A Cross-Cultural Comparison of Self-Efficacy as a Resilience Measure: Evidence From PISA 2018

Article *in* Youth & Society · August 2023

DOI: 10.1177/0044118X231186833

CITATIONS
0

READ
1

3 authors:

Sibel Kaya
University of Bedfordshire
52 PUBLICATIONS **531** CITATIONS
SEE PROFILE

Nurullah Eryilmaz
International Association for the Evaluation of Ed…
13 PUBLICATIONS **55** CITATIONS
SEE PROFILE

Dogan Yuksel
The Open University (UK)
77 PUBLICATIONS **487** CITATIONS
SEE PROFILE

Article

Youth & Society
1–25

DOI: 10.1177/0044118X231186833
journals.sagepub.com/home/yas
Sage

A Cross-Cultural Comparison of Self-Efficacy as a Resilience Measure: Evidence From PISA 2018

Sibel Kaya[1], Nurullah Eryilmaz[2], and Dogan Yuksel[3]

Abstract

This study explored the equivalence of resilience across countries and economies that participated in PISA 2018. A total of 79 countries and economies were divided into ten sub-groups based on their socio-demographic characteristics. Analysis of the comparability of the PISA self-efficacy scale as a measure of resilience across the participating countries/economies in the study was conducted using multi-group confirmatory factor analysis (MG-CFA). The results demonstrated that across all countries and economies, the configural invariance level, which is the lowest level of invariance, has been reached but the metric and scalar invariance levels have not been reached. Within-group results showed that all sub-groups presented a model fit for the metric level of invariance. However, only the Anglo countries were able to reach the strict invariance level. This finding indicates that the Anglo countries were more homogeneous in terms of their interpretation of self-efficacy in PISA, whereas other sub-groups were more heterogeneous. Confirming the notion of cultural affiliation of resilience, it was concluded that self-efficacy

[1]University of Bedfordshire, University Square, Luton, UK
[2]International Association for the Evaluation of Educational Achievement, Hamburg, Germany
[3]The Open University, Milton Keynes, UK

Corresponding Author:
Sibel Kaya, Academy for Learning and Teaching Excellence, University of Bedfordshire, University Square, Luton LU1 3JU, UK.
Email: sibel.kaya@beds.ac.uk

by itself might not be an adequate indicator of resilience. The current study has some recommendations for future research and how PISA can be more inclusive about the constructs it employs.

Keywords
resilience, self-efficacy, PISA, multi-group confirmatory factor analysis, invariance

Introduction

Resilience is "the ability of individuals (on their own and collectively) to navigate to the culturally relevant resources they need to do well when confronting adversity" (Ungar, 2015, p. 40). Resilience is an essential requirement for success, not only in school but also in life. Previous studies demonstrate that there is a strong relationship between resilience and emotional stability (Chandelkar & Shetty, 2019), perceived stress (Chandelkar & Shetty, 2019; Solomon, 2013), self-regulated learning (Paloș et al., 2020, Wang et al., 1994), engagement with learning environment (Benard, 1991; Finn & Rock, 1997), 21st-century skills (Baas, 2020) and academic performance (Abolmaali & Mahmudi, 2013; Jafri, 2013; Novotný & Křeménková, 2016; Solomon, 2013).

In recent years, leading theorists have agreed that resilience is not a trait that only certain children are naturally born with, but rather it is a dynamic state (Luthar et al., 2000; Masten, 2001; Rutter, 2012; Ungar, 2013) continually in flux with other variables. Individuals can develop resilience at any age through reciprocal relationships with their environments (Gillespie et al., 2007) and this process can vary in different contexts (Rutter, 2012). What makes a person resilient is the mixture of personal characteristics and the environment (Ungar, 2008; Ungar, 2013).

The resilience level of students and the conceptualization of resilience may differ across cultures. For instance, in Eastern cultures, resilience can be a collective effort toward the pursuit of justice and power; whereas, in Western cultures, it is usually interpreted as the pursuit of individual-level well-being (Panter-Brick & Eggerman, 2012). This matter has fundamental significance to making meaningful cross-cultural comparisons across countries. If the operationalization of the resilience construct relies on one specific cultural sight, making comparisons using the mean score of the resilience construct would not be valid across countries (He & Van de Vijver 2013). Furthermore, the measurement of resilience with a self-report assessment (Oshio et al., 2003, Tugade

& Fredrickson, 2004) may not be adequate to understand adolescent resilience. Rather, analytical approaches that examine relationships among risk and promotive factors are needed (Masten & Powell, 2003; Zimmerman & Arunkumar, 1994). Even though the conceptualization of resilience goes beyond personal characteristics, certain intrapersonal characteristics are key to developing resilience such as self-efficacy (DiCorcia & Tronick, 2011; Fergus & Zimmerman, 2005; Hays-Grudo & Morris, 2020; Henry et al., 2015; Luthans et al., 2007; Windle, 2011). The Programme for International Student Assessment (PISA) used self-efficacy as a measure of students' resilience beliefs in the 2018 assessment (Nelis et al., 2021), positive values on this index indicating higher resilience beliefs than the Organization for Economic Cooperation and Development (OECD) average (OECD, 2019).

PISA collects information on students' mathematics, science, reading literacy and various non-cognitive constructs (Steiner-Khamsi, 2019). Considering the large number of countries and economies participating in PISA (79 in the 2018 assessment), whether the scales that measure perceptions, beliefs, and attitudes of participants in PISA (in our case resilience beliefs) have the same meaning across different countries carry critical importance (He & Van de Vijver, 2013).

Resilience as Assets and Resources

Resilience depends on the co-existence of promotive factors and potential risks that can either contribute to a positive outcome or prevent a negative one. The promotive factors that can assist young people in avoiding the adverse consequences of risks can be categorized as *assets* and *resources* (Beauvais & Oetting, 1999). Assets are considered to be internal/intrapersonal such as competence and self-efficacy. Resources are external/interpersonal such as parental support and community organizations that encourage growth and well-being of young people in a positive way (Fergus & Zimmermann, 2005; Sacker & Schoon, 2007; Windle, 2011). As Fergus and Zimmermann (2005) state, there are some common assets and resources that support young people in overcoming risks such as self-efficacy and parental support regardless of the specific risk. They caution however, individual-level assets such as self-efficacy should not indicate that resilience is something that only exists within an individual or as a static characteristic of a person. Resources can shift the focus towards the impact of social and environmental factors on the development and health of adolescents, situating resilience theory within an ecological framework. This also departs from the idea of resilience as an unchanging, personal attribute (Sacker & Schoon, 2007).

The conceptualizations and mechanisms of resilience are shaped by the intersecting cultures within social ecologies. The continuous interactions between individuals and the interactions of individuals with different levels of environment facilitate the development of adaptive behavior (Wright & Masten, 2015). From a social-ecological perspective, resilience is defined as a process of adjusting to adversities where cultural orientations are taken into account (Masten, 2013; Ungar, 2011). Both internal and external characteristics play a crucial role in determining resilience in a person (Ungar, 2008). Transactions take place between individuals and the environment, but the adaptive capacity might be different in each individual as adolescents may not exhibit the same characteristics during and after facing a risk (Kuldas & Foody, 2022; Ungar, 2008). Ungar (2011) posits that a person's capacity to adjust during times of adversity is determined not just by the presence of resources such as social, cultural, and physical support, but also by the individual's capability to actively search for and make use of these resources. Therefore, Kuldas and Foody (2022) describe resilience as a "dynamic process, capacity, and outcome of developing person in a corresponding population, context, and risk" (p. 1369).

