RESEARCH NOTE

Measurement Equivalence Across Subnational Groups: An Analysis of the Conception of Nationhood in Switzerland

Oriane Sarrasin[1], Eva G. T. Green[1], André Berchtold[1] and Eldad Davidov[2]

[1]Institute of Social Sciences, University of Lausanne, Switzerland;
[2]Institute of Sociology, University of Zurich, Switzerland

Because of their higher status, members of national majorities (e.g., linguistic, ethnic, or religious) are likely to develop a stronger sense of ownership of the nation than members of national minorities (Sidanius, Feshbach, Levin, & Pratto, 1997), which is reflected in a stronger attachment to the nation as a whole (e.g., Staerklé, Sidanius, Green, & Molina, 2010). Accordingly, members of national majorities are expected to hold a stricter conception of nationhood, that is, to be stricter regarding who should or should not belong to the national community (Kunovich, 2009). Based on data from two large surveys (European Social Survey [ESS] 2002 and International Social Survey Programme [ISSP] 2003), the present study tests this expectation by comparing conceptions of nationhood across the two largest linguistic subnational groups in Switzerland.

However, observed differences between the groups can either reflect substantive differences or result from measurement artifacts, as distinctive features of the subnational groups (e.g., different languages; Davidov & De Beuckelaer, 2010) may impact the measurement itself. Therefore, measurement equivalence should be tested for before undertaking substantive mean comparisons (e.g., Steenkamp & Baumgartner, 1998). Although measurement equivalence testing is frequently carried out in cross-national studies (e.g., Davidov, 2009; Reeskens & Hooghe, 2010), it is rarely established when comparing data drawn from distinct national subgroups *within* countries (see, however, Billiet, 2003; Billiet, Maddens, & Beerten, 2003). Thus, this study has both a substantive and a methodological rationale: Beyond testing the assumption that members of national majorities hold a stricter conception of nationhood than

All correspondence concerning this article should be addressed to Oriane Sarrasin, Institute of Social Sciences, University of Lausanne, Vidy Building, CH–1015 Lausanne, Switzerland. E-mail: oriane.sarrasin@unil.ch

members of national minorities, it aims at illustrating the necessary steps of analysis—using exploratory and multigroup confirmatory factor analyses—for conducting valid latent mean comparisons across subnational groups.

Conceptions of Nationhood

In the present study, *conception of nationhood* refers to the different criteria evoked to establish and maintain the boundaries between those who belong to the national community and those who do not. The boundaries can be both tangible (e.g., being denied entry into the country) and intangible (being dismissed as a true member of the nation). Criteria to define such boundaries have been traditionally divided into two categories (Kohn, 1944). On the one hand, *ethnic* criteria entail an ascriptive representation of the members of the nation sharing a common ancestry and place of birth (Smith, 1991). Ethnic criteria such as having national ancestry or being a member of the main religious group of the nation are hard or impossible to obtain. On the other hand, *civic* criteria entail a voluntarist representation of the nation (Jones & Smith, 2001), where being a member of the nation requires adherence to political and cultural aspects of the nation (e.g., respecting institutions and speaking the language). Because they can be acquired, civic criteria might appear to be less strict than ethnic criteria. However, both criteria can be used and mobilized to exclude individuals or groups (e.g., immigrants) from the national ingroup (Green, 2009).

Conception of Nationhood Across Subnational Groups: The Example of Switzerland

Conceptions of nationhood at the country level (Smith, 1991) are reflected in national policies (e.g., naturalization; Brubaker, 1992) that can be described as predominantly ethnic or civic. However, no state is purely either ethnic or civic (Kuzio, 2002). Moreover, the degree of endorsement of both ethnic and civic criteria varies across individuals and national subgroups (Kunovich, 2009). The present study tests the prediction that members of national majorities, with higher status because of numerical advantage or better access to resources, hold a stricter conception of nationhood than members of national minorities. Because they are motivated to maintain restrictive boundaries of the national ingroup to safeguard their dominant position (Sidanius et al., 1997), members of national majorities are expected to express a stronger adhesion to both ethnic and civic criteria. With clearly distinct subnational groups, Switzerland is a highly relevant context for testing this prediction. Much like in other countries with a national majority and minorities (Staerklé et al., 2010), higher levels of blind and uncritical attachment to the nation (e.g., feeling of superiority regarding other nations) have been found among the German-speaking majority compared with the French-speaking minority (Green, Sarrasin, Fasel, & Staerklé, 2011).[1] Unsurprisingly, the willingness to exclude immigrants (Kriesi, Wernli,

