result supports the idea that test administration language could cue cultural mindsets—mental representations containing culture-congruent content, procedures, and goals (Oyserman & Lee, 2008). Also consistent with this idea, Chinese students studying in Canada reported values different from those of European Canadians only when responding in Chinese (Ross, Xun, & Wilson, 2002; see also Hong et al., 2000). The language of the personality measure could thus present a cultural cue that activates the goal of making culturally congruent responses.

Introducing stereotypes into research on CFS, Chen and Bond (2010) showed that Hong Kong Chinese-English bilinguals’ stereotypes of the typical personality traits of English and Chinese speakers were to some extent consistent with self-rated personality differences between responses in Chinese and responses in English. In the present research, we aimed to investigate whether, in our sample of bi-cultural Swedish-Finns, cueing different culturally based mindsets could lead to self-ratings that differed in how strongly they corresponded with the stereotype of Swedish-Finns. The Finn and Swedish-Finn cultural mindsets were cued by administering the same standard personality measure both in Finnish and in Swedish.

The Present Research

The present research asked Finnish and Swedish-Finnish students attending gymnasium (roughly corresponding to high school in the United States; see below for more details) to rate personality trait differences between Finns and Swedish-Finns. Based on research on intergroup relations, one could expect both groups to perceive their respective in-groups as having more desirable personality characteristics than the out-group. To investigate this, we examined the social desirability of both groups’ stereotypes of Finns and Swedish-Finns.

Besides the general occurrence of in-group bias, the SCM allows for the more specific prediction that the in-group will be judged as both warm and competent. Furthermore, due to the status differences described above, Finns could be expected to rate Finns and Swedish-Finns as similar in competence, but to rate Swedish-Finns as colder than Finns. By contrast, Swedish-Finns could be expected to rate both groups as similar in warmth, but to rate Finns as less competent. However, conflicting with these predictions, based on the literature on national character stereotypes one could expect Finns’ and Swedish-Finns’ ratings to converge on a shared stereotype of group differences.

To assess the kernel of truth hypothesis, we also gathered self-ratings from all our respondents. To enable mean level comparisons between the groups, all self-ratings were collected in Finnish (we tested the Finnish language skills of our Swedish-Finn participants to ensure that they were bilingual). Related to this issue, we also examined the possible influence of cueing cultural mindset on the personality self-ratings of our Swedish-Finn participants. For this purpose, our sample of Swedish-Finns was also administered a Swedish translation of our personality instrument. We expected responses given in Swedish to be more similar to perceived personality differences between Finns and Swedish-Finns.

Recent research on national stereotypes has conceptualized personality in terms of the Five-Factor-Model (FFM) of personality structure. This is currently the most popular framework within which studies on national character stereotypes are being conducted (e.g., McCrae,
Terracciano, Realo, & Allik, 2007; Realo et al., 2009; Terraciano et al., 2005; Terraciano & McCrae, 2007). Pertinent to our predictions derived from the SCM, both Warmth and Competence—the two fundamental dimensions of social perception identified by the SCM—are part of the FFM of personality structure. Within the FFM, Warmth is defined as the facet of Extraversion most relevant to interpersonal intimacy, and warm people are described as affectionate, friendly, cordial, and hearty (Costa & McCrae, 1992). This definition is quite similar to the one employed by SCM, which defines warmth as friendliness, helpfulness, sincerity, and trustworthiness (Fiske, Cuddy, & Glick, 2007). Competence, within the FFM, refers to being capable, sensible, prudent, and effective (Costa & McCrae, 1992). Within the literature on the SCM, the competence dimension reflects traits such as intelligence, skill, and efficacy (Fiske et al., 2007). The FFM and the SCM thus define warmth and competence very similarly, allowing us to employ the FFM facets warmth and competence in investigating hypotheses derived from the SCM.

METHOD

Participants and Procedure

Participants were recruited by their teachers from three Finnish language gymnasiums (located in Helsinki, Tampere, and Seinäjoki) and five Swedish language gymnasiums (located in Sipoo, Porvoo, Loviisa, Pori, and Vantaa) spread across Finland. Gymnasiums provide secondary education that prepares students for higher education. Currently, around half of the Finnish population will enroll at the gymnasium the year they turn sixteen, with the rest primarily seeking vocational education. The data were collected in 2009 and 2010. The questionnaires were administered on-line in various classes (e.g., psychology, philosophy, religion). Participation was voluntary. As an incentive to participate, participants were given feedback on their personal values profiles. Altogether 129 (44 men) students (out of the 140 who were approached) from Finnish language gymnasiums and 133 (41 men) students (out of the 144 who were approached) from Swedish language gymnasiums decided to take part in the research. The mean age of the participants was 17.0 years ($SD = 1.0$). Only those 119 students from Finnish language schools who reported their native tongue as Finnish were retained for further analyses. Regarding Swedish-Finns, we excluded both those 17 participants who reported another language than Swedish as their native tongue, and, because knowledge of the Finnish language was necessary to test some of our predictions, a further eight participants based on mediocre performance in the Finnish language test described below. This left us with 108 bilingual Swedish-Finns who reported Swedish as their mother tongue and performed excellently in our Finnish language skills test.