Ungar (2015) states that resilience can be interpreted differently in different cultures. For example, in impoverished societies, resilience may be defined as the ability to establish a job and gain income, yet in more affluent societies, it might be understood as the result of staying in school and continuing education. In a similar vein, in collectivist settings, young people are considered to be resilient when they attach to a network and become part of the community (Theron et al., 2013). Comparatively, in individualist settings, young people are considered resilient when they gain autonomy and financial independence (Lesko, 2001). Certain behaviors may have various connotations in different cultures regarding their influence on resilience. For instance, school dropout is considered to have a negative effect on resilience in most cultures, whereas it demonstrates dignity and independence among a group of African Canadian students (Dei et al., 1997).

Even though the combination of qualitative and quantitative methods is recommended in cross-cultural resilience research (He & Van de Vijver, 2015; Van de Vijver & Chasiotis, 2010), quantitative studies are more convenient in terms of handling equivalence issues. Equivalence is defined as "the level of comparability of scores across cultures" (He & Van de Vijver, 2015, p. 190). When there is no equivalence in quantitative scores, then the difference in observed scores may not be due to the genuine differences in targeted attributes but rather, due to the different psychological interpretations of items (He & Van de Vijver, 2015). The current study aims to test the equivalence of the resilience scale across countries and economies that participated in PISA 2018.

PISA used self-efficacy as a measure of students' resilience beliefs in the 2018 assessment. Thus, the current study attempts to investigate the extent to which mean comparisons of cross-cultural differences in young people's self-efficacy level as a measure of resilience. In doing so, we employed multiple-group confirmatory factor analysis (MG-CFA) which is an extensively engaged technique to investigate the degree to which an instrument is invariant across countries (Jöreskog, 1971; Millsap, 2012). This study aims to test the equivalence of self-efficacy, an intrapersonal asset of resilience, cross-culturally by using PISA 2018 data from 79 countries.

Measurement Invariance

Measurement tools or indicators should be interpreted in the same manner for all study groups to generate valid comparisons of average scores across groups (Rutkowski & Svetina, 2014). Failure to reach measurement invariance increases the possibility of inaccurate results and conclusions when comparing indicators across groups due to methodological artefacts. Therefore, regardless of participants' ethnicity, nationality, or socioeconomic backgrounds, it is necessary to establish reliable group comparisons between various circumstances and contexts (Byrne & van de Vijver, 2010).

Utilizing multiple-group confirmatory factor analysis is the traditional and most popular method for testing measurement invariance (Jöreskog, 1971; Putnick & Bornstein, 2016). Configural invariance is the lowest level of invariance. At this level, the concept should have a consistent factor structure across the groups as a starting point for the testing of measurement invariance (Millsap, 2012; Schmitt et al., 2011). Despite the existence of configural invariance, the latent construct is not comparable between the groups. It simply serves as a prerequisite for evaluating the following levels (Brown, 2015). Metric invariance is the next level of invariance. The cross-group factor loadings are expected to be equal at this level, indicating that the items' relationships with the latent factors are equivalent across the comparison groups (Chen, 2007; Cheung & Rensvold, 2002). We are able to conduct reliable between-country regression analysis once we reach metric invariance. However, this degree of invariance does not support comparing the latent construct's mean scores across groups.

When the item thresholds and factor loadings are the same across groups, the scalar level of invariance is obtained. This level of invariance lends credence to the idea that respondents from various groups interpret the response scale in a consistent manner. As a result, achieving a scalar degree of invariance allows for a reliable cross-group comparison of average latent scores (Millsap & Olivera-Aguilar, 2012).

Equal residual covariances, identical factor loadings, and identical intercepts between the groups are all implied by the strict level of invariance (Chen, 2007; Cheung & Rensvold, 2002). For each item in the scales, the variance not explained by the unobserved variables is the same across groups at this level of invariance (i.e. the same degree of error across groups) (Vandenberg & Lance, 2000). Overall, a construct must achieve the scalar level of invariance or above in order to be reliably compared across groups.

Measurement Invariance in International Large-scale Assessments (ILSAs)

Research on the measurement invariance of cognitive assessment has always been popular (Klieme, 2016; Oliveri & Ercikan, 2011), whereas very little attention has been paid to non-cognitive constructs' measurement invariance (Ding et al., 2023; He et al., 2019; Pedrero & Manzi, 2020). However, in recent years, there has been a growing interest in the measurement invariance of scales created from the responses to student background questionnaires (Ding et al., 2023; Eryilmaz et al., 2020; He et al., 2019; Pedrero & Manzi, 2020; Rutkowski & Svetina 2017; Rutkowski & Rutkowski, 2018; Sandoval-Hernandez et al., 2019; Wurster, 2022).

Recently, Ding et al. (2023) investigated the comparability of mathematics self-efficacy and self-concept constructs using PISA 2003 and 2012 cycles. They found that both factors and factor means of mathematics self-concept and self-efficacy were not comparable across participating countries. In another study, Pedrero and Manzi (2020) tested the invariance of six non-cognitive measures in TIMSS 2011 at the fourth-grade level. Based on their analyses, they concluded that the intrinsic motivation scales in mathematics and science showed a higher degree of cross-cultural equivalence, whereas the self-beliefs scale showed no degree of invariance. Moreover, He et al. (2019) investigated the cross-cultural comparability of non-cognitive measures in TIMSS and PISA, and they found that the three scales in both surveys (except the sense of school belonging in PISA) reached at least metric invariance. Lastly, Wurster (2022) investigated the measurement invariance of non-cognitive measures in TIMSS 2007, 2011, and 2015 for 26 countries, and measurement invariance analyses indicated a configural invariance across countries.

In the last few years, studies of measurement invariance of non-cognitive structures have become important as we have shown in the literature. We examined the measurement invariance of self-efficacy which was used as a measure of resilience in the PISA 2018 across groups of participating countries and economies.

Present Study

The present study aims to investigate the measurement invariance of the self-efficacy scale, which was used as a measure of resilience using the PISA 2018 survey. Analysis of the comparability of the resilience across participating countries and economies in the study is conducted using multi-group confirmatory factor analysis (MG-CFA). In the second stage, to determine which subgroups' conception of resilience is closest to the PISA conceptualization of resilience, and to evaluate the comparability between the subgroups, the participating countries and economies were divided into subgroups depending on their socio-demographic characteristics. The ten sub-groups were identified as Anglo, Central and Eastern European, Confucian, Latin American, Middle Eastern and North African, Nordic, Southern European, the Association of Southeast Asian Nations (ASEAN), Western and Central Asian, and Western European (refer to Table 2) following previous research on PISA and other ILSAs (e.g., Eryilmaz et al., 2020; GLOBE Foundation, n.d.; Isac et al., 2019; Liu 2020). The invariance of the resilience scale between these groups is investigated using MG-CFA.

