## A translation and validation of the Perceived Political Self-Efficacy (P-PSE) Scale for the use in German samples

Bromme, Laurits; Rothmund, Tobias; Caprara, Gian Vittorio

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

## INTERNATIONAL ADAPTATION OF MEASUREMENT INSTRUMENTS

**Open Access**

# A translation and validation of the Perceived Political Self-Efficacy (P-PSE) Scale for the use in German samples

Laurits Bromme[1*] 📷, Tobias Rothmund[2] and Gian Vittorio Caprara[3]

## Abstract

Caprara et al. (European Journal of Social Psychology 39:1002–1020, 2009) criticized existing measures of internal political efficacy for not taking into account psychological theories of self-efficacy and for the resulting low construct validity. As an alternative, they presented a ten-item measure called *Perceived Political Self-Efficacy (P-PSE) Scale*. Based on social cognitive theory, it adopts a psychological understanding of self-efficacy and captures the phenomenon in a more systematic and complete manner than previous measures of internal efficacy.

We translated the P-PSE scale to German and tested it in a German national quota sample, using quotas for age, gender and education ($N = 1025$). We provided evidence on the scale's construct validity (by testing its correlations towards related constructs) and on its criterion validity (by regressing political participation propensity on the P-PSE score). The scale explained $\Delta R^2 = 26\%$ of people's propensity for political participation over and above sociodemographic variables, and $\Delta R^2 = 12\%$ over and above previously existing measures, demonstrating its incremental value. We also tested cross-cultural measurement invariance towards an Italian sample, establishing configural, as well as partial metric and scalar invariance. In addition, we validated a four-item short version of the scale, which proved to be similarly valid as the full version. We argue, that these two measurement instruments provide a more adequate way of assessing internal political efficacy for research in German-speaking countries.

**Keywords:** Internal political efficacy, Self-efficacy, Social cognitive theory, Political participation, Translation, Construct validity

## Introduction

Political efficacy is defined as "an individual's perceived ability to participate in and influence the political system" (Yeich & Levine, 1994, p. 259). It is usually conceptualized as two-dimensional, with inter-related but distinct dimensions (Balch, 1974): *External* political efficacy refers to the belief that the political system is responsive to citizens' demands (Balch, 1974, p. 24); *internal* political efficacy refers to "an individual's perception of her/his abilities to execute political actions […]" (Sohl, 2014, p. 42). Internal political efficacy has been argued to be an important psychological predictor of political behavior (Bandura, 1997; Campbell, Gurin, & Miller, 1954, p. 187). Empirical research supports this prediction in regard to different forms of political behavior, such as participation in elections (Gallego & Oberski, 2012, p. 437), political protest (Chang & Chyi, 2009), and other forms of political action (e.g., Krampen, 1990; Vecchione & Caprara, 2009).

* Correspondence: bromme@uni-landau.de
[1]University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany
Full list of author information is available at the end of the article

Assumptions about the psychological underpinnings of self-efficacy beliefs can be drawn from the social cognitive theory (SCT) by Bandura (1991). The theory assumes, that people are capable of exercising control over their actions through self-reflection and self-regulation. Within the different mechanisms necessary for the exercise of control, self-efficacy beliefs play a central role: They are judgements of personal capability regarding some specific desirable behavior (Bandura, 1997) and influence a person's motivation, perseverance, performance, and subsequent consequences of the performance (Bandura, 1991).

The conceptual similarity between Bandura's (1997) self-efficacy concept and the concept of internal political efficacy is self-evident. However, both strings of literature have developed rather separate from each other. Only recently, Caprara, Vecchione, Capanna, and Mebane (2009) have developed a self-report measure of perceived political self-efficacy (P-PSE). In contrast to previously existing measures (e.g., Campbell et al., 1954; Niemi, Craig, & Mattei, 1991), it was developed based on the self-efficacy concept of the SCT framework by Bandura (1991, 1997) and follows the main principles for the construction of self-efficacy measures set by Bandura (2006), like the inclusion of different relevant and measurable tasks as benchmarks of successful behavior.

We aimed to extend this line of work by providing a German version of the P-PSE scale. Furthermore, we conducted empirical tests of the validity of the German translation of this scale. Most importantly, we tested whether this scale provides incremental validity over and above existing measures of internal political efficacy.

## Theoretical background

In a review of definitions, Sohl (2014) illustrated how the exact meaning of the concept of internal efficacy differs between studies. She identified three different components that are regularly used to define internal efficacy: (1) the perception "that one can exert influence (affect political outcomes)" (p. 36-37), (2) the "perceived ability to […] execute political actions" (p. 37), and (3) a perception of "understanding politics/the political system" (p. 37). Definitions differ in which component they include. For example, Niemi et al. (1991) focus on beliefs about the "competence to understand and participate effectively" (p. 1407), but do not explicitly mention a successful influence on outcomes as a necessary component of internal efficacy. In contrast, Balch (1974) focuses solely on the perceived availability of "means of influence" (p. 24), without mention of specific abilities or understanding.

We argue, that a focus on the second component (i.e., internal efficacy as a *perceived ability*) is the most useful, because (a) it best separates internal efficacy from related concepts, like external efficacy (e.g., Balch, 1974; Cohen, Vigoda, & Samorly, 2001), political awareness (Zaller, 1992), and political sophistication (e.g., Luskin, 1987); and (b) it matches with the psychological concept of perceived self-efficacy from SCT (Bandura, 1991, 1997).