Participants first completed the S5 in self-rating format in their non-native language (besides Swedish-Finns completing the S5 in Finnish, Finns completed the S5 in Swedish; however, the latter data were not analyzed due to Finns’ poor Swedish language skills—this lack of proficiency was revealed both in the psychometric properties of the measure and a language skills test). Participants then completed, on a separate page, a personal values measure (administered as a filler and for feedback purposes), as well as the S5 in their native language. Finally, they completed the stereotype questionnaire, as well as a language skills test.
Measures

**Personality self-ratings.** Personality was measured with the Short Five (S5; Konstabel, Lönnqvist, Walkowitz, Konstabel, & Verkasalo, 2012), a 60-item measure designed to measure the five factors and 30 facets of the FFM (Costa & McCrae, 1992). Each facet is measured by both a positively and negatively keyed item (adding up to 12 items per personality factor), and each item is responded to on a scale from –3 (*The description is completely wrong*) to 3 (*The description is completely right*). Personality facet and factor scores were obtained by first reverse coding the negatively keyed items, and then computing the mean. The items of the S5 are designed with an eye on comprehensiveness. For instance, the negatively keyed item that measures the warmth facet of extraversion is “I do not like to associate with people much; I am considered a rather cold and distant person.” Means and standard deviations of self-ratings on the S5 facet and factor scales are presented in Table 1.

The Finnish language S5 has been empirically validated as a measure of the personality space depicted by the Five-Factor-Model (Konstabel et al., 2012). We translated the S5 into Swedish for purposes of the present research. The first translation from Finnish into Swedish was done by the first author (bilingual). This translation was scrutinized by the third author (also bi-lingual), and the translation was modified after discussion. To evaluate the correspondence of the Swedish translation of the S5 with the Finnish version, we first looked at the correlations between the measures. At the factor level, the correlations between the Swedish and Finnish versions were all between .86 and .90. Regarding facet level correlations, these were all between .60 (A2; for abbreviations, see first column of Table 1) and .88 (E2), with the exception of self-discipline, which showed a correlation of only .38. To examine this surprisingly low correlation, we looked separately at the positively and negatively keyed item. The positively keyed items correlated $r = .62$ across languages, but the negatively keyed items were uncorrelated, $r = -.07$. This implies that the Swedish translation of the negatively keyed item should be revised. The average facet level correlation, excluding self-discipline, was $r = .74$.

We also evaluated the factor structure of our Swedish translation of the S5. After Procrustes rotation of a five factor solution targeted at the normative North American structure of the NEO-PI-R (Costa & McCrae, 1992), the factor congruence coefficients between the North American factors and the factors obtained from the current data were .93 (N), .85 (E), .91 (O), .89 (A), and .90 (C). For comparison purposes, the extensively validated Finnish S5 showed, in this particular sample, congruence coefficients of 0.93, 0.88, 0.92, 0.88, and 0.93, respectively. Comparison of these congruence coefficients suggests that the Swedish translation of the S5 mimics the normative structure of the NEO-PI-R as well as does the Finnish S5. The high correlations between the Finnish and Swedish language versions, together with the results of the factor analysis, suggest that the Swedish translation of the S5 adequately measures the same constructs as does the Finnish S5.

**Stereotypes.** Stereotypes were measured using the 30 positively keyed items of the S5. We decided to use such a shortened version of the S5 (a) in order to reduce participant fatigue, (b) because the 30 positively keyed items do almost as well as the full set of 60 items in terms of criterion validity (see Figure 1 in Konstabel et al., 2012), and (c) using only one item per trait to be measured should be enough when averaging responses over several respondents (Denissen, Geenen, Selfhout, & Van Aken, 2008). The instruction for the stereotype measure was: “Evaluate
TABLE 1
Descriptive Statistics for the S5 Self-Ratings in Groups of Finns and Swedish-Finns, and *t*-test for the Differences in Means Between the Two Groups When Responding in Finnish