Cross-cultural comparison of constructs may pose challenges, yet this endeavor often relies on model-fit evaluation criteria. We can test if more homogeneous groups achieve a better level of invariance by utilizing subgroup analysis. Our research has two goals: first, we aim to inform secondary users of PISA 2018 data about how the self-efficacy scale can be compared across cultures as a measure of resilience. Second, we offer data for the scales' conceivable future development. This study's findings can help ILSAs in developing more cross-culturally comparable instruments. This study displays how surveys may potentially have measurement invariance in the cases of cross-context/country comparison. Given the significance of comparative analysis in educational research, particularly when ILSAs are employed, the current study is of particular importance. This study expands the invariance tests to include non-cognitive student variables, which adds to our understanding of the degree of comparability between cultures.

Methods

Data

This study used the data from PISA 2018 cycle. Overall, 79 countries and economies participated in PISA in 2018. Details about participating countries and economies were presented in Table 1.

Table 1. Selection of Country/Economy Clusters.

Anglo	ASEAN	Central and Eastern European	Confucian	Latin American	Middle Eastern and North African	Nordic	Southern European	Western and Central Asian	Western European
-Australia	-Brunei Darussalam	-Bulgaria	-BSJZ China	-Argentina	-Israel	-Denmark	-Albania	-Baku Azerbaijan	-Austria
-Canada	-Indonesia	-Belarus	-Chinese Taipei	-Brazil	-Jordan	-Finland	-Bosnia and Herzegovina	-Kazakhstan	-Belgium
-Ireland	-Malaysia	-Czech Rep.	-HongKong	-Chile	-Lebanon	-Iceland	-Croatia	-Turkey	-France
-New Zealand	-Philippines	-Estonia	-Japan	-Colombia	-Morocco	-Norway	-Greece	-Georgia	-Germany
-UK	-Thailand	-Hungary	-Korea	-Costa Rica	-Qatar	-Sweden	-Kosovo	-Tatarstan	-Luxemburg
-US		-Latvia	-Macao	-Dominican Rep.	-UAE		-Malta		-Netherlands
		-Lithuania	-Singapore	-Mexico	-Saudi Arabia		-Montenegro		-Portugal
		-Moldova	-Vietnam	-Panama			-North Macedonia		-Spain
		-Moscow Reg.		-Peru			-Slovenia		-Switzerland
		-Poland		-Uruguay			-Serbia		
		-Romania					-Italy		
		-Russian Fed.							
		-Slovak Rep.							
		-Ukraine							

Note. We considered country groupings used in previous studies (see Eryilmaz et al., 2020; GLOBE Foundation, n.d.; Isac et al., 2019; Liu 2020).

Variables

Resilience in PISA 2018 is based on five self-efficacy items, reflecting students' abilities to cope with difficulties. The students were asked to rate the following statements on a four-point Likert scale, ranging from "Strongly disagree" and "Disagree," to "Agree" and "Strongly agree":

- *I usually manage one way or another.*
- *I feel proud that I have accomplished things.*
- *I feel that I can handle many things at a time.*
- *My belief in myself gets me through hard times.*
- *When I'm in a difficult situation, I can usually find my way out of it.*

Analytical Strategy

Three steps that make up the analytical technique in this study are 1. Internal consistency analysis (reliability), 2. Confirmatory factor analysis (CFA), 3. Multi-group confirmatory factor analysis (MG-CFA).

Internal consistency (reliability). First, we examined the internal consistency (reliability) of the self-efficacy scale as the resilience measure in each country and economy separately. Since Cronbach's alpha is the commonly used reliability metric, we reported the results for each system's Cronbach's alpha values.

Confirmatory factor analysis (CFA). Second, a confirmatory factor analysis (CFA) was conducted to evaluate the consistency of the model data across each country/economy. To assess the consistency of the model data within each system, root mean square error of approximation (RMSEA), chi-squared test, comparative fit index (CFI), and Tucker-Lewis index (TLI) are frequently used. It is critical to emphasize that the closer the RMSEA and SRMR values are to 0 and the closer the CFI and TLI values are to 1 the better the model fit. An acceptable model fit was considered for $CFI > 0.90$; $TLI > 0.90$; $RMSEA < 0.10$; and $SRMR < 0.08$ (Hu & Bentler, 1999; Rutkowski & Svetina, 2014).

Multi-group confirmatory factor analysis (MG-CFA). Finally, we used a multi-group confirmatory factor analysis (MG-CFA) to assess the model's fit for invariance at different levels and the resilience scales' cross-cultural comparability. One of the most widely used research methods to examine measurement invariance is MG-CFA, which is a basic extension of CFA (Hair et al.,

2010; Kline, 2011). We applied this approach to ten distinct homogeneous country and economy sub-groups, where homogeneity is determined by each country and economy's socio-demographic traits.

Following Brown (2015)'s recommendation to estimate configural, metric, scalar, and stringent invariance models from the least restrictive to the most restricted model, we estimated the measurement invariance models (for more information, see the literature review).

The self-efficacy scale as the resilience measure was first subjected to MG-CFA for the 79 sampled countries and economies. Then, the same process was applied separately to each of the ten sub-groups. For each invariance level, the goodness of fit metrics were evaluated and contrasted with the preceding, less constrained model. To assess the goodness of fit indices, researchers suggest the cut-off value of $CFI > 0.95$ and $RMSEA < 0.1$ (Brown, 2015; Rutkowski & Svetina, 2014). The change in fit between one model and the next was assessed using the comparative fit index (CFI). The absolute difference in CFI was used to evaluate the change in fit. Rutkowski and Svetina (2014) suggested the criterion of -0.020 for ΔCFI to determine metric invariance (Rutkowski & Svetina, 2014). For scalar and strict invariance, the traditional cut-off value of -0.010 was used for ΔCFI (Cheung & Rensvold, 2002; Desa, 2014; Rutkowski & Svetina, 2014).

Estimates were created using R statistical software. Among R packages, lavaan (Rosseel, 2012) and lavaan.survey (Oberski, 2014) are considered most suitable for complicated survey designs such as PISA (R Core Team, 2019).

Results

Internal Consistency (Reliability)

All countries and economies have a Cronbach's alpha coefficient above 0.6, for the resilience measure used in PISA 2018. In other words, the internal consistency of self-efficacy as the resilience measure was confirmed for all countries and economies participating in PISA 2018.

Confirmatory Factor Analysis (CFA)

Country/economy level CFA models are shown in Table 2. All countries and economies met the criteria provided by Hu and Bentler (1999). The next step is to test the measurement invariance of the resilience scale across countries and economies.

Table 2. Confirmatory Factor Analysis for the Resilience Scale for Each Country/Economy.