The SCT framework assumes that people take an agentic role in planning and executing their own behavior, and that self-efficacy beliefs (i.e., beliefs in the personal capability of mastering desirable behavior) play a key role for initiating and executing specific actions (Bandura, 1991, 1997). For any domain of functioning, a person can judge his or her domain-specific self-efficacy. From a perceived self-efficacy perspective, internal political efficacy can thus be understood as the *belief that oneself is capable of mastering the necessary tasks to successfully participate in the political process* (cf. Bandura, 1997).

### Traditional measures of internal efficacy

While the importance of the concept of internal political efficacy is undisputed, there has been controversy on how to measure it appropriately (for an overview, see Morrell, 2003, pp. 591–595). Most researchers use internal efficacy measures based on the National Election Study (NES) scale introduced by Niemi et al. (1991). The four items of the NES scale are "I consider myself to be well qualified to participate in politics", "I feel that I have a pretty good understanding of the important political issues facing our country", "I feel that I could do as good a job in public office as most other people", and "I think that I am better informed about politics and government than most people"[1] (Niemi et al., 1991, p. 1408). Although there is no exact translation of the NES scale for the use in German surveys, two adaptations have been validated in Germany: a three-item scale by Vetter (1997) and a two-item scale by Beierlein, Kemper, Kovaleva, and Rammstedt (2014). Their items closely resemble the ones that have been used to create the NES scale (cf. Craig, Niemi, & Silver, 1990). The correlational patterns towards related variables (e.g., political interest, participation propensity, and sociodemographic variables) indicate that both scales (henceforth "Vetter scale" and "Beierlein scale") measure the same construct as the English NES scale (cf. Arzheimer, 2005; Beierlein et al., 2014).

---

[1]Niemi et al. (1991) suggested the following response categories: "*agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat,* and *disagree strongly*" (p. 1408).

Caprara et al. (2009) identify two conceptual weaknesses in the traditional measures of internal efficacy, which potentially limit their predictive validity. First, although self-efficacy is a psychological construct, with an extensive body of psychological theory and research surrounding the topic (for an extensive review, see Bandura, 1997), the NES scale has not been constructed using the theoretical insight into the psychological nature of the concept. Second, the NES items address a relatively small set of politically relevant skills, with a focus on understanding (rather than participating in) the political process (for a similar critique of the NES scale see Bandura, 1997, pp. 483–484, and Caprara & Vecchione, 2017, p. 287).

### The P-PSE scale as an alternative measure

In response to these limitations, Caprara et al. (2009) constructed the P-PSE scale, which is based on the guidelines that Bandura (2006) created for the construction of domain-specific self-efficacy measures: Its items should be formulated in terms of capability judgements, their content should represent all tasks relevant for what is considered successful behavior, and the items should include varying levels of task difficulty. Applying these principles to the political domain, Caprara et al. (2009) suggested two main sets of skills necessary to successfully participate in a representative democracy: (1) to voice and effectively promote one's own political opinions and (2) to execute control over elected officials. From these skills, they deduced ten items referring to concrete tasks of political participation, for each of which respondents are asked to rate their capability of mastering it (see Table 1). The scale's variety of specific tasks contrasts with the NES scale, which relies on a more narrow understanding of political competence—rather focusing on knowledge and comprehension of politics. The P-PSE scale can therefore be argued to cover the phenomenon of internal efficacy better than its preceding alternatives, and hence, to offer a more content valid measure of the concept (cf. Fontaine, 2005, p. 804).

Although published in English, only an Italian version of the full P-PSE scale has been validated so far (see Table 5 in Appendix for the Italian wording). Using several Italian samples, the authors confirmed the scale's internal consistency, reliability, and criterion validity (Caprara et al., 2009; Vecchione et al., 2014). Most importantly, they provided evidence of incremental validity over the NES scale in predicting various types of political participation behaviors.

In addition, Vecchione et al. (2014) suggested a short version of the P-PSE scale, by selecting a subset of four items (3, 4, 8, and 10), which they argue to

**Table 1** Item wordings of the original P-PSE scale

For each of the following items, please rate how confident you are in your ability to execute the specific action or behavior described.

(1) State your own political opinion openly, even in clearly hostile settings.

(2) Make certain that the political representatives you voted honor their commitments to the electorate.

(3) Promote public initiatives to support political programs that you believe are just.

(4) Maintain personal relationships with representatives of national government authorities.

(5) Play a decisive role in the choice of the leaders of political movements to which you belong, or to which you are near.

(6) Carry out an effective information campaign for the political movement or party with which you concur regarding beliefs and programs.

(7) Actively promote the election of political candidates in which you trust.

(8) Promote effective activities of information and mobilization in your own community (of work, friends, and family), to sustain political programs in which you believe.

(9) Collect a substantial amount of money to sustain the activities of your party.

(10) Use the means you have as a citizen to critically monitor the actions of your political representatives.

Source: Caprara et al. (2009, p. 1007). Participants rate the items on a five-point Likert scale from "not at all" to "completely". The scale-level P-PSE score is the arithmetic mean of the ten single-item scores. The scale can be administered as survey (online or offline) or interview (telephone or face-to-face)

adequately represent the content of the full scale. They conducted studies on Italian, Spanish, and Greek samples confirming the validity of this short scale.