[1]In 2000, 72.5% of the Swiss citizens had German and 21% French as first language (Swiss Federal Statistical Office, 2011). The two other official linguistic groups (i.e., Italian, 4.3%, and Rumantsch, 0.6%) could not be included because the surveys used in the present study do not contain a sufficient number of respondents to perform valid comparisons.

Sciarini, & Gianni, 1996) is also stronger in the German-speaking region. Accordingly, higher commitment to both ethnic and civic conceptions of nationhood is likely to be expressed by members of the German-speaking majority than by members of the French-speaking minority.

The prediction will be tested with data from two recent large international surveys (ESS 2002 and ISSP 2003) that included sets of items measuring conceptions of nationhood (Table 1). Although the cross-national comparability of these scales has received considerable methodological attention (Heath, Martin, & Spreckelsen, 2009; Medina, Smith, & Long, 2009; Reeskens & Hooghe, 2010), no study, to our knowledge, has so far examined whether they can be validly compared or pooled in countries with distinct subnational groups. Although some common causes of non-equivalence are less likely to be problematic within a country (e.g., data collection and sampling methods; Heath et al., 2009), distinctive features of the subnational groups, in particular different languages (Davidov & De Beuckelaer, 2010), may jeopardize measurement equivalence.

Measurement Equivalence Testing

Various methodological and substantive differences between a set of samples under study may lead to artificially significant or nonsignificant mean differences (Steenkamp & Baumgartner, 1998). Differences can occur in terms of the organization of the survey (e.g., responses rates, data collection method, or sampling procedures; Heath et al., 2009), but they can also be because of either the survey items themselves (i.e., some items do not have the same meaning across the groups) or to response styles (Cheung & Rensvold, 2000). In such cases, establishing measurement equivalence is a necessary step before comparing or pooling data from distinct groups.

Measurement equivalence is generally established by a hierarchy of tests ordered by level of strictness, with the least restrictive level being called *configural* equivalence. The only requirement is a similar pattern of salient and nonsalient factor loadings across the groups under comparison (i.e., the same items load, as well as do not load, on the same factors in each group). Although multigroup confirmatory factor analyses (MGCFA) are used to test whether configural (and higher forms of) equivalence is reached, it is advised (by Byrne, 2010) to study the factor structure within each group before proceeding with MGCFA to detect whether the pattern of salient and nonsalient factor loadings is similar across the groups. For instance, Heath et al. (2009) performed exploratory factor analyses (EFAs) on the ISSP conception of nationhood scale for each country separately. Results showed that the "having the national citizenship" item hampered configural equivalence (Reeskens & Hooghe, 2010), either by loading on the ethnic factor (i.e., salient loading on the ethnic factor and nonsalient loading on the civic factor) or cross-loading (i.e., salient loading on both factors). The authors explained such variation by the conditions for obtaining citizenship, which differ to a great extent across countries (Brubaker, 1992). It is recommended to discard items that hamper configural equivalence (Reeskens & Hooghe, 2010), as they likely result in model rejection when using MGCFA. In the current study, we examine whether items tapping civic and ethnic conceptions of nationhood,

respectively, load on the same factor in the German- and the French-speaking regions of Switzerland.

The next steps of measurement invariance are directly testable through MGCFA.[2] The second and more restrictive level, *metric* equivalence, requires item loadings to be equivalent across groups. Thus, if metric equivalence is established, items displaying strong loadings in one group should display similar loadings in the other group(s) under comparison (Selig, Card, & Little, 2008). Partial (rather than full) metric equivalence will be established if at least two items per latent variable still display equal factor loadings across groups (Byrne, Shavelson, & Muthén, 1989). As full or partial metric equivalence relies on item covariation only, comparison of relations between latent variables is possible at this level (Van de Vijver & Leung, 1997). We examine the extent to which civic and ethnic item loadings are similar in the German- and French-speaking regions in Switzerland. However, comparative studies frequently elaborate hypotheses on mean differences—in the present case, a stricter conception of nationhood among national majorities than among national minorities—which require not only equality of item loadings across groups but also equality of item intercepts.