Country/Economy	CFI	TLI	RMSEA	SRMR	Chi-square	*df*	Reliability	*n*
Albania	0.988	0.975	0.063	0.015	121.852	5	0.82	5893
Azerbaijan (Baku)	0.989	0.978	0.080	0.015	152.79	5	0.89	4636
Argentina	0.971	0.942	0.079	0.023	317.081	5	0.75	9916
Australia	0.967	0.934	0.092	0.026	497.473	5	0.78	11745
Austria	0.964	0.928	0.094	0.029	274.477	5	0.77	6101
Belgium	0.925	0.851	0.105	0.037	248.838	5	0.67	4404
Bosnia and Herz	0.977	0.955	0.087	0.023	217.445	5	0.83	5569
Brazil	0.976	0.951	0.069	0.021	199.727	5	0.74	8117
Brunei Darussalam	0.943	0.887	0.080	0.03	199.765	5	0.65	6150
Bulgaria	0.947	0.894	0.150	0.035	465.868	5	0.85	4079
Belarus	0.947	0.893	0.100	0.032	280.49	5	0.73	5486
Canada	0.958	0.916	0.103	0.029	1050.037	5	0.79	19744
Chile	0.957	0.915	0.121	0.031	472.953	5	0.82	6376
Chinese Taipei	0.964	0.928	0.098	0.030	341.628	5	0.79	7075
Colombia	0.974	0.949	0.088	0.023	252.357	5	0.80	6371
Costa Rica	0.952	0.904	0.112	0.033	401.86	5	0.79	6279
Croatia	0.954	0.908	0.104	0.031	336.16	5	0.77	6150
Czech Republic	0.926	0.852	0.120	0.038	462.736	5	0.73	6344
Denmark	0.961	0.921	0.093	0.029	278.63	5	0.76	6307
Dominican Republic	0.980	0.960	0.094	0.021	131.286	5	0.85	2843
Estonia	0.948	0.895	0.116	0.033	342.164	5	0.79	5002
Finland	0.962	0.924	0.104	0.028	287.374	5	0.80	5193
France	0.914	0.828	0.128	0.041	461.337	5	0.73	5548
Georgia	0.980	0.960	0.080	0.022	143.044	5	0.82	4364
Germany	0.954	0.908	0.094	0.030	172.412	5	0.73	3817
Greece	0.971	0.942	0.079	0.024	185.546	5	0.75	5856
Hong Kong	0.981	0.962	0.077	0.019	174.064	5	0.81	5654
Hungary	0.966	0.932	0.083	0.025	199.553	5	0.75	5654
Iceland	0.968	0.935	0.124	0.026	231.408	5	0.87	2934
Indonesia	0.952	0.905	0.109	0.033	680.243	5	0.77	11458
Ireland	0.934	0.867	0.112	0.035	338.594	5	0.73	5300
Israel	0.964	0.929	0.098	0.027	283.225	5	0.79	5739
Italy	0.922	0.844	0.114	0.040	678.824	5	0.72	10317
Kosovo	0.977	0.954	0.082	0.022	152.687	5	0.81	4348
Japan	0.927	0.853	0.138	0.041	568.329	5	0.78	5943

(continued)

Table 2. (continued)

Country/ Economy	CFI	TLI	RMSEA	SRMR	Chi-square	*df*	Reliability	*n*
Kazakhstan	0.972	0.944	0.092	0.031	743.22	5	0.79	17623
Jordan	0.965	0.930	0.109	0.027	493.904	5	0.81	8159
Korea	0.977	0.953	0.090	0.025	270.36	5	0.82	6543
Lebanon	0.983	0.965	0.065	0.019	96.142	5	0.77	4351
Latvia	0.968	0.936	0.078	0.024	152.737	5	0.74	4885
Lithuania	0.956	0.913	0.104	0.029	340.916	5	0.78	6244
Luxemburg	0.955	0.911	0.111	0.030	293.063	5	0.80	4638
Macao	0.963	0.926	0.084	0.026	136.328	5	0.72	3738
Malaysia	0.982	0.963	0.052	0.018	85.906	5	0.70	5930
Malta	0.952	0.903	0.109	0.031	182.904	5	0.78	3014
Mexico	0.956	0.912	0.113	0.032	351.932	5	0.80	5445
Moldova	0.969	0.938	0.074	0.024	145.526	5	0.73	5114
Montenegro	0.986	0.971	0.068	0.018	141.204	5	0.82	5833
Morocco	0.962	0.924	0.095	0.028	191.185	5	0.77	4164
Netherlands	0.961	0.922	0.085	0.027	138.403	5	0.73	3700
New Zealand	0.935	0.870	0.114	0.036	378.429	5	0.74	5765
Northern Macedonia	0.962	0.925	0.109	0.030	305.925	5	0.83	5025
Panama	0.960	0.921	0.103	0.028	187.618	5	0.79	3448
Peru	0.969	0.939	0.086	0.027	160.364	5	0.77	4208
Philippines	0.994	0.988	0.038	0.011	54.077	5	0.78	6755
Poland	0.969	0.939	0.088	0.026	211.368	5	0.78	5375
Portugal	0.956	0.912	0.092	0.028	233.23	5	0.75	5375
Qatar	0.980	0.960	0.067	0.021	260.117	5	0.76	11510
Romania	0.970	0.940	0.081	0.024	160.826	5	0.75	4809
Russian Federation	0.977	0.954	0.077	0.022	203.498	5	0.79	6783
Saudi Arabia	0.980	0.961	0.077	0.021	167.296	5	0.81	5530
Serbia	0.983	0.966	0.075	0.019	157.965	5	0.82	5422
Singapore	0.949	0.897	0.099	0.031	326.819	5	0.74	6537
Slovak Republic	0.947	0.894	0.117	0.032	365.456	5	0.78	5299
Vietnam	0.929	0.857	0.077	0.030	164.658	5	0.57	5329
Slovenia	0.960	0.920	0.107	0.029	340.956	5	0.81	5828
Spain	0.938	0.877	0.119	0.036	2269.628	5	0.75	32171
Sweden	0.963	0.925	0.110	0.028	303.889	5	0.82	4903
Switzerland	0.941	0.882	0.106	0.034	280.766	5	0.73	4890
Thailand	0.944	0.888	0.118	0.036	582.072	5	0.76	8354
UAB	0.964	0.929	0.110	0.028	1034.323	5	0.84	17073

(continued)

Table 2. (continued)

Country/ Economy	CFI	TLI	RMSEA	SRMR	Chi-square	*df*	Reliability	*n*
Turkey	0.986	0.972	0.074	0.018	187.302	5	0.85	6638
Ukraine	0.960	0.920	0.086	0.026	215.717	5	0.73	5742
UK	0.961	0.922	0.096	0.029	572.266	5	0.77	12407
USA	0.962	0.925	0.098	0.028	221.253	5	0.79	4548
Uruguay	0.935	0.871	0.140	0.039	383.683	5	0.80	3863
BSJZ China	0.986	0.971	0.066	0.018	261.277	5	0.80	11901
Moscow Region	0.977	0.954	0.074	0.023	54.361	5	0.78	1816
Tatarstan	0.975	0.949	0.081	0.024	176.739	5	0.79	5212

Multi-Group Confirmatory Factor Analysis (MG-CFA)

MG-CFA analyzed the degree of invariance for the PISA 2018 resilience measure among countries and economies, and within subgroups. The results for invariance levels and the change in the CFI for all countries and economies can be found in Table 3. Accordingly, the configural invariance has been reached since the fit values were within acceptable ranges (CFI and TLI ≥ 0.95, RMSEA ≤ 0.10). A metric level of invariance, however, was not supported. The criteria of ΔCFI ≤ 0.020 when moving from configural to metric invariance has not been met. Therefore, based on MG-CFA results, we can state that all the countries and economies have reached the configural invariance level but they have not reached the metric invariance level.