### Translation and methodology

Since the original scale was published in English, we used the English version as source instrument for the translation, thereby ensuring that future translations to other languages can be based on the same source instrument. After one German native speaker translated the scale to German, we reviewed the scale in a three-person team. The team consisted of two Germans and one English native speaker, all of whom are fluent in the other language, and who—as suggested by the Best Practice Guidelines for Cross-Cultural Surveys—unite different levels of discipline expertise (Survey Research Center, 2016, p. 245). Following the guidelines, we aimed to "keep the content of the questions semantically similar; keep the question format similar within the bounds of the target language; [and] retain measurement properties, including the range of response options offered" (Survey Research Center, 2016, pp. 233–234). In

order to achieve these goals, we based our translation on an asking-the-same-questions-and-translation approach (ASQT; Survey Research Center, 2016, p. 234), and tried to stay as close to the content of the original items as possible. As in the original instrument, the translated version asks about respondents' perceived capability to execute different political activities in the introduction. By instruction, each item displays a capability judgment ("I feel capable to…"), which is to be rated on a five-point agreement scale. The translated items are displayed in Table 5 in Appendix, together with the instruction and response categories.

### Measures

In order to validate the translated scale, we conducted an online survey including the translated P-PSE items and several scales of related constructs. The P-PSE items were administered using a five-point Likert agreement scale with only the extreme categories labelled ("completely disagree" and "completely agree").

For validation, we included the German internal efficacy scales by Beierlein et al. (2014; Spearman Brown coefficient = 0.83) and Vetter (1997; McDonald's $\omega = 0.78$), a five-item political interest scale (Otto & Bacherle, 2011; McDonald's $\omega = 0.94$), a three-item external efficacy scale (Vetter, 1997; McDonald's $\omega = 0.74$), a three-item scale of general self-efficacy (Beierlein, Kemper, Kovaleva, & Rammstedt, 2013; McDonald's $\omega = 0.88$), a self-placement of left–right orientation (based on GESIS, 2015), and a list of eleven items asking about the respondents' political participation behavior (e.g., "During the last two years, how often did you actively participate in a political party or movement?"; McDonald's $\omega = 0.83$). All measures are provided in detail in the online supplementary materials.

### Sample

The survey was conducted in October 2016 by the professional sampling agency Respondi (www.respondi.com). The sampling process followed a plan with representative quota for the German adult population regarding age, gender, and formal education. All respondents declared their informed consent before starting the survey and received a financial incentive for their participation by the sampling agency. After exclusion of careless responders (cf. Meade & Craig, 2012; for documentation, see supplementary material) and listwise exclusion of missing values ($N = 64$), a total of $N = 1025$ cases was used for analysis. The sample consisted of 51.7% females vs. 48.3% males. The mean age was 51.7 years ($SD = 16.5$). Of the participants, 36.7% reported low levels of formal education ('Hauptschule' or no degree at all), 30.5% reported

medium levels ('Realschule'), and 32.8% reported high levels ('Abitur' or 'Fachabitur'). The sample distributions closely approximated population parameters, even after the exclusion of careless responders and missing value cases (for a detailed sample description see supplementary materials).

In order to assess cross-cultural invariance of the scale, we used an Italian sample from Caprara et al. (2009, Study 1), which was kindly provided to us by the authors. The data were collected via face-to-face questionnaire in Italy in 2008. The participants were recruited by psychology majors, who conducted the interviews as part of a course assignment (for more details, see Caprara et al., 2009, p. 1006). All respondents participated voluntarily. After listwise exclusion of missing values ($N = 30$), the Italian sample had a total of $N = 1654$ valid cases. Although a convenience sample, distributions were diverse regarding gender, age, and formal education: The sample consisted of 54.4% females vs. 45.6% males. Respondents' mean age ranged from 19 to 89 years ($M = 44.7$; $SD = 17.6$). Of the respondents, 20.7% had concluded elementary or junior high school, 55.6% had concluded high school, and 23.8% had achieved some university degree. A detailed sample description and comparison to population distributions is provided in the online supplementary materials. The Italian P-PSE scale consisted of the ten original items (see Table 5 in Appendix for the Italian wording and response categories) and yielded an internal consistency of McDonald's $\omega = 0.92$.

### Psychometric properties of the translated scale

The German P-PSE scale resulted in a mean score of $M = 2.78$ ($SD = 0.88$). Item–total correlations varied between $r = 0.49$ (item 1) and $r = 0.79$ (items 6 and 7), with a mean correlation of $r = 0.67$ ($SD = 0.11$). Item difficulty ranged from 0.24 (item 9) to 0.62 (item 1), with a mean difficulty of 0.44 ($SD = 0.11$). As suggested by Vecchione et al. (2014), we used the items 3, 4, 8, and 10 as a short version of the P-PSE scale, which resulted in a mean score of $M = 2.90$ ($SD = 1.01$). Other item-level statistics and intercorrelations are reported in the Table 4 in Appendix.

### Objectivity

The translated P-PSE scale contains a written instruction, Likert scale response options, and a simple aggregation rule to obtain the scale-level score (see Table 5 in Appendix). For paper-and-pencil and computer-based questionnaires, these features sufficiently ensure objectivity regarding administration and scoring (Lösel, 1999).

## Reliability

In order to assess reliability, we calculated the internal consistency estimator Omega (ω) described by McDonald (1999), which has been demonstrated to perform better than Cronbach's α (Dunn, Baguley, & Brunsden, 2014). All data analysis was conducted in R version 3.5.1 (R Core Team, 2017) via RStudio version 1.1.456 (RStudio Team, 2015). The internal consistency of the P-PSE scale was ω = 0.91 (95% CI [0.91, 0.92]), which indicates high internal consistency. The four-item short scale resulted in a lower, yet acceptable internal consistency of ω = 0.84 (95% CI [0.83, 0.86]).