Thus, the comparison of latent variable means requires a more restrictive level of equivalence, named *scalar* equivalence, where, in addition, the intercepts of the different items measuring the construct(s) are constrained to be equivalent. If scalar equivalence is reached, "the pattern of high and low scores on a set of indicators is the same" (Selig et al., 2008, p. 97). Because of their strictness, scalar equivalence models are hard to reach (Steenkamp & Baumgartner, 1998), and researchers tend to frequently rely on *partially* scalar equivalent scores. Although some studies have shown that a subset (i.e., two) of equivalent items per concept was sufficient for achieving meaningful latent score comparisons (Baumgartner & Steenkamp, 1998; Byrne et al., 1989), other scholars maintain (De Beuckelaer, 2005; Van de Vijver & Leung, 1997) that latent mean comparisons are not fully reliable unless *full* scalar equivalence is reached. As solving this debate is beyond the scope of the present study, we adopt the following pragmatic procedure: Before drawing conclusions regarding the higher restrictiveness of the national majority compared with the minority, latent means obtained from partially scalar equivalent models (where at least two items per civic and ethnic construct of nationhood have equal intercepts across groups) will be carefully compared with latent means obtained from fully scalar equivalent models (where *all* items measuring each construct have equal intercepts across groups).

Method

In both the ESS 2002 and ISSP 2003,[3] we selected data from Swiss citizens living in the German- and French-speaking regions. Items are presented in Table 1.

---

[2]Although diverse methods allow the examination of nonequivalence (Hui & Triandis, 1985; Medina et al., 2009), MGCFA is often considered as most appropriate because "a number of specific aspects to the measurement equivalence issue are readily testable within a CFA framework" (Vandenberg & Lance, 2000, p. 6).

[3]Detailed information and data can be freely accessed at http://ess.nsd.uib.no/ess/round1/ (ESS) and http://www.gesis.org/issp/issp-modules-profiles/national-identity/2003/ (ISSP).

European Social Survey 2002

In the ESS (ESS, 2002), respondents had to rate how important (from 0 = *extremely unimportant* to 10 = *extremely important*) eight characteristics *should be in deciding whether someone born, brought up, and living outside Switzerland should be able to come and live here*. Six characteristics from the original scale were considered for the present study (two items—"being wealthy" and "having family in the host country"—were excluded, as they do not correspond to the ethnic–civic conceptualization; Green, 2009). Our final ESS sample contained 1,748 Swiss citizens (912 women and 836 men), with 1,350 respondents (77.2%) living in the German-speaking part and 398 respondents (22.8%) living in the French-speaking part. Respondents' ages ranged from 14 to 102 years ($M = 48.13$, $SD = 17.51$).

International Social Survey Programme 2003

In the ISSP (ISSP, 2003), respondents had to rate how important eight characteristics were *for being truly Swiss*. The ISSP coding (from 1 = *very important* to 4 = *not at all important*) was reversed so that high scores indicated that a characteristic was important. The final sample contained 902 Swiss citizens (463 women and 439 men), with 745 (82.6%) living in the German-speaking part and 157 living in the French-speaking part (17.4%). Respondents' ages ranged from 18 to 96 years ($M = 50.48$, $SD = 17.28$).

Results

Description of Analyses

In a preliminary step, EFAs (principal component factor analyses with a Promax rotation) were performed with Stata 10.0 to evaluate whether similar factor structures were present in the two linguistic regions. Based on suggestions by Van De Vijver and Leung (1997), items were standardized within each region before analysis. EFA has the power to reveal the number and content of the dimensions in each region. However, they do not convey the similarity (or congruence) of the factor solution *between* regions, as the spatial orientation of factors is arbitrary. This was examined in a second step with Procrustean analysis (or targeted rotations; Barrett, 1986). The degree of congruence, that is, the similarity in the loading patterns after a rotation to maximize agreement, was estimated with the Procrustes statistic provided by Stata (ranging from 0 to 1, with 0 indicating a perfect congruence; StataCorp, 2007). Items that loaded on both factors in the EFA (Table 1) were discarded, as they are likely to hamper configural equivalence.