The findings for the measurement invariance models namely, configural, metric, scalar, and strict invariance, for all the sub-groups of countries and economies, are presented in Table 4. Only Anglo countries and economies sub-group confirmed an excellent model fit for the strict invariance level. All the fit values and the change in CFI were within acceptable ranges (e.g. CFI ≥ 0.95; TLI ≥ 0.95; RMSEA ≤ 0.10; ΔCFI ≤ 0.020) for this sub-group from the least restrictive to the most restrictive level. For other country groups, the general adjustment criteria have been met only at the metric invariance level (e.g. CFI ≥ 0.95; TLI ≥ 0.95; RMSEA ≤ 0.10; ΔCFI ≤ 0.020). It can be concluded that country and economy sub-groups other than the Anglo sub-group support a model fit for the metric level of invariance. Table 4 shows each group's invariance level model fit in bold.

Discussion

This study explored the equivalence of self-efficacy as a measure of resilience across countries and economies that participated in PISA 2018. A total

Table 3. Multiple Group Configural, Metric, and Scalar Invariance of the Resilience Scale.

Model	$\chi 2$	*df*	CFI	TLI	RMSEA	SRMR	Model comparison	$\Delta\chi 2$	Δdf	ΔCFI
Configural	13038.82	385	0.961	0.923	0.069	0.029	Metric against configural	6804.568	304	
Metric	20197.43	689	0.940	0.933	0.064	0.072	Scalar against configural	49776.79	608	−0.021
Scalar	61177.78	993	0.816	0.857	0.093	0.116	Scalar against metric	41565.85	304	−0.124

Note. CFI = comparative fit index; TLI = tucker lewis index; RMSEA = root mean square error approximation; ΔCFI = change in CFI.

Table 4. Invariance Test for Different Country/Economy Subgroups (Invariance level model fit in bold).

Country/ economy subgroup	Invariance level	Chi-square	*df*	RMSEA	CFI	TLI	ΔCFI
Anglo Countries	Configural invariance	2790.646	25	0.101	0.958	0.916	
	Metric invariance	2880.4	41	0.080	0.957	0.948	−0.001
	Scalar invariance	3105.871	57	0.071	0.954	0.959	−0.003
	Strict invariance	**3339.162**	**77**	**0.063**	**0.951**	**0.968**	**−0.003**
ASEAN	Configural invariance	1668.513	25	0.092	0.959	0.917	
	Metric invariance	**1959.586**	**41**	**0.078**	**0.952**	**0.941**	**−0.007**
	Scalar invariance	5583.2	57	0.112	0.861	0.878	−0.091
	Strict invariance	8908.908	77	0.122	0.779	0.856	−0.082
Central and Eastern European	Configural invariance	3613.492	70	0.099	0.957	0.915	
	Metric invariance	**4163.123**	**122**	**0.080**	**0.951**	**0.944**	**−0.006**
	Scalar invariance	10940.87	174	0.110	0.871	0.896	−0.080
	Strict invariance	13873.22	239	0.105	0.836	0.904	−0.035
Confucian Countries	Configural invariance	2241.622	40	0.091	0.967	0.934	
	Metric invariance	**2933.266**	**68**	**0.080**	**0.957**	**0.950**	**−0.010**
	Scalar invariance	7965.937	96	0.112	0.883	0.902	−0.074
	Strict invariance	12149.34	131	0.118	0.821	0.891	−0.062
Latin America	Configural invariance	2941.921	50	0.101	0.962	0.924	
	Metric invariance	**3463.484**	**86**	**0.083**	**0.956**	**0.948**	**−0.006**
	Scalar invariance	6217.813	122	0.094	0.920	0.934	−0.036
	Strict invariance	9669.035	167	0.100	0.875	0.925	−0.045
Middle Eastern and North African	Configural invariance	2558.958	35	0.094	0.970	0.940	
	Metric invariance	**3501.529**	**59**	**0.085**	**0.959**	**0.951**	**−0.011**
	Scalar invariance	6589.225	83	0.099	0.923	0.934	−0.036
	Strict invariance	11473.99	113	0.112	0.865	0.916	−0.058
Nordic	Configural invariance	1118.657	20	0.107	0.963	0.926	
	Metric invariance	**1327.689**	**32**	**0.092**	**0.956**	**0.945**	**−0.007**
	Scalar invariance	2581.45	44	0.109	0.914	0.922	−0.042
	Strict invariance	3122.096	59	0.104	0.897	0.930	−0.017
Southern European	Configural invariance	2841.563	55	0.094	0.968	0.936	
	Metric invariance	**3598.214**	**95**	**0.080**	**0.960**	**0.954**	**-0.008**
	Scalar invariance	7583.875	135	0.098	0.915	0.930	−0.045
	Strict invariance	11549.48	185	0.103	0.870	0.923	−0.045
Western Central Asia	Configural invariance	1408.607	25	0.085	0.979	0.958	
	Metric invariance	**1880.979**	**41**	**0.076**	**0.972**	**0.966**	**−0.007**
	Scalar invariance	6333.241	57	0.120	0.904	0.915	−0.068
	Strict invariance	7825.934	77	0.114	0.881	0.923	−0.023
Western European Countries	Configural invariance	4466.592	45	0.112	0.941	0.882	
	Metric invariance	**5073.768**	**77**	**0.091**	**0.933**	**0.922**	**−0.008**
	Scalar invariance	9247.947	109	0.103	0.878	0.899	−0.055
	Strict invariance	11210.27	149	0.098	0.852	0.911	−0.026

Note. CFI = comparative fit index; TLI = tucker lewis index; RMSEA = root mean square error approximation; ΔCFI = change in CFI.

of 79 countries and economies were divided into ten sub-groups based on their socio-demographic characteristics as categorized previously in recent analyses (e.g., Eryilmaz et al., 2020; GLOBE Foundation, n.d.; Isac et al., 2019; Liu, 2020). Analysis of the comparability of the PISA self-efficacy as a measure of resilience scale across the participating countries and economies in the study was conducted using multi-group confirmatory factor analysis (MG-CFA). The results demonstrated that across all the countries and economies, the configural invariance level has been reached but the metric and scalar invariance levels have not been reached. Within-group results showed that all sub-groups presented a model fit for the metric level of invariance. However, only the Anglo countries were able to reach the strict invariance level. This finding indicates that Anglo countries were more homogeneous in terms of their interpretation of the self-efficacy construct in PISA, whereas other sub-groups were more heterogeneous.

Considering the previous assumptions regarding the variance in resilience in different cultures (Ungar, 2015), our findings provide further support regarding cultural similarities among Anglo countries, and differences between Anglo and other countries demonstrating individualist and collectivist characteristics (e.g., Lesko, 2001; Theron et al., 2013). Specifically, one of the main implications of our analysis is that resilience might be understood differently in impoverished and/or collectivist societies and individualist societies. In this respect, utilizing a general scale about the perceptions, beliefs, and attitudes of its participants in ILSAs may not present the most vivid understanding of the construct used in specific settings and cultures. This one-size-fits-all approach might be overly simplified and has the risk of losing sight of the larger picture.