## Dimensionality

We tested unidimensionality of the scales using confirmatory factor analysis (CFA) with one latent factor. CFA was conducted using the *lavaan* package (version 0.6-3; Rosseel, 2012) with diagonally weighted least squares estimation due to the items' ordinal level of measurement (Kline, 2016, pp. 257–258). As recommended by Hu and Bentler (1999), we assessed model fit by jointly considering the comparative fit index (CFI; acceptable fit > 0.95) and standardized root mean-square residual (SRMR; acceptable fit < 0.08). Both indices corroborated the hypothesis of a single latent factor for the full scale (CFI = 0.993; SRMR = 0.048), as well as the short scale (CFI = 0.999; SRMR = 0.022; see the upper part of Table 2 for an overview).

## Measurement invariance

The P-PSE full scale has only been validated in Italy so far (although the short scale has also been validated in Spanish and Greek samples). Therefore, we tested cross-cultural invariance of the scale by comparing our German sample ($n_1 = 1025$) to the Italian sample ($n_2 = 1654$). Using multigroup CFAs, we tested for *configural invariance* (same factor structure across samples), followed by *metric invariance* (same factor loadings across samples), *scalar invariance* (same item intercepts across samples), and *residual invariance* (same error variances across samples; for a similar procedure, see Baumert et al., 2014; for an overview of measurement invariance conventions, see Putnick & Bornstein, 2016). We assessed model fit of the configural invariance model using the same criteria as before (Hu & Bentler, 1999). All subsequent models are each nested within its preceding model (e.g., metric invariance within configural invariance) and were therefore assessed in comparison to the preceding model. To judge whether fit differences between nested models are substantial, we used the cutoff criteria by Chen (2007). For large sample sizes, she recommends the use of ΔCFI = 0.01 as main criterion, and ΔRMSEA = 0.015 and ΔSRMR = 0.01 (except for metric invariance, where ΔSRMR = 0.03) as additional criteria.

Using the same specifications as before, we found full configural invariance between the two samples (CFI = 0.996; RMSEA = 0.035; SRMR = 0.036; see the lower part of Table 2). However, full metric invariance could not be established (e.g., ΔCFI = 0.015). We identified Item 2 to differ most strongly in its loading across samples, and—allowing this item's loading to differ freely between samples, as suggested by Vandenberg and Lance (2000, p. 57)—established a model of partial metric invariance. This model met two of

**Table 2** Confirmatory factor analysis (CFA) and measurement invariance (MI) model fit

| Model | $\chi^2$ (df) | CFI | RMSEA [90% CI] | SRMR | Model comparison | | | | | Decision |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Ref. | $\Delta\chi^2$ (Δdf) | ΔCFI | ΔRMSEA | ΔSRMR | |
| CFA 1: Full scale | 103.58* (35) | 0.993 | 0.044 [0.034, 0.054] | 0.048 | - | - | - | - | - | Accept |
| CFA 2: Short scale | 4.34 (2) | 0.999 | 0.034 [0.000, 0.078] | 0.022 | - | - | - | - | - | Accept |
| Invariance | | | | | | | | | | |
| MI 1: configural | 185.20* (70) | 0.996 | 0.035 [0.029, 0.041] | 0.036 | - | - | - | - | - | Accept |
| MI 2: metric | 603.73* (79) | 0.981 | 0.070 [0.065, 0.076] | 0.064 | MI 1 | 418.53* (9) | 0.015 | 0.035 | 0.028 | Reject |
| MI 2a: partial metric | 431.67* (78) | 0.987 | 0.058 [0.053, 0.064] | 0.054 | MI 1 | 246.47* (8) | 0.009 | 0.023 | 0.018 | Accept |
| MI 3: scalar | 1331.43* (87) | 0.954 | 0.103 [0.099, 0.108] | 0.088 | MI 2a | 899.76* (9) | 0.033 | 0.045 | 0.034 | Reject |
| MI 3a: partial scalar | 593.78* (84) | 0.981 | 0.067 [0.062, 0.072] | 0.061 | MI 2a | 162.11* (6) | 0.006 | 0.009 | 0.007 | Accept |
| MI 4: residual | 682.74* (94) | 0.978 | 0.068 [0.064, 0.073] | 0.067 | MI 3a | 88.96* (10) | 0.003 | 0.001 | 0.006 | Accept |

Note. CFA: N = 1025; MI: $n_1$ = 1025; $n_2$ = 1654. We did not consider the $\chi^2$ test for model rejection, because its sensitivity increases with sample size: In large samples—as is the case in this study—even small model–data discrepancies produce significant results (Bollen & Long, 1993; Cheung & Rensvold, 2002)
*CFI* comparative fit index, *RMSEA* Root Mean Square Error of Approximation, *SRMR* Standardized Root Mean-square Residual, *Ref.* reference model
*p < 0.001

our fit criteria (ΔCFI = 0.009; ΔSRMR = 0.018), but exceeded the cut off value of the third one (ΔRMSEA = 0.023). Since ΔCFI is recommended as the main criterion (Chen, 2007, p. 501) and absolute model fit was still good in terms of the Hu and Bentler (1999) criteria, we decided to accept this model of partial metric invariance. Since the majority of item loadings was invariant (Vandenberg & Lance, 2000, p. 38; see also Putnick & Bornstein, 2016), we proceeded in testing scalar invariance. Again, the full scalar invariance model failed our criteria (e.g., ΔCFI = 0.033). Thus, we identified the three items which most strongly diverged in their intercepts (Items 1, 2, and 3) and—allowing these to differ between groups—established a model of partial scalar invariance (ΔCFI = 0.006; ΔRMSEA = 0.009; ΔSRMR = 0.007). Finally, we tested the residual invariance model against the preceding model and found it to fit relatively well (ΔCFI = 0.003; ΔRMSEA = 0.001; ΔSRMR = 0.006).