<table>
<thead>
<tr>
<th>Personality facets</th>
<th>M (Finns)</th>
<th>M (Swedish-Finns)</th>
<th>SD (Finns)</th>
<th>SD (Swedish-Finns)</th>
<th>t(df = 221–224)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response language</strong></td>
<td>Fin</td>
<td>Fin</td>
<td>Swe</td>
<td>Fin</td>
<td>Fin</td>
</tr>
<tr>
<td>Anxiety (N1)</td>
<td>−0.64</td>
<td>−0.68</td>
<td>−0.82</td>
<td>2.88</td>
<td>2.51</td>
</tr>
<tr>
<td>Angry hostility (N2)</td>
<td>−0.17</td>
<td>−0.19</td>
<td>−0.46</td>
<td>3.07</td>
<td>2.93</td>
</tr>
<tr>
<td>Depression (N3)</td>
<td>−1.52</td>
<td>−1.51</td>
<td>−1.60</td>
<td>3.07</td>
<td>3.00</td>
</tr>
<tr>
<td>Self-consciousness (N4)</td>
<td>−1.36</td>
<td>−1.16</td>
<td>−1.27</td>
<td>2.94</td>
<td>2.79</td>
</tr>
<tr>
<td>Impulsiveness (N5)</td>
<td>0.68</td>
<td>0.06</td>
<td>0.10</td>
<td>2.69</td>
<td>2.68</td>
</tr>
<tr>
<td>Vulnerability (N6)</td>
<td>−1.07</td>
<td>−1.50</td>
<td>−1.40</td>
<td>2.52</td>
<td>2.35</td>
</tr>
<tr>
<td>Warmth (E1)</td>
<td>3.01</td>
<td>3.27</td>
<td>2.85</td>
<td>2.48</td>
<td>2.33</td>
</tr>
<tr>
<td>Gregariousness (E2)</td>
<td>1.31</td>
<td>1.21</td>
<td>1.10</td>
<td>2.93</td>
<td>2.84</td>
</tr>
<tr>
<td>Assertiveness (E3)</td>
<td>1.32</td>
<td>1.49</td>
<td>1.23</td>
<td>2.80</td>
<td>2.60</td>
</tr>
<tr>
<td>Activity (E4)</td>
<td>0.12</td>
<td>0.43</td>
<td>0.44</td>
<td>2.46</td>
<td>2.51</td>
</tr>
<tr>
<td>Excitement-seeking (E5)</td>
<td>1.76</td>
<td>1.50</td>
<td>1.66</td>
<td>2.59</td>
<td>2.70</td>
</tr>
<tr>
<td>Positive emotions (E6)</td>
<td>2.39</td>
<td>2.37</td>
<td>2.68</td>
<td>2.73</td>
<td>2.47</td>
</tr>
<tr>
<td>Fantasy (O1)</td>
<td>2.75</td>
<td>2.97</td>
<td>2.56</td>
<td>2.58</td>
<td>2.44</td>
</tr>
<tr>
<td>Aesthetics (O2)</td>
<td>1.74</td>
<td>2.41</td>
<td>1.87</td>
<td>3.00</td>
<td>3.13</td>
</tr>
<tr>
<td>Feelings (O3)</td>
<td>3.26</td>
<td>3.26</td>
<td>3.07</td>
<td>2.51</td>
<td>2.23</td>
</tr>
<tr>
<td>Actions (O4)</td>
<td>1.42</td>
<td>0.87</td>
<td>1.13</td>
<td>2.21</td>
<td>2.66</td>
</tr>
<tr>
<td>Ideas (O5)</td>
<td>0.95</td>
<td>1.10</td>
<td>1.42</td>
<td>2.28</td>
<td>2.69</td>
</tr>
<tr>
<td>Values (O6)</td>
<td>2.54</td>
<td>1.99</td>
<td>2.59</td>
<td>2.54</td>
<td>2.52</td>
</tr>
<tr>
<td>Trust (A1)</td>
<td>1.92</td>
<td>1.77</td>
<td>1.65</td>
<td>2.51</td>
<td>2.25</td>
</tr>
<tr>
<td>Straight-forwardness (A2)</td>
<td>1.58</td>
<td>1.98</td>
<td>1.73</td>
<td>2.53</td>
<td>2.41</td>
</tr>
<tr>
<td>Altruism (A3)</td>
<td>2.13</td>
<td>2.80</td>
<td>2.30</td>
<td>2.18</td>
<td>2.29</td>
</tr>
<tr>
<td>Compliance (A4)</td>
<td>0.91</td>
<td>1.21</td>
<td>1.15</td>
<td>2.93</td>
<td>2.93</td>
</tr>
<tr>
<td>Modesty (A5)</td>
<td>0.61</td>
<td>0.75</td>
<td>0.51</td>
<td>2.43</td>
<td>2.65</td>
</tr>
<tr>
<td>Tender-mindedness (A6)</td>
<td>2.52</td>
<td>2.28</td>
<td>2.13</td>
<td>2.40</td>
<td>2.23</td>
</tr>
<tr>
<td>Competence (C1)</td>
<td>1.26</td>
<td>2.26</td>
<td>1.80</td>
<td>2.27</td>
<td>2.07</td>
</tr>
<tr>
<td>Order (C2)</td>
<td>0.52</td>
<td>1.27</td>
<td>1.20</td>
<td>3.09</td>
<td>3.16</td>
</tr>
<tr>
<td>Dutifulness (C3)</td>
<td>2.36</td>
<td>2.83</td>
<td>2.29</td>
<td>2.11</td>
<td>2.03</td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Striving (C4)</td>
<td>1.46</td>
<td>1.64</td>
<td>1.11</td>
<td>2.23</td>
<td>2.68</td>
</tr>
<tr>
<td>Self-discipline (C5)</td>
<td>0.14</td>
<td>0.85</td>
<td>0.19</td>
<td>2.41</td>
<td>2.76</td>
</tr>
<tr>
<td>Deliberation (C6)</td>
<td>0.64</td>
<td>1.36</td>
<td>0.86</td>
<td>2.61</td>
<td>2.60</td>
</tr>
<tr>
<td>Personality factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism (N)</td>
<td>−0.69</td>
<td>−0.84</td>
<td>−0.89</td>
<td>2.04</td>
<td>1.81</td>
</tr>
<tr>
<td>Extraversion (E)</td>
<td>1.66</td>
<td>1.69</td>
<td>1.66</td>
<td>1.96</td>
<td>1.82</td>
</tr>
<tr>
<td>Openness (O)</td>
<td>2.12</td>
<td>2.09</td>
<td>2.10</td>
<td>1.45</td>
<td>1.58</td>
</tr>
<tr>
<td>Agreeableness (A)</td>
<td>1.60</td>
<td>1.80</td>
<td>1.60</td>
<td>1.65</td>
<td>1.49</td>
</tr>
<tr>
<td>Conscientiousness (C)</td>
<td>1.06</td>
<td>1.69</td>
<td>1.25</td>
<td>1.56</td>
<td>1.66</td>
</tr>
</tbody>
</table>