ILSAs provide rich information on students' academic performance as well as various background variables related to learning across countries. This information allows comparisons between different countries with the assumption that measurement scales are universal and not influenced by the individual characteristics of students (Pedrero & Manzi, 2020). However, the findings of the current study, alongside previous studies (e.g., Ding et al., 2023; He et al., 2019; Pedrero & Manzi, 2020; Rutkowski & Rutkowski, 2018; Sandoval-Hernandez et al., 2019) have shown that the ILSAs' background measurement scales are not completely invariant across different countries. ILSAs aim to monitor the development and quality of education systems and promote evidence-based policymaking (e.g., Harju-Luukkainen et al., 2020) with the aim of providing cross-national comparisons to support educational decisions and reforms (Ding et al., 2023). However, we should consider the reliability and validity of the analyses in various cultures and

contexts so that we can have more research-informed results for policy development and decision-making practices.

In the current study, all countries and economies have reached the configural invariance meaning that the one-factor structure of the resilience construct was the same over the groups. However, since the metric invariance has not been achieved, we cannot conclude that self-efficacy scale items' relationships with the one-factor structure are comparable across sub-groups. In other words, the factor loadings of self-efficacy items might be different (Ding et al., 2023). Since metric and scalar invariances have not been reached, correlational comparisons across groups or ranking of groups based on the scale mean scores (i.e. self-efficacy) are not advised (He et al., 2019).

Some constructs in ILSAs have better cross-cultural comparability compared to others. For example, Eryilmaz and Sandoval Hernandez (2021) found that principals' instructional leadership scale was invariant across subgroups of countries. Researchers deem it necessary to test the measurement invariance before conducting comparative analyses across cultures or countries (Ding et al., 2023; He et al., 2019; Wurster, 2022), specifically for self-report Likert scales due to possible differences in respondents' interpretation of constructs (He et al., 2019). Failing to reach scalar-level invariance indicates possible issues at the item level (Pedrero & Manzi, 2020; Wurster, 2022). The item-level issues might be related to translation, content, or cultural differences in interpretations of certain expressions (Pedrero & Manzi, 2020).

We should also be informed that some scales that are validated and analyzed using a Western perspective might fail to explain the construct in other cultures and contexts especially if the construct has very strong culturally-loaded items and perceptions. In our study, only for Anglo countries and economies sub-group, our results confirmed an excellent model fit for the strict invariance level. This means that the interpretation of the resilience scale by participants was similar among Anglo countries but varied in other sub-groups. In nine other groups, even though the earlier levels of the analysis revealed an acceptable range of results, they did not provide robust results in revealing the testing measurement invariance of the resilience scale used in PISA 2018.

Ungar (2013) states that developing a cross-cultural measure of resilience is highly unlikely since it is bound to the context and environment in which individuals live. Most of the resilience instruments reflect individualistic interpretations of coping with adversities which are based on Western theories. Considering the items of the resilience scale in PISA (i.e. "My belief in myself gets me through hard times"), they represent an individualistic view of the construct which is more congruent with the Western worldview. Hence,

the items were interpreted similarly in Anglo countries in this study. Ungar (2013) further posits that even though resilience mechanisms are complex, introducing some culturally relevant items into the resilience scale might help capture the voices of youth who were raised in collectivist traditions. "A more unbiased and culturally relevant assessment of resilience would import into measures of children of Anglo-European ancestry questions about their capacity to resist cultural hegemony and experience their rights" (Ungar, 2013, p. 261). Therefore, we believe that including more items that reflect a collectivist view of resilience might help improve scale invariance.

Limitations and Future Research Suggestions

Although this study used large-scale international data and an efficient method of multiple-group confirmatory factor analysis (MG-CFA) to test the measurement invariance across countries there are important limitations to consider. First, since the study used self-reported data, it may be subject to common self-reported data biases (Rosenman et al., 2011). However, the main objective of the study was to show the potential drawback of using self-reported data when measuring resilience. It is worthwhile to investigate other self-reported measures of PISA such as well-being.

Second, even though the participating countries and economies were divided into ten subgroups depending on their socio-demographic characteristics the interpretation of self-efficacy might differ across countries within the same group. A more refined cross-country invariance analysis can be conducted to reveal similarities and differences within the same socio-demographic group.

Third, reaching high levels of measurement invariance can be difficult with large groups since invariance violations increase as the sample size increases (Ding et al., 2023; He et al., 2019). In order to deal with the non-invariance issue in the context of ILSAs, a less strict partial invariance (Byrne & van de Vijver, 2010; van de Vijver et al., 2019) or approximate invariance with an alignment method that tolerates trivial differences are recommended (Asparouhov & Muthén, 2014; Ding et al., 2023; Glassow et al., 2021; He et al, 2019; Munck et al., 2018).

Finally, this study used a construct-level measurement invariance analysis. This approach might overlook item-level invariance issues. Therefore, an examination of item-level invariance is recommended (Ding et al., 2023; Schulze & Pohl, 2021; Wurster, 2022). The determination of non-invariant items allows the revision of questionnaires and inquiry into why certain items may have different meanings across countries (Wurster, 2022).

Conclusion

The analysis of our study has some significant methodological contributions to the secondary research studies that benefit from the results of various ILSAs. We argue that the researchers should be informed about the measurement invariance testing before making any causal comparisons among various countries and economies. Our analysis also reinforces the use of MG-CFA when an indicator can be culturally sensitive in different contexts and cultures as reported in some other studies with other scales (e.g., Ding et al., 2023; He et al., 2019).

Specifically, our study sheds more light on the use of self-efficacy as an indicator of resilience in PISA assessment and might provide insights to the researchers who are going to work on cross-country comparisons. Notwithstanding self-efficacy is considered a crucial asset of resilience, cross-cultural studies need to take into account the unique characteristics of the population and the context when using self-efficacy as an indicator of resilience. The introduction of other promotive factors such as resources related to family and society into ILSA resilience scales is recommended for more reliable cross-country comparisons. After setting up the measurement invariance, either in specific counties or sub-categories of countries, analysis with resilience and other variables might provide more accurate conceptualizations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Sibel Kaya https://orcid.org/0000-0001-8417-3627

References

Abolmaali, K., & Mahmudi, R. (2013). The prediction of academic achievement based on resilience and perception of the classroom environment. *Open Science Journal of Education*, *1*(1), 7–12.

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495–508. https://doi.org/10.1080/10705511.2014.919210

Baas, M. I. A. (2020). *The relationship between psychological capital and twenty first century skills in an educational setting* [Master's thesis, University of Twente]. University of Twente Repository. https://purl.utwente.nl/essays/81432

Beauvais, F., & Oetting, E. R. (1999). Drug use, resilience, and the myth of the golden child. In M. D. Glantz, & J. L. Johnson (Eds.), *Resilience and development: Positive life adaptations. Longitudinal research in the social and behavioral sciences: An interdisciplinary series* (pp. 101–107). Springer. https://doi.org/10.1007/0-306-47167-1_5

Benard, B. (1991). *Fostering resiliency in kids: Protective factors in the family, school, and community*. Department of Education. https://eric.ed.gov/?id=ED335781

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guildford Press.