Summarizing, we can say that the ten P-PSE items load on a single latent factor in both samples (Germany and Italy), and that nine out of ten items do so with equal factor loadings across samples, i.e., the scale shows partial metric invariance. The observed non-invariance of Item 2 ("Make certain that the political representatives you voted honor their commitments to the electorate"), indicates a cross-cultural difference in how much the described task (i.e., monitoring elected representatives) relates to the latent construct of internal efficacy, with slightly lower standardized loadings in Germany ($\lambda_2 = 0.53$) than in Italy ($\lambda_2 = 0.78$). Apparently, German voters

perceive the control of elected officials to be less related to internal efficacy beliefs compared to Italian voters. One possible explanation might be the Italians' experience of relatively frequent snap elections (1994, 1996, and 2008) and frequent changes of their government leader in recent years, which might be perceived as evidence that elected officials are actually controlled by the people. Since only one item displays metric non-invariance, interpretation of the overall P-PSE mean score can be assumed to be unaffected (Steenkamp & Baumgartner, 1998; Steinmetz, 2013). Full scalar invariance was impeded by three out of ten items, with higher standardized intercepts in Germany ($v_1 = 3.07$, $v_2 = 2.90$, $v_3 = 2.36$) than in Italy ($v_1 = 2.09$, $v_2 = 2.05$, $v_3 = 1.80$). Since between-group differences in item intercepts can affect the comparability of observed mean scores (Steinmetz, 2013), we suggest to use latent modelling when group mean comparisons are of the essence, where partial scalar invariance is a sufficient prerequisite (Steenkamp & Baumgartner, 1998; see also Steinmetz, 2013).

## Validity
We assessed construct, criterion and incremental validity of the translated P-PSE scale and its four-item short scale.

### Construct validity
In order to assess construct validity of the translated scale, we postulated a nomological network of theoretically related (convergent validity) and unrelated (discriminant validity) constructs and tested the

**Table 3** Pearson's correlations of the P-PSE scales with theoretically related (convergent validity) and unrelated (discriminant validity) constructs

| Construct | P-PSE full scale | | P-PSE short scale | |
|---|---|---|---|---|
| | r | p value | r | p value |
| Sociodemographic variables | | | | |
| Age | 0.07 | 0.020 | 0.12 | < 0.001 |
| Gender (0 = female, 1 = male) | 0.20 | < 0.001 | 0.19 | < 0.001 |
| High level of formal education | 0.13 | < 0.001 | 0.14 | < 0.001 |
| Vetter scale | 0.60 | < 0.001 | 0.60 | < 0.001 |
| Beierlein scale | 0.61 | < 0.001 | 0.60 | < 0.001 |
| Political interest | 0.59 | < 0.001 | 0.61 | < 0.001 |
| External efficacy | 0.24 | < 0.001 | 0.22 | < 0.001 |
| General self-efficacy | 0.27 | < 0.001 | 0.25 | < 0.001 |
| Left–right orientation | 0.00 | 0.988 | 0.00 | 0.942 |
| Extremism (based on left–right orientation) | 0.15 | < 0.001 | 0.13 | < 0.001 |

Note. Values of gender and level of formal education are point-biseral correlations

correlations between these constructs and the P-PSE score (Hartig, Frey, & Jude, 2008, pp. 148–154). All results are displayed in Table 3.

As expected, we found high correlations with both preexisting internal efficacy scales. There was also a high correlation with political interest, which is in line with previous studies (e.g., Craig et al., 1990, p. 305; Foschi & Lauriola, 2014, p. 350). We found medium-sized correlations with external efficacy and general self-efficacy. Again, this is very plausible, since both constructs are conceptually related to internal efficacy (Balch, 1974; Bandura, 1997). In line with previous findings and theoretical assumptions (e.g. Caprara et al., 2009), the P-PSE score was independent from left–right orientation, but revealed a small correlation with ideological extremity—operationalized as the squared z-standardized left–right score. Regarding sociodemographic variables, men and highly educated people scored higher on the P-PSE scale than women and people with lower levels of formal education, which is the typical pattern of internal efficacy (e.g., Arzheimer, 2005, p. 199). Additionally, we found a small positive correlation with age. Concluding, the correlations towards external criteria reveal the expected pattern for a measure of internal efficacy: high correlations with other internal efficacy measures and political interest, medium-sized correlations with related self-belief variables, and null-correlations with independent constructs. The same pattern emerged when using the four-item short scale (see Table 3).