Next, measurement equivalence across German- and French-speaking Swiss samples was tested by means of MGCFA with the Mplus 5.1 software (Muthén & Muthén, 2008). Items were assigned to factors based on our initial allocation of items to factors and consideration of the EFA results. After confirming configural equivalence with MGCFA models, item loadings were constrained to be equivalent, thus testing for metric equivalence. Where models for metric equivalence adequately fitted the data, scalar equivalence was tested in a further step; in this case, item

intercepts were also constrained to be equal across linguistic groups. Factor latent means were compared only in case of full or partial scalar equivalence. All statistical analyses used in the present study are designed for continuous data; however, as responses in ISSP are recorded on 4-point scales, MGCFA for ordinal data were also performed (for an application with Mplus, see Davidov, Datler, Schmidt, & Schwartz, 2011). Results were similar to those obtained with MGCFA for continuous data.

Though no absolute rule exists (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004), confirmatory models are usually considered as fitting the data adequately when both the comparative fit index (CFI) and Tucker–Lewis index (TLI) are higher than .95 and both the standardized root mean square residual (SRMR) and root mean square error of approximation (RMSEA) are lower than .05 (though values between .05 and .08 are considered acceptable by some scholars; for a review, see Schermelleh-Engel, Moosbrugger, & Müller (2003)). These indices reveal whether a certain model fits the data; they are however less informative regarding the change between two steps in equivalence measurement testing. For this, examination of fit indices (Chen, 2007; Cheung & Rensvold, 2002) showed that small changes (between the steps) in CFI and RMSEA reliably indicate that a further step in equivalence is reached. Based on Chen (2007), a stricter level of equivalence was considered achieved when both the decrease in CFI was equal to or smaller than .010 and the increase in RMSEA was equal to or smaller than .015 (Byrne & Stewart, 2006). If fit indices fell outside the recommended values, equality constraints on some parameters were relaxed based on the modification indices (MIs) indicating the possible reductions in χ^2 when relaxing equality constraints.

Measurement Equivalence Testing

European Social Survey. The initial structure was chosen as a function of EFA results (Table 1): Being White and being Christian were assigned to a first *ethnic* dimension, and commitment to the way of life, good educational qualifications, speaking the language, and work skills were assigned to a second *civic* dimension. Though commitment to the way of life loaded in the EFA to a similar extent on both dimensions in the German-speaking region, we placed it on the civic dimension in the confirmatory analyses because this item is usually considered as a characteristic that can be acquired (Green, 2009). The Procrustean analysis yielded strong congruence (Procrustes statistics = .043), indicating that the factor solution was similar across the two regions.

The two first models that tested for configural and metric equivalence resulted in good fit indices (Table 2). Next, the model testing for full scalar equivalence failed to reach an acceptable fit as the changes in CFI (.044) and RMSEA (.027) fell beyond the recommended criteria (Chen, 2007). The MI showed that the χ^2 value could be considerably reduced ($\Delta\chi^2 = 96.73$) if the equality constraint for the intercept of the commitment to the way of life item would be relaxed. As a result, partial scalar equivalence was supported by the data: Fit indices were acceptable, and changes in indices (ΔCFI = .002, ΔRMSEA = .003) were within the range recommended by Chen (2007). In a last step, we compared the latent means of ethnic and civic

Table 1

Factor Loadings for Exploratory Factor Analysis with Promax Rotation by Linguistic Region (ESS 2002 Immigration Criteria; ISSP 2003 Criteria for Being a "True" Swiss)

Survey and criteria	German		French	
	Ethnic	Civic	Ethnic	Civic
ESS 2002				
Being white	**.90**	−.15	**.90**	
Being Christian	**.80**		**.81**	.10
Be committed to the way of life in Switzerland	.31	**.37**	−.17	**.72**
Having good educational qualifications	−.15	**.90**		**.83**
Being able to speak German, French, or Italian		**.82**		**.72**
Having work skills that Switzerland needs		**.80**	.12	**.75**
ISSP 2003				
Having Swiss ancestry	**.85**	−.11	**.86**	−.16
Being born in Switzerland	**.82**		**.85**	−.10
Having lived in Switzerland for most of one's life	**.72**		**.59**	.18
Being a Christian	**.71**	−.21	**.80**	
Having Swiss citizenship	**.68**	.13	**.71**	
Feeling Swiss	**.51**	.40	.38	**.44**
Respecting Swiss political institutions and laws	−.16	**.85**	−.17	**.90**
To be able to speak (Swiss languages)		**.74**		**.84**