Byrne, B. M., & van de Vijver, F. J. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, *10*(2), 107–132. https://doi.org/10.1080/15305051003637306

Chandelkar, S., & Shetty, V. (2019). Resilience, emotional stability and perceived stress among 10th graders. *Indian Journal of Mental Health*, *6*(2), 203–207.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Dei, G. J. S., Massuca, J., Mclsaac, E. & Zine, J. (1997). *Reconstructing "drop-out": A critical ethnography of the dynamics of Black students' disengagement from school.* University of Toronto Press.

Desa, D. (2014). *Evaluating measurement invariance of TALIS 2013 complex scales: Comparison between continuous and categorical multiple-group confirmatory factor analyses* (Working Papers No. 103). OECD Education. https://doi.org/10.1787/5jz2kbbvlb7k-en

DiCorcia, J. A., & Tronick, E. (2011). Quotidian resilience: Exploring mechanisms that drive resilience from a perspective of everyday stress and coping. *Neuroscience & Biobehavioral Reviews*, *35*(7), 1593–1602. https://doi.org/10.1016/j.neubiorev.2011.04.008

Ding, Y., Yang Hansen, K., & Klapp, A. (2023). Testing measurement invariance of mathematics self-concept and self-efficacy in PISA using MGCFA and the alignment method. *European Journal of Psychology of Education*, *38*, 709–732. https://doi.org/10.1007/s10212-022-00623-y

Eryilmaz, N., Rivera-Gutiérrez, M., & Sandoval Hernández, A. (2020). Should different countries participating in PISA interpret socioeconomic background in

the same way? A measurement invariance approach. *Revista Iberoamericana de Educación*, *84*(1), 109–133. https://rieoei.org/RIE/article/view/3981

Eryilmaz, N., & Sandoval Hernandez, A. (2021). Improving cross-cultural comparability: Does school leadership mean the same in different countries? *Educational Studies*, 1–22. https://doi.org/10.1080/03055698.2021.2013777

Fergus, S., & Zimmerman, M. A. (2005). Adolescent resilience: A framework for understanding healthy development in the face of risk. *Annual Review of Public Health*, *26*, 399–419. https://doi.org/10.1146/annurev.publhealth.26.021304.144357

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, *82*(2), 221–234. https://doi.org/10.1037/0021-9010.82.2.221

Gillespie, B., Chaboyer, W., & Wallis, M. (2007). Development of a theoretically derived model of resilience through concept analysis. *Contemporary Nurse*, *25*(1–2), 124–135. https://doi.org/10.5172/conu.2007.25.1-2.124

Glassow, L. N., Rolfe, V., & Yang Hansen, K. (2021). Assessing the comparability of teacher-related constructs in TIMSS 2015 across 46 education systems: An alignment optimization approach. *Educational Assessment, Evaluation and Accountability*, *33*(1), 105–137. https://doi.org/10.1007/s11092-020-09348-2

GLOBE Foundation. (n.d.). *GLOBE: Global leadership and organizational behavior effectiveness*. Retrieved March 1, 2021, from https://globeproject.com

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Harju-Luukkainen, H., McElvany, N., & Stang, J. (2020). *Monitoring student achievement in the 21st century*. Springer.

Hays-Grudo, J., & Morris, A. S. (2020). *Adverse and protective childhood experiences: A developmental perspective*. American Psychological Association. https://doi.org/10.1037/0000177-000

He, J., Barrera-Pedemonte, F., & Buchholz, J. (2019). Cross-cultural comparability of noncognitive constructs in TIMSS and PISA. *Assessment in Education: Principles, Policy & Practice*, *26*(4), 369–385. https://doi.org/10.1080/0969594X.2018.1469467

He, J., & Van de Vijver, F. J. (2015). The value of keeping an open eye for methodological issues in research on resilience and culture. In L. C. Theron, L. Lienebberg, & M. Ungar (Eds.), *Youth resilience and culture* (pp. 189–201). Springer.

He, J., & Van de Vijver, F. J. R. (2013). Methodological issues in cross-cultural studies in educational psychology. In G. A. D. Liem, & A. B. I. Bernardo (Eds.), *Advancing cross-cultural perspectives on educational psychology: A festschrift for Dennis McInerney* (pp. 39–56). Information Age Publishing.

Henry, C. S., Morris, A. S., & Harrist, A. W. (2015). Family resilience: Moving into the third wave. *Family Relations*, *64*(1), 22–43. https://doi.org/10.1111/fare.12106

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

Isac, M. M., Palmerio, L., & van der Werf, M. P. C. (2019). Indicators of (in) tolerance toward immigrants among European youth: An assessment of measurement invariance in ICCS 2016. *Large-Scale Assessments in Education*, *7*(1), 1–21. https://doi.org/10.1186/s40536-019-0074-5

Jafri, M. H. (2013). A study of the relationship of psychological capital and students' performance. *Business Perspectives and Research*, *1*(2), 9–16. https://doi.org/10.1177/2278533720130202

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*(4), 409–426. https://doi.org/10.1007/BF02291366

Klieme, E. (2016). *TIMSS 2015 and PISA 2015: How are they related on the country level*. https://pisa.dipf.de/de/pdf-ordner/Klieme_TIMSS2015andPISA2015.pdf

Kline, R. B. (2011). *Principles and practices of structural equation modelling* (3rd ed.). Guilford Press.

Kuldas, S., & Foody, M. (2022). Neither resiliency-trait nor resilience-state: Transactional Resiliency/e. *Youth & Society*, *54*(8), 1352–1376. https://doi.org/10.1177/0044118X211029309

Lesko, N. (2001). *Act your age! A cultural construction of adolescence*. Routledge Falmer.

Liu, Y. (2020). Focusing on the practice of distributed leadership: The international evidence from the 2013 TALIS. *Educational Administration Quarterly*, *56*(5), 779–818. https://doi.org/10.1177/0013161X20907128

Luthans, F., Avolio, B. J., Avey, J. B., & Norman, S. M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psychology*, *60*(3), 541–572. https://doi.org/10.1111/j.1744-6570.2007.00083.x

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development*, *71*(3), 543–562. https://doi.org/10.1111/1467-8624.00164

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, *56*(3), 227–238. https://doi.org/10.1037/0003-066X.56.3.227

Masten, A. S. (2013). Global perspectives on resilience in children and youth. *Child Development*, *85*(1), 6–20. https://doi.org/10.1111/cdev.12205

Masten, A. S., & Powell, J. L. (2003). A resilience framework for research, policy, and practice. In S. Luthar (Ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 1–28). Cambridge University Press.

Millsap, R. E. (2012). *Statistical approaches to measurement invariance.* Routledge.

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modelling* (pp. 380–392). Guilford Press.

Munck, I., Barber, C., & Torney-Purta, J. (2018). Measurement invariance in comparing attitudes toward immigrants among youth across Europe in 1999 and 2009:

The alignment method applied to IEA CIVED and ICCS. *Sociological Methods & Research*, *47*(4), 687–728. https://doi.org/10.1177/0049124117729691

Nelis, S. M., Gilleece, L., Fitzgerald, C., & Cosgrove, J. (2021). *Beyond achievement: Home, school and wellbeing findings from PISA 2018 for students in DEIS and non-DEIS schools*. Educational Research Centre.