Note. *p < .05, **p < .001.
to what extent the following items describe a typical Swedish-Finn as compared to a typical Finn. The items were rated from −3 (the description is much more accurate of a typical Finnish speaking Finn) to 3 (the description is much more accurate of a typical Swedish-Finn). The inter-rater reliabilities (assessed using average measures intraclass correlation coefficients) were, in both samples, .96 (both 95% CIs: .93–.98; see below for details on how these statistics were computed). Within both samples, the ratings thus reliably tapped into a shared stereotype. The items assessing warmth (E1) and competence (C1) were “He or she likes people; He or she is friendly and open talking to strangers” and “He or she is sensible and competent; He or she can find practical, quick, and effective solutions to problems,” respectively.

**Language skills.** To assure that our Swedish-speaking participants were fluent also in Finnish, we administered a 10-item language test that was designed for the purposes of the present research. Each item consisted of an infrequent Finnish language word selected by the first author, and four alternative Swedish translations of the word. As expected, almost all Swedish-Finns...
showed good knowledge of the Finnish language, with half of the participants making no mistakes, despite the afterwards revealed ambiguity of one of the items. Those eight participants who made more than two mistakes on the ten items were excluded from further analyses.

Measuring Profile Agreement

Profile agreement indicates the extent to which two profiles are similar across a range of attributes. In the present case, the profiles to be compared were personality profiles. For example, we compared Finns’ stereotype ratings with Swedish Finns’ stereotype ratings, as well as stereotype ratings with self-ratings. In order to quantify profile agreement, we computed, using the double-entry method (Kenny, Kashy, & Cook, 2006), the intraclass correlation coefficient (ICC). This statistic is generally used to assess the accuracy of national character stereotypes (e.g., McCrae et al., 2008; Terracciano et al., 2005; Realo et al., 2007). It assesses profile agreement across the 30 personality facets that constitute the FFM (Griffin & Gonzalez, 1995).

The confidence intervals for the ICCs were estimated by bootstrap resampling (e.g., Efron & Gong, 1983). Bootstrap is a resampling method used to calculate standard errors of a statistic when parametric methods are not available or their assumptions are not met. To find the bootstrap estimate, several random “resamples” are taken from the original sample; sampling is done with replacement; that is, in the resample, some cases will be included two or more times, whereas others will be excluded. In each such resample, the value of the statistic is computed, and the procedure is repeated for several times (in our case, 10,000 times). A simple method to calculate the bootstrap standard error is to compute the standard deviation of the statistic over the bootstrap resamples; the confidence intervals can be derived from the standard error, or be estimated from the quantiles of the resampling distribution (we used the latter method).

In order to address some of our research questions, we needed to know whether two ICCs differed from each other. We compared ICCs using the permutation test. This tests the between-group differences in ICCs for significance. In this case, random permutations of the group indicator are used to derive the null distribution of the test statistic. If the observed difference is very unlikely to occur with random permutations of the group indicator, then one can infer that the observed difference is very unlikely to occur by chance alone.

RESULTS

Convergence of the Stereotypes

We first focused only on stereotype ratings. Finns’ and Swedish-Finns’ stereotypes are presented in Figure 1. Recall that stereotype ratings were collected by asking participants to directly compare Finns with Swedish-Finns, with negative numbers suggesting that the trait is more typical of a Finnish speaking Finn, and positive numbers suggesting that the trait is more typical of a Swedish-Finn. On the x-axis are the stereotype ratings made by Finns, and on the y-axis are the stereotype ratings made by Swedish-Finns. Hence, traits in the top right corner are traits that both Finns and Swedish-Finns agree that are more typical of Swedish-Finns. Opposite to those, in the bottom left corner, are traits that both groups agree are more typical of Finnish speaking Finns. In the top left corner and in the bottom right corner are traits on which there is disagreement.
between the two groups of raters. More specifically, in the top left corner are traits that both groups have rated as more descriptive of their own in-group. By contrast, traits rated as more typical of the other group would have appeared in the bottom right corner, had there been any such traits.