Note. Loadings below .10 are not displayed. The highest loading per item is in boldface. ESS = European Social Survey; ISSP = International Social Survey Programme.

conceptions across the two linguistic groups. Changes in fit indices were acceptable for all models but were at the limit (i.e., ΔCFI = .010) for both the model constraining the two means to equality and the model constraining only the ethnic mean. As the changes were considerably larger than for the model constraining only the civic mean to equality (ΔCFI = .002, ΔRMSEA = .001), this model was retained for latent mean comparison. Although adhesion to civic criteria did not differ across the two linguistic groups (κ = 6.20, kappa denotes the latent factor mean in CFA notations), German-speaking respondents expressed a higher support to ethnic criteria (κ = 2.99) compared with French-speaking respondents (κ = 2.36).

International Social Survey Programme. The initial structure was chosen as a function of the EFA results (Table 1): Having ancestry, being born, having lived in Switzerland for most one's life, being a Christian, and having Swiss citizenship were assigned to a first *ethnic* dimension. Although previous cross-national studies on the ISSP items tapping the conception of nationhood (Heath et al., 2009; Reeskens & Hooghe, 2010) revealed strong cross-loadings for the "having the national citizenship" item, in the Swiss sample, this item was clearly a part of the ethnic dimension. Respecting Swiss political institutions and laws and being able to speak the language were assigned to a second *civic* dimension. Again, a strong congruence between the factor solutions across the two regions (Procrustes statistics = .047) was revealed. Finally, the "feeling Swiss" item was not used in the MGCFA because EFA revealed strong cross-loadings in both linguistic groups.

Table 2

Fit Indices for Models Testing Configural, Metric, Partial Scalar, Full Scalar, and Latent Mean Equivalence

Survey	Model	Df	χ^2	CFI	TLI	RMSEA	SRMR
ESS	Configural	16	97.774	.967	.938	.076	.033
Six items	Metric	20	103.356	.966	.949[a]	.069[a]	.037
	Partial scalar (one item relaxed)	23	111.431	.964	.953[a]	.066[a]	.036[a]
	Full scalar	24	216.479	.922	.902	.096	.080
	Partial scalar, civic mean = equivalent	24	116.814	.962	.953	.067	.039
	Partial scalar, ethnic mean = equivalent	24	136.652	.954	.943	.073	.046
	Partial scalar, both means = equivalent	25	136.727	.954	.945	.072	.046
ISSP	Configural	26	64.181	.974	.958	.057	.030
7 items	Metric	31	72.226	.972	.962[a]	.054[a]	.038
	Partial scalar (two items relaxed)	34	78.657	.970	.963[a]	.054	.042
	Full scalar	36	109.207	.951	.942	.067	.055
	Partial scalar, civic mean = equivalent	35	97.008	.958	.950	.063	.043
	Partial scalar, ethnic mean = equivalent	35	103.351	.954	.945	.066	.057
	Partial scalar, both means = equivalent	36	113.016	.948	.939	.069	.058

[a]When testing for a stricter level of equivalence, fit indices are expected to worsen (i.e., lower CFI and TLI and higher RMSEA and SRMR). However, changes in the other direction (i.e., higher CFI and TLI and lower RMSEA and SRMR) are possible because most fit indices are a function of the number of degrees of freedom (see Cheung & Rensvold, 2002, for a detailed explanation). CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