Novotný, J. S., & Křeménková, L. (2016). The relationship between resilience and academic performance at youth placed at risk. *Československá Psychologie*, *60*(6), 553–566.

Oberski, D. (2014). lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, *57*(1), 1–27. https://doi.org/10.18637/jss.v057.i01

OECD. (2019). *PISA 2018 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/b25efab8-en

Oliveri, M. E., & Ercikan, K. (2011). Do different approaches to examining construct comparability in multilanguage assessments lead to similar conclusions? *Applied Measurement in Education*, *24*(4), 349–366. https://doi.org/10.1080/08957347.2011.607063

Oshio, A., Kaneko, H., Nagamine, S., & Nakaya, M. (2003). Construct validity of the adolescent resilience scale. *Psychological Reports*, *93*, 1217–1222. https://doi.org/10.2466/pr0.2003.93.3f.1217

Paloș, R., Sava, S. L., & Vîrgă, D. (2020). The role of teacher support, students' need satisfaction, and their psychological capital in enhancing students' self-regulated learning. *Studia Psychologica*, *62*(1), 44–57. https://doi.org/10.31577/sp.2020.01.790

Panter-Brick, C., & Eggerman, M. (2012). Understanding culture, resilience, and mental health: The production of hope. In M. Ungar (Ed.), *The social ecology of resilience* (pp. 369–386). Springer.

Pedrero, V., & Manzi, J. (2020). Self-beliefs, engagement and motivation in science and mathematics: Are they universal? *International Journal of Educational Research*, *101*, 101562. https://doi.org/10.1016/j.ijer.2020.101562

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team (2019). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing.

Rosenman, R., Tennekoon, V., & Hill, L. G. (2011). Measuring bias in self-reported data. *International Journal of Behavioural and Healthcare Research*, *2*(4), 320–332. https://doi.org/10.1504/IJBHR.2011.043414

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rutkowski, L., & Rutkowski, D. (2018). Improving the comparability and local usefulness of international assessments: A look back and a way forward. *Scandinavian Journal of Educational Research*, *62*(3), 354–367. https://doi.org/10.1080/00313831.2016.1261044

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Rutkowski, L., & Svetina, D. (2017). Measurement invariance in international surveys: Categorical indicators and fit measure performance. *Applied Measurement in Education*, *30*(1), 39–51. https://doi.org/10.1080/08957347.2016.1243540

Rutter, M. (2012). Resilience as a dynamic concept. *Development and Psychopathology*, *24*(2), 335–344. https://doi.org/10.1017/S0954579412000028

Sacker, A., & Schoon, I. (2007). Educational resilience in later life: Resources and assets in adolescence and return to education after leaving school at age 16. *Social Science Research*, *36*(3), 873–896. https://doi.org/10.1016/j.ssresearch.2006.06.002

Sandoval-Hernandez, A., Rutkowski, D., Matta, T., & Miranda, D. (2019). Back to the drawing board: Can we compare socioeconomic background scales? *Revista de Educacion*, *383*, 37–61. https://doi.org/10.4438/1988-592X-RE-2019-383-400

Schmitt, N., Golubovich, J., & Leong, F. T. (2011). Impact of measurement invariance on construct correlations, mean differences, and relations with external correlates: An illustrative example using Big Five and RIASEC measures. *Assessment, 18*(4), 412–427. https://doi.org/10.1177/1073191110373223

Schulze, D., & Pohl, S. (2021). Finding clusters of measurement invariant items for continuous covariates. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(2), 219–228. https://doi.org/10.1080/10705511.2020.1771186

Solomon, O. (2013). *Exploring the relationship between resilience, perceived stress and academic achievement*. [Unpublished project]. Manchester Metropolitan University.

Steiner-Khamsi, G. (2019). Conclusions: What policy-makers do with PISA. In F. Waldow, & G. Steiner-Khamsi (Eds.) *Understanding PISA's attractiveness critical analyses in comparative policy studies* (pp. 233–248). Bloomsbury Publishing.

Theron, L. C., & Theron, A. M. C. (2013). Positive adjustment to poverty: How family communities encourage resilience in traditional African contexts. *Culture & Psychology*, *19*(3), 391–413. http://doi.org/10.1177/1354067X13489318

Tugade, M., & Fredrickson, B. (2004). Resilient individual use positive emotions to bounce back from negative emotional experiences. *Journal of Personality and Social Psychology*, *86*, 320–333. http://doi.org/10.1037/0022-3514.86.2.320

Ungar, M. (2008). Resilience across cultures. *The British Journal of Social Work*, *38*(2), 218–235. https://doi.org/10.1093/bjsw/bcl343

Ungar, M. (2011). The social ecology of resilience: Addressing contextual and cultural ambiguity of a nascent construct. *American Journal of Orthopsychiatry*, *81*(1), 1–17. https://doi.org/10.1111/j.1939-0025.2010.01067.x

Ungar, M. (2013). Resilience, trauma, context, and culture. *Trauma, Violence, & Abuse*, *14*(3), 255–266. https://doi.org/10.1177/1524838013487805

Ungar, M. (2015). Resilience and culture: The diversity of protective processes and positive adaptation. In L. C. Theron, L. Lienebberg, & M. Ungar (Eds.), *Youth resilience and culture* (pp. 37–48). Springer.

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*(1), 4–70. https://doi.org/10.1177/109442810031002

van de Vijver, F. J. R., Avvisati, F., Davidov, E., Eid, M., Fox, J. -P., Donn'e, N. L., Lek, K., Meuleman, B., Paccagnella, M., & van de Schoot, R. (2019). *Invariance analyses in large-scale studies*. (Working Papers No. 201). OECD Education. https://doi.org/10.1787/254738dd-en

van de Vijver, F. J. R., & Chasiotis, A. (2010). Making methods meet: Mixed designs in crosscultural research. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B. E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 455–473). Wiley.

Wang, M. C., & Gordon, E. W. (Eds.). (1994). *Educational resilience in inner-city America: Challenges and prospects*. Routledge.

Windle, G. (2011). What is resilience? A review and concept analysis. *Reviews in Clinical Gerontology*, *21*(2), 152–169. https://doi.org/10.1017/S0959259810000420

Wright, M. O. D., & Masten, A. S. (2015). Pathways to resilience in context. In L. C. Theron, L. Lienebberg, & M. Ungar (Eds.), *Youth resilience and culture* (pp. 3–22). Springer.

Wurster, S. (2022). Measurement invariance of non-cognitive measures in TIMSS across countries and across time. An application and comparison of multigroup confirmatory factor analysis, Bayesian approximate measurement invariance and alignment optimization approach. *Studies in Educational Evaluation*, *73*, 101143. https://doi.org/10.1016/j.stueduc.2022.101143

Zimmerman, M. A., & Arunkumar, R. (1994). Resiliency research: Implications for schools and policy. *Social Policy Report*, *8*(4), 1–20.

Author Biographies

Sibel Kaya is a Research Fellow at the University of Bedfordshire, Luton, UK.

Nurullah Eryilmaz is a Research Analyst at the International Association for the Evaluation of Educational Achievement, Hamburg, Germany.

Dogan Yuksel is a Research Associate at the Open University, Milton Keynes, UK.