### Criterion validity
One predominant aspect of internal efficacy is its predictive value regarding political participation behavior (Bandura, 1997; Krampen, 1990; Vecchione & Caprara, 2009). We therefore assessed the scale's criterion validity (Hartig et al., 2008, p. 156) by measuring its relationship towards the propensity to participate in politics. Similar to other researchers in the field (e.g. Kaase, 1999; Peterson, Speer, & McMillan, 2008), we asked about past involvement in eleven different activities of political participation behavior (e.g., stating one's political opinion or signing a political online-petition) and used these items to build an index of political participation propensity (McDonald's $\omega = 0.83$). We used a hierarchical regression model including the control variables age, gender, and education to estimate the scale's criterion validity. As expected, the P-PSE scale explained a substantial amount of variance in respondents' participation propensity over and above the sociodemographic variables in its full ten-item version ($\beta = 0.28$, $p < 0.001$, $\Delta R^2 = 0.26$), and in its four-item short version ($\beta = 0.25$, $p < 0.001$, $\Delta R^2 = 0.26$).

### Incremental validity
In their validation study, Caprara et al. (2009) showed that the original P-PSE scale accounted for unique variance in several indicators of political participation over and above the traditional NES scale. In order to corroborate the incremental value of the scale in the German context, we aimed at replicating this finding with regards to the Vetter and Beierlein scales. We estimated two more hierarchical regression models of political participation propensity. Each model included the before-mentioned control variables and one of the traditional measures before adding the translated P-PSE scale in a second step. The P-PSE scale increased the explained variance by $\Delta R^2 = 0.12$ compared to the Vetter scale and by $\Delta R^2 = 0.13$ compared to the Beierlein scale. The four-item short scale revealed the same incremental value compared to the traditional scales. Detailed results of all regression models are documented in the online supplementary materials.

## Discussion and conclusion
Research on internal political efficacy has not yet come to a consensus about how to measure the construct—many of the previous measures have been criticized for several reasons (e.g., Bandura, 1997, pp. 483–484; Morrell, 2003, p. 595). Largely neglected by scholars studying political efficacy, SCT (Bandura, 1991) has offered a psychological and systematic perspective on self-efficacy beliefs. Based on this theory, Caprara et al. (2009) created the P-PSE scale as a new measure of internal efficacy, which—constructed in terms of capabilities related to relevant participation behavior—offers an arguably more content valid alternative to the established measures of internal efficacy. We translated and validated the scale for the use in German samples. One limitation of our study concerns the labelling of the response categories in terms of *agreement*. Ratings in terms of *strength of confidence* might have been more consistent, and researchers might want to try this as an alternative to our response categories in the future. Nevertheless, analogous to Caprara et al. (2009), the results confirm the reliability and construct validity of the translated scale. Analyses of measurement invariance revealed that the translated scale yields the same factorial structure, as well as partial metric and scalar invariance compared to the original scale by Caprara et al. (2009). Regarding the most important external criterion—political participation propensity—the scale surpasses the established internal efficacy measures, thereby attesting to its potential value for the study of political behavior. In addition, a four-item short version of the scale resulted in similar results—though with a small decrease in internal consistency—and hence offers an economical alternative especially suited for the application in large surveys.

## Appendix

**Table 4** Item-level statistics and intercorrelations of the translated P-PSE scale

| Item | M | SD | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1. | 3.46 | 1.13 | 0.49 | 0.38 | 0.45 | 0.39 | 0.32 | 0.36 | 0.37 | 0.45 | 0.24 | 0.40 |
| 2. | 3.22 | 1.11 | | 0.51 | 0.42 | 0.44 | 0.39 | 0.38 | 0.37 | 0.38 | 0.25 | 0.50 |
| 3. | 2.91 | 1.23 | | | 0.72 | 0.56 | 0.46 | 0.60 | 0.68 | 0.64 | 0.47 | 0.52 |
| 4. | 2.72 | 1.20 | | | | 0.74 | 0.59 | 0.66 | 0.63 | 0.54 | 0.52 | 0.56 |
| 5. | 2.40 | 1.12 | | | | | 0.66 | 0.66 | 0.57 | 0.47 | 0.48 | 0.49 |
| 6. | 2.56 | 1.20 | | | | | | 0.79 | 0.75 | 0.63 | 0.60 | 0.56 |
| 7. | 2.60 | 1.27 | | | | | | | 0.79 | 0.70 | 0.60 | 0.54 |
| 8. | 3.01 | 1.25 | | | | | | | | 0.74 | 0.51 | 0.60 |
| 9. | 1.95 | 1.07 | | | | | | | | | 0.61 | 0.42 |
| 10. | 2.94 | 1.20 | | | | | | | | | | 0.69 |

Note. $N = 1025$. The upper triangle displays intercorrelations. Values in the diagonal are corrected item-total-correlations