Initial models testing for both configural and metric equivalence yielded good fit indices (Table 2). Though the indices obtained for the final model testing for full scalar equivalence fell within the acceptable range, the decrease in CFI (.021) exceeded Chen's (2007) recommendations. Based on MI, equality constraints on item intercepts were released until the change in CFI compared with the metric equivalence model was equal or smaller than .010. Accordingly, intercepts for two items, "have lived in Switzerland" and "being born in Switzerland," were relaxed. The resulting partial scalar equivalence model was supported by the data. When comparing latent means, none of the three models with mean equivalence constraints (see last three lines of Table 2) showed acceptable change in fit indices: The change in CFI for the model that received the better indices, that is, the model constraining only the civic mean to equality, was .012. In line with our prediction, it can thus be concluded that German-speaking respondents expressed a higher support

for both ethnic ($\kappa = 2.93$) and civic criteria ($\kappa = 3.49$) than French-speaking respondents (ethnic: $\kappa = 2.53$; civic: $\kappa = 3.21$).

Discussion

The present study demonstrates that, in Switzerland, members of the German-speaking majority tend to be stricter than members of the French-speaking minority regarding who should belong to the national community. Strict criteria received stronger support from members of the national majority than from members of the national minority, but this appeared to be the case only for ethnic criteria in the ESS and for both types of criteria in the ISSP. The differences in findings suggest that researchers should be cautious when drawing conclusions based on a single data set. As several explanations can be put forward (e.g., similar but not identical questions and a greater number of criteria in ISSP), it is difficult to pinpoint the exact reasons for such differences in findings without replication of the analyses using different data sets across various points in time.

Importantly, the results revealed that differences in measurement across subnational regions, here in the form of nonequivalent item intercepts, are likely to occur. This confirms that testing for measurement equivalence is a necessary step before substantive within-country analyses, as nonequivalence may hamper meaningful comparisons of subnational groups' scores and render the pooling of subnational groups unwarranted. We examine potential reasons for nonequivalent item intercepts and whether these nonequivalent intercepts were likely to bias the conclusions drawn in the present study.

Nonequivalent Item Intercepts

MGCFA analyses revealed that the intercept of the ESS "commitment to the way of life" item was not equivalent across regions. As a nonequivalent intercept indicates a systematic bias for an item in a group, insight into the causes of such nonequivalence can be gained through the examination of specific aspects of each group. In the present case, an inaccurate translation (clearly different wording) is likely to have caused the nonequivalent intercept. Although the German item asks if someone has to adopt ("annehmen") the Swiss way of life when they come and live in Switzerland, the French item includes two verbs: respect and integrate ("respecter et intégrer"). Respecting the local way of life does not necessarily involve adopting it and is likely to be perceived as a basic requirement to live in any society, which could plausibly explain the stronger support received for this characteristic in the French-speaking sample.

Although in the present case, we can suspect with a certain degree of certainty that translation problems are the cause of nonequivalent measurement, the situation might be less obvious in other cases. For instance, in the present study, analyses revealed two nonequivalent intercepts in the ISSP data. Further inquiries did not highlight translation issues. For such unclear cases, methods based on item response theory may be used because they provide unique information regarding items (e.g., several

parameters are estimated per item instead of one intercept; Meade & Lautenschlager, 2004), and may provide insight into the cause of nonequivalence.

Comparing Latent Means in Case of Partial Scalar Equivalence

For both ESS and ISSP models, testing for partial scalar equivalence fitted the data better than models testing for full scalar equivalence. The validity of conclusions drawn from partially equivalent latent variables is under debate: Some scholars affirm that such comparisons are valid (e.g., Byrne et al., 1989), whereas others have doubts (e.g., De Beuckelaer, 2005). We recommend comparing the latent means provided by the models testing for full and partial scalar equivalence (for similar suggestions see Chen, 2008). If the rank order of the means does not change (i.e., if the means are ordered in the same way), then one may rely on the partial scalar equivalence models. Additional analyses (not provided) reveal that in both ESS and ISSP data, latent means are ordered in the same way in all models: civic criteria receive the strongest support, first in the German-speaking group and then in the French-speaking group, followed by support for ethnic criteria, again first in the German-speaking group and then in the French-speaking group.

Conclusion

Using a Swiss example, the present study demonstrated that testing for measurement equivalence is a necessary step before comparing or pooling data from distinct sub-national groups. In the absence of this step, all items in the conception of nationhood scales might have been used to estimate means (i.e., average support to ethnic and civic criteria in the two groups). This would have resulted in biased comparisons and conclusions. It is thus crucial to examine whether similar factors are present in the groups under study and, in a second step, whether measurement can be considered as equivalent across groups. Finally, the findings underscore the need to cautiously examine latent means obtained from partially equivalent models and compare them with those obtained from fully equivalent models.