The first thing to notice in Figure 1 is the remarkable level of agreement on which traits are more typical of Finns and which traits are more typical of Swedish-Finns. For 24 out of the 30 facets, Finns and Swedish-Finns agreed on whether the trait was more descriptive of Finns (bottom left quadrant) or Swedish-Finns (top right quadrant). The remaining six facets were such that both groups thought them more descriptive of themselves (top left quadrant). The ICC between Finns’ and Swedish Finns’ stereotype ratings was .77 (95% CI: .63–.83). Such high inter-group agreement is consistent with our prediction regarding the convergence of national character stereotypes.

Content of the Stereotypes

The second thing to notice about Figure 1 is the content of the traits. The 16 traits in the top right quadrant (more typical of Swedish-Finns) are primarily the constituents of extraversion, openness to experience, and agreeableness. These are all highly desirable personality traits (e.g., Paulhus, 2002). However, in the bottom left quadrant (more typical of Finnish speaking Finns), are all six facets of neuroticism, the only one of the five factors to be considered undesirable. The six neuroticism facets are accompanied by assertiveness (E3) and modesty (A5). The former trait suggests some degree of indifference towards others (in the sense of neglecting others’ opinions and viewpoints), and the latter could be interpreted as a sign of lacking confidence or self-esteem. Both groups thus appeared to rate the personality of a typical Swedish-Finn more positively than the personality of a typical Finn. To confirm this interpretation, we compared the stereotype ratings with social desirability ratings collected from an independent batch of gymnasium students. More specifically, we computed the ICCs between Finns’ and Swedish-Finns’ stereotype ratings with the social desirability ratings. The stereotype ratings were positively correlated with the social desirability ratings of the items (the ICC between Finns’ stereotype ratings and social desirability ratings was .17; 95% CI: .07–.27 and the ICC between Swedish-Finns’ stereotype ratings and social desirability ratings was .51; 95% CI: .42–.60). The result that also Finns viewed Swedish-Finns as having more desirable personalities conflicted with our expectation that both groups would rate their in-group more positively. But consistent with the notion of in-group bias, Finns’ stereotype of Swedish-Finns was not as positive as Swedish-Finns’ stereotype (the difference between the two ICCs, .17 and .51, was statistically significant at \( p < .0001 \)).

Returning to Figure 1, examining the stereotypes at the level of the five factors reveals some differences between the groups. Finns’ and Swedish-Finns’ stereotypes were dissimilar regarding conscientiousness—both groups perceived this personality factor to be more typical of their in-group (upper left quadrant). Besides this differences in perceptions of conscientiousness (\( M \) (Swe) = 1.71 vs. \( M \) (Fin) = −1.47, \( t = 4.85, p < .001 \)), Swedish-Finns also rated the typical representative of their group as relatively higher in openness to experience (\( M \) (Swe) = 4.84 vs. \( M \) (Fin) = 3.13, \( t = 2.97, p < .01 \)) and agreeableness (\( M \) (Swe) = 3.55 vs. \( M \) (Fin) = 0.39, \( t = 5.52, p < .001 \)) than did Finns. In fact, regarding agreeableness, Finns’ ratings
did not reveal a statistically significant difference between ratings of Finns and Swedish-Finns (as shown by one sample t-test tested against $M = 0$, indicative of no differences between the two groups; $t = -1.02, p = .31$). In sum, although both groups of raters agreed that the personality of a typical Swedish-Finn was more socially desirable, differences between the two groups of raters revealed, as expected, that divergences from the shared stereotype were in-group favoring.

Based on the hypothesized compensatory nature of out-group stereotypes, we expected Finns to rate Finns and Swedish-Finns as similar on competence, but to rate Swedish-Finns as lower on warmth. Opposite to these expectations, Finns rated themselves as higher on competence (the following t-test results are all from one sample t-tests tested against $M = 0$: $M = -0.53; t = -4.2, p < .0001$), but Swedish-Finns as higher on warmth ($M = 1.39; t = 12.1, p < .0001$). The results regarding Finns’ perceptions of intergroup differences in competence and warmth were thus not consistent with our predictions based on the SCM. We further expected Swedish-Finns to rate Finns and Swedish-Finns as similar on warmth, but to rate Finns as lower on competence. Consistent with expectations, Swedish-Finns did believe that they were typically more competent than Finns ($M = 0.53; t = 5.01, p < .0001$). However, contrary to expectations, Swedish-Finns also saw themselves as higher on warmth ($M = 1.47; t = 11.30, p < .0001$). In sum, both groups of raters believed their own in-group to be higher on competence, and both groups agreed that Swedish-Finns were higher on warmth.

The Kernel of Truth Hypothesis

We next investigated the kernel of truth hypothesis. Following previous research (e.g., Terraciano et al., 2005; Realo et al., 2009), we first compared, within samples, the shape of the stereotype profile with the shape of the aggregate self-ratings. These correlations revealed that the Swedish-Finns’ average self-rated personality profile, when responding in Finnish, resembled the stereotype profile that indicated how they thought themselves to differ from Finns. The ICC between Swedish-Finns’ self-ratings and stereotype ratings was .35 (95% CI: .28–.41). By contrast, Finns’ average self-rated personality profile was not very similar to their stereotype of group differences; the ICC between Finns’ self-ratings and stereotype ratings was .16 (95% CI: .09–.22). Furthermore, comparison of mean levels of self-ratings (presented in Table 1) between Swedes and Finns did not support the kernel of truth hypothesis—on the level of personality factors, the only difference between the two groups was that Swedish-Finns rated themselves as higher on conscientiousness, the only personality factor on which both groups perceived themselves to be higher.