**Table 5** German and Italian item wording

| No. | German | Italian |
|------|--------|---------|
| Intro | Nun geht es um die Fähigkeiten, die nötig sind, um politische Handlungen auszuführen. Ich fühle mich in der Lage… | Di seguito sono elencate alcune situazioni nelle quali un cittadino può trovarsi nell'esercizio dei suoi diritti di partecipazione all'attività politica. La preghiamo di valutare quanto lei ritiene di essere capace di affrontare queste diverse situazioni. Risponda, utilizzando la scala riportata di seguito. |
| (1) | Meine eigene politische Meinung offen zu bekunden, auch in einem offensichtlich feindseligen Umfeld. | Dichiarare la sua opinione politica apertamente, anche nei contesti decisamente ostili |
| (2) | Mich zu vergewissern, dass die politischen Repräsentanten, für die ich gestimmt habe, ihre Wahlversprechen halten. | Fare sì che i rappresentanti politici per i quali ha votato onorino gli impegni presi con l'elettorato |
| (3) | Werbung für politische Bewegungen zu machen, die ich gut finde. | Farsi promotore di iniziative pubbliche a sostegno dei programmi politici che lei ritiene giusti |
| (4) | Persönlichen Kontakt mit Abgeordneten oder Mitarbeitern der Regierung aufzunehmen und zu pflegen. | Intrattenere rapporti personali con rappresentanti degli organi di governo nazionale |
| (5) | Die Wahl von Anführern einer politischen Bewegung entscheidend zu beeinflussen. | Svolgere un ruolo decisivo o comunque rilevante nella scelta dei dirigenti dei movimenti politici ai quali si sente vicino o fa parte |
| (6) | Eine effektive Öffentlichkeitskampagne für eine politische Bewegung durchzuführen, mit deren Zielen oder Überzeugungen ich übereinstimme. | Svolgere un'efficace azione di propaganda per il movimento o lo schieramento politico di cui condivide le ispirazioni e i programmi |
| (7) | Aktiv Wahlwerbung für politische Kandidaten zu machen, denen ich vertraue. | Contribuire attivamente all'elezione dei candidati politici in cui ripone la sua fiducia |
| (8) | Freunde oder Bekannte erfolgreich zu informieren und zu mobilisieren, um ein politisches Programm zu unterstützen, von dem ich überzeugt bin. | Promuovere azioni efficaci di informazione e mobilitazione negli ambiti in cui opera (lavoro, amici, famiglia), a sostegno dei programmi politici in cui crede |
| (9) | Eine bedeutsame Menge Geld zu sammeln, um damit eine politische Bewegung zu unterstützen. | Raccogliere consistenti somme di denaro a sostegno delle attività del movimento politico del quale fa parte |
| (10) | Die Mittel zu nutzen, die mir als Bürger zur Verfügung stehen, um das Tun der politischen Vertreter kritisch zu überwachen. | Utilizzare i mezzi che lei possiede come cittadino per vigilare criticamente sull'operato dei suoi rappresentanti politici |
| Scale | 1 = stimme überhaupt nicht zu | 5 = stimme vollkommen zu | 1 = per nulla capace | 5 = del tutto capace |

Note. The scale-level P-PSE score is the arithmetic mean of the ten single-item score

## Supplementary information

## Authors' contributions
LB collected the data in Germany, analyzed all data, and was the primary person responsible in writing the manuscript. TR helped with the interpretation of the data and wrote parts of the theory and results sections. GVC helped in planning the study and collected the Italian data. The authors read and approved the final manuscript.

## Availability of data and materials
The datasets analyzed during the current study are available at https://osf.io/89j7h/ and can be accessed for scientific purposes.

## Ethics approval and consent to participate
Participants consented to their participation in the anonymous survey. Approval by an ethics committee was not necessary.

## Consent for publication
Not applicable.

## Competing interests
The authors declare that they have no competing interests.

## Author details
[1]University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany. [2]Friedrich Schiller University of Jena, Jena, Thuringia, Germany. [3]Sapienza University of Rome, Rome, Italy.

## References
Arzheimer, K. (2005). "Politikverdrossenheit" - eine Frage der Persönlichkeit?: Der Zusammenhang zwischen Persönlichkeitsfaktoren und Verdrossenheitseinstellungen. In S. Schumann (Ed.), *Persönlichkeit: Eine vergessene Größe der empirischen Sozialforschung* (pp. 193–207). Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH.

Balch, G. I. (1974). Multiple indicators in survey research: the concept "sense of political efficacy". *Political Methodology, 1*(2), 1–43.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Greenwich: Information Age Publishing.

Baumert, A., Beierlein, C., Schmitt, M., Kemper, C. J., Kovaleva, A., Liebig, S., & Rammstedt, B. (2014). Measuring four perspectives of justice sensitivity with two items each. *Journal of Personality Assessment, 96*(3), 380–390.

Beierlein, C., Kemper, C. J., Kovaleva, A., & Rammstedt, B. (2013). Kurzskala zur Erfassung allgemeiner Selbstwirksamkeitserwartungen (ASKU). *Methoden, Daten, Analysen, 7*(2), 251–278. https://doi.org/10.12758/mda.2013.014.

Beierlein, C., Kemper, C. J., Kovaleva, A., & Rammstedt, B. (2014). Political Efficacy Kurzskala (PEKS). In D. Danner & A. Glöckner-Rist (Eds.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen* https://doi.org/10.6102/zis34.

Bollen, K. A., & Long, J. S. (1993). Introduction. In K. A. Bollen & J. S. Long (Eds.), *Sage focus editions: vol. 154. Testing structural equation models* (pp. 1–9). Newbury Park: Sage.

Campbell, A., Gurin, G., & Miller, W. E. (1954). *The voter decides*. White Plains: Row, Peterson and Company.

Caprara, G. V., & Vecchione, M. (2017). *Personalizing politics and realizing democracy. Series in political psychology*. Oxford: Oxford University Press.

Caprara, G. V., Vecchione, M., Capanna, C., & Mebane, M. (2009). Perceived political self-efficacy: Theory, assessment, and applications. *European Journal of Social Psychology, 39*(6), 1002–1020. https://doi.org/10.1002/ejsp.604.

Chang, H.-C., & Chyi, H. I. (2009). Voting with their feet: The relationship between political efficacy and protest propensity among Hong Kong residents. *Taiwan Journal of Democracy, 5*(2), 127–144.