References

Baumgartner, H., & Steenkamp, J.-B. E. M. (1998). Multi-group latent variable models for varying numbers of items and factors with cross-national and longitudinal applications. *Marketing Letters, 9*, 21–35. doi:10.1023/A:1007911903032.

Barrett, P. (1986). Factor comparison: An examination of three methods. *Personality and Individual Differences, 7*, 327–340. doi:10.1016/0191-8869(86)90008-5.

Billiet, J. (2003). Cross-cultural equivalence with structural equation modeling. In J. A. Harkness, F. J. R. Van de Vijver & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 247–264). New York: John Wiley.

Billiet, J., Maddens, B., & Beerten, R. (2003). National identity and attitude toward foreigners in a multinational state: A replication. *Political Psychology, 24*, 241–257. doi:10.1111/0162-895X.00327.

Brubaker, R. (1992). *Citizenship and nationhood in France and Germany*. Cambridge, MA: Harvard University Press.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. London: Laurence Erlbaum Associates.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issues of partial measurement invariance. *Psychological Bulletin, 105*, 456–466. doi:10.1037/0033-2909.105.3.456.

Byrne, B. M., & Stewart, S. M. (2006). Teacher's corner: The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling: A Multidisciplinary Journal, 13*, 287–321. doi:10.1207/s15328007sem1302_7.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 464–504. doi:10.1080/10705510701301834.

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*, 1005–1018. doi:10.1037/a0013193.

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31*, 187–212. doi:10.1177/0022022100031002003.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233–255. doi:10.1207/S15328007SEM0902_5.

Davidov, E. (2009). Measurement equivalence of nationalism and constructive patriotism in the ISSP: 34 Countries in a comparative perspective. *Political Analysis, 17*, 64–82. doi:10.1093/pan/mpn014.

Davidov, E., Datler, G., Schmidt, P., & Schwartz, S. H. (2011). Testing the invariance of values in the Benelux countries with the European Social Survey: Accounting for ordinality. In E. Davidov, P. Schmidt & J. Billiet (Eds.), *Methods and applications in cross-cultural analysis*. NY: Routledge.

Davidov, E., & De Beuckelaer, A. (2010). How harmful are survey translations? A test with Schwartz's Human Values Instrument. *International Journal of Public Opinion Research, 22*, 485–510. doi:10.1093/ijpor/edq030.

De Beuckelaer, A. (2005). *Measurement invariance issues in international management research* (Unpublished doctoral dissertation). Hasselt University, Diepenbeek.

European Social Survey (2002). ESS Round 1: European Social Survey Round 1 Data (2002). Data file edition 6.2. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

Green, E. G. T. (2009). Who can enter? A multilevel analysis on public support for immigration criteria across 20 European countries. *Group Processes & Intergroup Relations, 12*, 41–60. doi:10.1177/1368430208098776.

Green, E. G. T., Sarrasin, O., Fasel, N., & Staerklé, C. (2011). Nationalism and patriotism as predictors of immigration attitudes in Switzerland: A municipality-level analysis. *Swiss Political Science Review, 17*, 369–393. doi:10.1111/j.1662-6370.2011.02030.x.

Heath, A., Martin, J., & Spreckelsen, T. (2009). Cross-national comparability of survey attitude measures. *International Journal of Public Opinion Research, 21*, 293–315. doi:10.1093/ijpor/edp034.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fix indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. doi:10.1080/10705519909540118.

Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology. *Journal of Cross-Cultural Psychology, 16*, 131–152. doi:10.1177/0022002185016002001.

International Social Survey Programme (2003). *National Identity II, ZA 3910* [codebook]. Cologne, Germany: GESIS.

Jones, F. L., & Smith, P. (2001). Diversity and commonality in national identities: An exploratory analysis of cross-national patterns. *Journal of Sociology, 37*, 45–63. doi:10.1177/144078301128756193.

Kohn, H. (1944). *The idea of nationalism: A study in its origins and background*. New York: Macmillan.