We also investigated the kernel of truth hypothesis in terms of the general desirability of the personality profiles—because both groups perceived Swedish-Finns to have more desirable personalities, the kernel of truth hypothesis would imply that Swedish-Finns also in self-ratings reveal a more desirable personality profile. These analyses were the only ones to yield different conclusions depending on whether we computed ICCs or Pearson’s rs (see Footnote 2). The ICCs suggested that Finns’ and Swedish-Finns’ aggregate self-rating personality profiles (when both groups responded in Finnish) were equally similar to the social desirability profile (the ICC between Finns’ self-ratings and desirability ratings was .74; 95% CI: .67–.80 and the ICC between Swedish-Finns’ self-ratings and desirability ratings was .81; 95% CI: .76–.85). However, the Pearson’s rs suggest, in support of the kernel of truth-hypothesis, that the Swedish-Finns’ self-ratings were somewhat more desirable than the Finns’ self-ratings (the Pearson’s r between Finns’
self-ratings and desirability ratings was 0.86, 95% CI: 0.80–0.89, and the ICC between Swedish-Finns’ self-ratings and desirability ratings was 0.90, CI: 0.86–0.92; the difference between the two correlations was statistically significant at \( p < .05 \).

**Cultural Frame Shifting**

Finally, we investigated the possible effects of CFS. We expected Swedish-Finns’ self-ratings to more strongly resemble their stereotype of Swedish-Finns when they responded in Swedish. Consistent with this expectation, when they responded in Swedish, the resemblance between stereotypes and aggregate self-ratings was stronger than when they responded in Finnish (the ICC between Swedish-Finns’ self-ratings in Swedish and stereotype ratings was .81 (95% CI: .71–.85); the ICC between Swedish-Finns’ self-ratings in Finnish and stereotype ratings was .75 (95% CI: .66–.79). However, this difference did not reach statistical significance (\( p = .08 \), using a permutation test with 10000 permutations; because we had a hypothesis based on CFS about the direction of the effect, we used a one-sided test).

**DISCUSSION**

The results supported the view that in-group and out-group national stereotypes tend to converge. The shared stereotype of Swedish-Finns depicted them as having more favorable personality characteristics in comparison to Finns. However, there was also some evidence of in-group bias, as Swedish-Finns’ stereotype of Swedish-Finns was more socially desirable than Finns’ corresponding stereotype. On the level of the five factors, in-group bias was evidenced by both groups rating themselves as higher on conscientiousness, and by Swedish-Finns perceiving in-group favoring differences in openness to experience and agreeableness as larger (in fact, regarding the latter, Finns perceived the two groups as similar). However, we found no evidence for the compensatory nature of out-group stereotypes—both groups agreed that Swedish-Finns were higher on warmth, and both groups perceived themselves to be higher on competence. These results suggest a possible boundary condition to the SCM—ratings of national character stereotypes may not be determined by the competitiveness of intergroup relations (thought to be pertinent to judgments of warmth) nor by intergroup status differences (thought to be pertinent to judgments of competence).

The evidence regarding the kernel of truth hypothesis was mixed. Finns’ self-ratings were not similar to their stereotype of group differences, and actual personality differences between Finns and Swedish-Finns were not consistent with the stereotypical group differences. However, congruent with the kernel of truth hypothesis, Swedish-Finns’ self-rated personality profiles were, consistent with their stereotype of group differences, more desirable than Finns’ self-rated personality profiles, although this result was only statistically significant when employing Pearson’s \( r_s \) instead of ICCs. Furthermore, similarity between Swedish-Finns’ aggregate self-rated personality profile and their stereotype of group differences supported the kernel of truth hypothesis. Finally, although the result failed to reach conventional levels of statistical significance, the similarity between these two profiles may have grown stronger when Swedish-Finns responded in Swedish suggesting that the accuracy of stereotypes vis-à-vis self-ratings may in part depend
on the circumstances under which the self-ratings are gathered—specifically, on which cultural mindset is activated.

Sources of National Character Stereotypes

There is an abundance of social psychological literature showing that people generally tend to evaluate in-groups more positively than out-groups and attribute more desirable traits to the in-group than the out-group (Hewstone et al., 2002). Consistent with this, Finns did rate the typical Finn more positively than did Swedish-Finns. However, in contrast to expectations, both Finns and Swedish-Finns attributed more socially desirable personality characteristics to Swedish-Finns.