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.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: a Multidisciplinary Journal, 9*(2), 233–255.

Cohen, A., Vigoda, E., & Samorly, A. (2001). Analysis of the mediating effect of personal-psychological variables on the relationship between socioeconomic status and political participation: A structural equations framework. *Political Psychology, 22*(4), 727–757.

Craig, S. C., Niemi, R. G., & Silver, G. E. (1990). Political efficacy and trust: A report on the NES pilot study items. *Political Behavior, 12*(3), 289–314. https://doi.org/10.1007/BF00992337.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399–412.

Fontaine, J. R. J. (2005). Equivalence. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 1, pp. 803–813). Amsterdam: Elsevier. https://doi.org/10.1016/B0-12-369398-5/00116-X.

Foschi, R., & Lauriola, M. (2014). Does sociability predict civic involvement and political participation? *Journal of Personality and Social Psychology, 106*(2), 339–357. https://doi.org/10.1037/a0035331.

Gallego, A., & Oberski, D. (2012). Personality and political participation: the mediation hypothesis. *Political Behavior, 34*(3), 425–451. https://doi.org/10.1007/s11109-011-9168-7.

GESIS. (2015). *GESIS panel standard edition: ZA5665 Datenfile version 11.0.0*. Köln. https://doi.org/10.4232/1.12360.

Hartig, J., Frey, A., & Jude, N. (2008). Validität. In H. Moosbrugger & A. Kelava (Eds.), *Springer-Lehrbuch. Testtheorie und Fragebogenkonstruktion* (1st ed., pp. 135–163). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

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.

Kaase, M. (1999). Interpersonal trust, political trust and non-institutionalised political participation in Western Europe. *West European Politics, 22*(3), 1–21. https://doi.org/10.1080/01402389908425313.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford Press.

Krampen, G. (1990). Sequenzanalytische Befunde zur Entwicklung politischer Handlungsorientierungen im Jugendalter. *Zeitschrift f. Entwicklungspsychologie u. Pädagogische Psychologie, 22*(4), 325–340.

Lösel, F. (1999). Persönlichkeitsdaten (Tests). In R. S. Jäger & F. Petermann (Eds.), *Psychologische Diagnostik: Ein Lehrbuch* (4th ed., pp. 362–380). Beltz: PsychologieVerlagsUnion.

Luskin, R. C. (1987). Measuring political sophistication. *American Journal of Political Science, 31*(4), 856–899. https://doi.org/10.2307/2111227.

McDonald, R. P. (1999). *Test theory: a unified treatment*. Mahwah: Lawrence Earlbaum Associates.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*(3), 437–455. https://doi.org/10.1037/a0028085.

Morrell, M. E. (2003). Survey and experimental evidence for a reliable and valid measure of internal political efficacy. *Public Opinion Quarterly, 67*(4), 589–602. https://doi.org/10.1086/378965.

Niemi, R. G., Craig, S. C., & Mattei, F. (1991). Measuring internal political efficacy in the 1988 national election study. *The American Political Science Review, 85*(4), 1407–1413. https://doi.org/10.2307/1963953.

Otto, L., & Bacherle, P. (2011). Politisches Interesse Kurzskala (PIKS) – Entwicklung und Validierung. *Politische Psychologie, 1*(1), 19–35 Retrieved from http://www.psychologie-aktuell.com/popsy.

Peterson, N. A., Speer, P. W., & McMillan, D. W. (2008). Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. *Journal of Community Psychology, 36*(1), 61–73. https://doi.org/10.1002/jcop.20217.

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: DR, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004.

R Core Team. (2017). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing Retrieved from https://www.R-project.org/.

Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36 Retrieved from http://www.jstatsoft.org/v48/i02/.

RStudio Team. (2015). *RStudio. Integrated development for R*. Boston: RStudio, Inc Retrieved from http://www.rstudio.com/.

Sohl, S. (2014). *Youths' political efficacy: Sources, effects and potentials for political equality (dissertation)*. Örebro: Örebro University.

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of consumer research, 25*(1), 78–90.

Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology, 9*(1), 1–12. https://doi.org/10.1027/1614-2241/a000049.

Survey Research Center. (2016). *Guidelines for best practice in cross-cultural surveys*. Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan Retrieved from http://www.ccsg.isr.umich.edu/.

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.

Vecchione, M., & Caprara, G. V. (2009). Personality determinants of political participation: The contribution of traits and self-efficacy beliefs. *Personality and Individual Differences, 46*(4), 487–492. https://doi.org/10.1016/j.paid.2008.11.021.

Vecchione, M., Caprara, G. V., Caprara, M. G., Alessandri, G., Tabernero, C., & González-Castro, J. L. (2014). The Perceived Political Self-Efficacy Scale–Short Form (PPSE-S). *Cross-Cultural Research, 48*(4), 368–384. https://doi.org/10.1177/1069397114523924.

Vetter, A. (1997). *Political efficacy — Reliabilität und Validität*. Wiesbaden: Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-322-95354-4.

Yeich, S., & Levine, R. (1994). Political efficacy: Enhancing the construct and its relationship to mobilization of people. *Journal of Community Psychology, 22*(3), 259–271. https://doi.org/10.1002/1520-6629(199407)22:3<259::AID-JCOP2290220306>3.0.CO;2-H.

Zaller, J. (1992). *The nature and origins of mass opinion*. Cambridge: Cambridge University Press.

## Publisher's Note