Kriesi, H., Wernli, B., Sciarini, P., & Gianni, M. (1996). *Le clivage linguistique. Problèmes de compréhension entre les communautés linguistiques en Suisse* [The Swiss linguistic cleavage. Comprehension difficulties between the Swiss linguistic communities]. Bern: Swiss Federal Statistical Office.

Kunovich, R. M. (2009). The sources and consequences of national identification. *American Sociological Review, 74*, 573–593. doi:10.1177/000312240907400404.

Kuzio, T. (2002). The myth of the civic state: A critical survey of Hans Kohn's framework for understanding nationalism. *Racial and Ethnic Studies, 25*, 20–39. doi:10.1080/01419870120112049.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 320–341. doi:10.1207/s15328007sem1103_2.

Meade, A. W., & Lautenschlager, G. J. (2004). A Monte-Carlo study of confirmatory - factor analytic tests of measurement equivalence/invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 60–72. doi:10.1207/S15328007 SEM1101_5.

Medina, T. R., Smith, S. N., & Long, J. S. (2009). Measurement models matter: Implicit assumptions and cross-national research. *International Journal of Public Opinion Research, 21*, 333–361. doi:10.1093/ijpor/edp037.

Muthén, L. K., & Muthén, B. (2008). Mplus (Version 5.1) [Computer software]. Los Angeles, CA: Muthén & Muthén.

Reeskens, T., & Hooghe, M. (2010). Beyond the civic-ethnic dichotomy: Investigating the structure of citizenship concepts across thirty-three countries. *Nations and Nationalism, 16*, 579–597. doi:10.1111/j.1469-8129.2010.00446.x.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23–74.

Selig, J. P., Card, N. A., & Little, T. D. (2008). Latent variable structural equation modeling in cross-cultural research: Multigroup and multilevel approaches. In F. J.

R. Van de Vijver, D. A. Van Hemert & Y. H. Poortinga (Eds.), *Multilevel analysis of individuals and cultures* (pp. 93–119). New York: Laurence Erlbaum.

Sidanius, J., Feshbach, S., Levin, S., & Pratto, F. (1997). The interface between ethnic and national attachment: Ethnic pluralism or ethnic dominance? *Public Opinion Quarterly, 61*, 102–133. doi:10.1086/297789.

Smith, A. D. (1991). *National identity*. Reno: University of Nevada Press.

Staerklé, C., Sidanius, J., Green, E. G. T., & Molina, L. (2010). Ethnic minority-majority asymmetry in national attitudes around the world: A multilevel analysis. *Political Psychology, 31*, 491–519. doi:10.1111/j.1467-9221.2010.00766.x.

StataCorp. (2007). *Stata 10 Base Reference Manual*. College Station, TX: Stata Press.

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78–90. doi:10.1086/209528.

Swiss Federal Statistical Office (2011). Resident population according to main language. Retrieved from http://www.bfs.admin.ch/bfs/portal/de/index/themen/01/05/blank/key/sprachen.html.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–70. doi: 10.1177/109442810031002.

Van de Vijver, F. J. R., & Leung, L. (1997). *Methods and data analysis for cross-cultural research*. London: Sage.

Biographical Notes

Oriane Sarrasin is a postdoctoral researcher (Swiss National Science Foundation) at the Institute of Social Sciences, University of Lausanne, Switzerland. Her research focuses on different forms of prejudice (i.e., immigration attitudes and sexism), national attachment, and measurement equivalence of attitudes in surveys.

Eva G. T. Green is senior lecturer at the Institute of Social Sciences, University of Lausanne. At the cross-roads of social and political psychology, her research interests include prejudice and immigration attitudes, links between national and ethnic identity, and scientific communication among lay people, as well as measurement quality in survey data.

André Berchtold is associate professor of statistics at the Institute of Social Sciences, University of Lausanne. He is mainly interested in applying statistics to health research and developing new statistical models for the social sciences.

Eldad Davidov is Professor of Sociology at the University of Zurich, Switzerland. His research interests are applications of structural equation modeling to survey data, especially in cross-cultural and longitudinal research. Applications include human values, national identity, and attitudes toward immigrants and other minorities.