The out-group favoritism shown by Finns, a generally uncommon phenomenon, could be explained by the stereotype of Swedish-Finns containing a kernel of truth. Supporting the idea that the stereotype was somewhat accurate, Swedish-Finns’ average self-rated personality profile resembled the stereotype profile that indicated how Swedish-Finns thought themselves to differ from Finns. One interesting aspect of the results is that the shared stereotype regarding personality differences between Finns and Swedish-Finns corresponded to some extent with actual behavioral and demographic differences between these two groups. The groups agreed that Swedish-Finns are higher in extraversion and openness to experience, and lower in neuroticism. In a recent meta-analysis on the effects of personality on three important life outcomes (Roberts et al., 2007), high extraversion and low neuroticism were both related to lower mortality and higher occupational attainment, whereas low neuroticism was additionally related to lower divorce rates. As noted in the Introduction, Swedish-Finns actually do enjoy lower mortality, better work-related outcomes, and lower divorce rates (Finnäs, 1997; Hyypää & Mäki, 2003; Saarela & Finnäs, 2003). Regarding the shared stereotype of Swedish-Finns’ higher openness to experience, this is also perfectly consistent with research showing that Swedish-Finns are more politically liberal (McRae, 1999)—an ideology predicted by personality traits such as open-mindedness, creativity, curiosity, and novelty-seeking (Carney, Jost, Gosling, & Potter, 2008).

In light of such demographic and behavioral differences between Finns and Swedish-Finns, one explanation for the shared stereotype could be that the stereotype, in fact, accurately reflects personality differences. Our results are thus similar to those reported on by Heine, Buchtel, & Norenzayan (2008), who presented evidence that perceptions of national character are quite valid as measures of between-culture differences in behavior. More specifically, Heine and colleagues (2008) showed that perceptions of national character, but not self-report personality questionnaires, predicted between-culture behavioral and demographic indices: auto-stereotypes of conscientiousness predicted, across 17 to 55 nations, walking speed, postal workers’ speed, accuracy of clocks in public banks, per capita GDP, and longevity.

Perceptions of differences between Swedish-Finns and Finns are likely to at least in part have originated from the above described behavioral and demographic differences between the two groups. In essence, especially Swedish-Finns but also Finns may be vaguely aware, based on, for instance, media coverage of research, or populist nationalistic outbursts towards the allegedly favored minority, that Swedish-Finns tend to do better than Finns regarding some life outcomes. People typically overemphasize the role of individuals’ personality traits, and underemphasize the role of varying social environments when interpreting and explaining others’ behavior.
and Nisbett, 1991), suggesting that the above differences in behavioral and demographic variables may be seen as signs of differences in underlying personality dispositions—i.e., knowledge that Swedish-Finns tend to be more positive life outcomes might have influenced stereotype ratings—such knowledge may have led participants to rate the personality of Swedish-Finns more positively.

One reason that national character stereotypes have in some (e.g., Realo et al., 2009), but not other (e.g., Church & Katigbak, 2002) studies been found to be accurate vis-à-vis self-ratings of personality could be that self-ratings of personality are malleable and systematically influenced by cultural mindsets. Our results, which suggest that the Swedish language constituted a cultural cue that activated a specific cultural mindset, support this idea. When responding in Swedish, a certain stereotype consistent self-concept or set of cultural goals or values could have been made salient, and this could have influenced self-ratings of personality. In essence, there would be no one true personality, but one’s standing on personality dimensions would, to some extent, depend on the moment-to-moment salience of different self-concepts (for such an interpretation of the dimensions of individualism and collectivism, see Oyserman and Lee, 2008). Consistent with the idea that one’s standing even on such fundamental characteristics as personality dimensions can depend on the salience of different self-concepts, some recent research shows that people report higher levels of Agreeableness when considering how they act when they are around strangers, as compared to when considering how they act around family and friends (White et al., 2012). Similarly, cultural cues could increase the salience of a specific self-concept and lead to responses more congruent with that self-concept.

Limitations and Conclusions

In a literature dominated by large-scale multi-culture studies (Terraciano et al., 2005; Realo et al., 2009), one can ask what small-scale studies such as the present one can contribute. One advantage of smaller studies is that they allow for the use of experts; i.e., people who are actually familiar with the cultures being rated. The present study had the further advantage that participants most likely shared the same cultural norms, eliminating or at least attenuating any reference-group effects that could otherwise have distorted the results or complicated their interpretation (although such effects may anyway be weak, see Mõttus et al., 2012). Small-scale studies also allow focusing on particularly interesting groups, such as high-status minorities. Results stemming from research on such atypical groups may be particularly informative in illuminating the boundary conditions and limitations of purportedly general models. In particular, the present results show that low status groups, even when constituting the vast majority, may show out-group favoritism in ratings of personality characteristics.

Our results suggest possible boundary conditions for the SCM of social perception. If national character stereotypes, as we argue, at least in some instances reflect real differences between groups of people, then the SCM may not be pertinent in explaining or predicting such stereotypes. Investigating the boundary conditions of the SCM could be a fruitful avenue for future research.

One important contribution of the present research comes from its use of bilingual participants. This allowed us to investigate whether CFS could help explain the mixed evidence regarding the correspondence between national character stereotypes and personality self-ratings. Although the results were in the direction predicted by CFS, they were only marginally statistically significant. However, this is not sufficient grounds for more generally rejecting CFS; our results could be...
interpreted as merely indicating that the Swedish-Finnish culture is rather similar to the Finnish culture. Responses in another language, such as English, could have produced different results. Whether self-ratings of personality can vary as a function of cultural mindset should be the focus of future research in cross-cultural psychology. Such a result, if robust, could have far-reaching implications for research domains in cross-cultural psychology dealing with the personality profiles of cultures. This type of research has stirred a lot of controversy, for instance due to problems of scalar equivalence across languages—comparing mean score differences between groups that have completed the measure in different languages requires the difficult and seldom performed task of establishing scalar invariance or equivalence across languages (Chen, 2008). One way to circumvent this problem is the increasingly common practice of administering the tests in English across different language groups (Burmester, Kwang, & Gosling, 2011). However, the present results do not fully support such a practice, as they tentatively suggest that test scores even on such fundamental characteristics as personality traits may be influenced by the cultural mindsets associated with the language the test is administered in.

The results of the present study supported the idea that group membership influences perceptions of national character—Finns perceived the typical Finn somewhat more favorably than did Swedish-Finns. However, the results even more strongly supported the idea that auto- and hetero-stereotypes tend to be highly similar—despite some differences, both Finns and Swedish-Finns ascribed more favorable personality characteristics to Swedish-Finns. We also found mixed support for the accuracy of national character stereotypes when compared to aggregate personality profiles: Swedish-Finns’ self-ratings were more desirable than Finns’ self-ratings, and resembled their stereotype of Swedish-Finns. Related to this, Swedish-Finns’ personality self-ratings were (although only marginally statistically significantly) more consistent with their stereotype of the typical Swedish-Finn when these self-ratings were collected in Swedish rather than Finnish, hinting at the influence of cultural mindset on personality. Future research in cross-cultural personality psychology should further investigate the influence of cultural mindsets.

NOTES

1. The instructions and rating scale, in which participants were directly asked to compare Finns and Swedish-Finns, were essential for being able to investigate the expected compensation effect (Yzerbyt et al., 2005). The expected negative correlation between warmth and competence judgments could appear to conflict with the firmly established halo-effect (Rosenberg, Nelson, & Vivekananthan, 1968; Thorndike, 1920). According to the halo-effect a general good vs. bad feeling will color the judgment of the separate qualities of the target. Resolving this conflict, Judd et al. (2005) showed that when participants judged a single target, either described as high or low on warmth or competence, a clear halo effect emerged: those described as high on one dimension were judged to be high also on the other dimension. The compensatory effect between warmth and competence judgments only appeared when participants were instructed to compare two targets with each other—this method was therefore employed in the present research.

2. ICCs are quite similar to Pearson’s correlation coefficients, but there is one important difference. Whereas Pearson’s rs are only sensitive to the shapes of the profiles, ICCs are sensitive to both the shapes of the profiles and to differences in elevation. This has been argued to make ICCs more appropriate for calculating profile similarity. Empirically verifying this assertion, McCrae (2008) compared different measures of profile agreement and found double-entry ICC to be the best. However, Furr (2010) could not replicate these findings, and argued that the Pearson’s rs could instead be used—being responsive only to the shape of the profile, the Pearson’s r is clearer to interpret. We therefore
ran all analyses using both ICCs and Pearson’s rs. In general, the correlations between profiles tended to be somewhat stronger when using the Pearson’s rs, suggesting the occurrence of mean-level differences between the profiles that were captured only by the ICCs. However, as none of our conclusions changed regardless of which method we used, except in one set of analyses for which we will report both ICCs and Pearson’s rs, we report only the ICCs. We prefer the ICCs because they should improve comparability with previous research on national stereotypes (e.g., Terraciano et al., 2005; Terracciano & McCrae, 2007, 2008; Realo et al., 2009). Regarding the interpretation of ICCs, similarly to Pearson’s rs, the square of the ICC is an estimate of the amount of variance of one variable that another variable can account for.

3. The desirability ratings were gathered in 2006–2007 from three Finnish language gymnasiums (located in Raasebro, Lumo and Helsinki) different to those from which the stereotype ratings were gathered. The procedure was similar to the one used to gather stereotype ratings: the questionnaires were administered on-line in various classes, participation was voluntary, and as an incentive, participants were given feedback on their personal values profiles. Altogether 281 (95 men, gender data missing for one class consisting of 26 students) students out of the around 290 who were approached decided to take part in the research. The mean age of the participants was 16.9 years (SD = 0.9). For desirability ratings, each of the sixty S5 items was rated with the following instructions: “Descriptions of people are often evaluative. Some characteristics are judged as socially desirable, and others as undesirable. Indicate for each item whether the characteristic described in the item is socially desirable, undesirable, or neutral”. The ratings were made on scale from –3 (wholly undesirable) to 3 (wholly desirable).

**AUTHOR NOTES**

Jan-Erik Lönnqvist is affiliated with the Swedish School of Social Science, University of Helsinki. Kenn Konstabel is affiliated with the Department of Psychology, University of Tartu. Nellystiina Lönnqvist is affiliated with the Institute of Behavioural Sciences, University of Helsinki. Markku Verkasalo is affiliated with the Institute of Behavioural Sciences, University of Helsinki.

**REFERENCES**


Received January 22, 2013
Accepted September 6, 2013