Received: 16 March 2021 | Accepted: 16 March 2022

DOI: 10.1111/bjso.12538

the british psychological society
promoting excellence in psychology

ARTICLE

Attitudes towards refugees: Introducing a short three-dimensional scale

Patrick F. Kotzur[1] | **Maria-Therese Friehs**[2] | **Peter Schmidt**[3,4] | **Ulrich Wagner**[5] | **Steffen Pötzschke**[6] | **Bernd Weiß**[6]

[1]Durham University, Durham, UK

[2]University of Koblenz-Landau, Mainz, Germany

[3]University of Gießen and Centre for International Development and Environmental Research (ZEU), Gießen, Germany

[4]University of Mainz, Mainz, Germany

[5]University of Marburg, Marburg, Germany

[6]Leibniz-Institute for Social Sciences (GESIS), Mannheim, Germany

CorrespondencePatrick F. Kotzur, Department of Psychology, Durham University, South Road, Durham DH1 3LE, UK.
Email: patrick.f.kotzur@durham.ac.uk

Abstract

Many researchers subscribe to the three-component conceptualization of attitudes, the idea that attitudes have cognitive, affective, and behavioural (intentional) components. Yet, these components are rarely considered simultaneously in scales, especially those measuring attitudes towards refugees. Moreover, it is debated how these components relate to one another. We present the development and validation of a six-item short-scale to measure attitudes towards refugees based on three surveys (Study 1: N = 330; Study 2a: N = 2,083; Study 2b: N = 2,174). We assessed the performance of this scale with respect to three rivalling attitude conceptualizations (one-factor, three-factor, and second-order factor model). We found that a three-factor or second-order factor conceptualization fitted best to the data. The scale had excellent psychometric properties. We hope that our work stimulates a wave of relevant research on attitudes towards refugees that applies this scale, and contributes to the debate on the conceptualization of attitudes in general.

KEYWORDS
alternative measurement models, attitudes, confirmatory factor analysis, intergroup relations, refugees, scale development

INTRODUCTION

Europe, the United States, and other regions in the world have experienced a tremendous rise in first-time applications for asylum in recent years (Berry et al., 2015). As a result, immigration became one of the most prominent topics in political debates (Green-Pedersen & Otjes, 2017), a circumstance to which the social sciences and psychology, in particular, are still catching up (Wagner & Greipl, 2017). Indisputably, though, refugees have become and continue to be a highly relevant attitude target (e.g., Cowling et al., 2019; Echterhoff et al., 2020; Kotzur & Wagner, 2021), requiring particular scientific attention. When researchers give their attention to this important topic, it is imperative that they can draw on reliable and validated measures to assess attitudes towards refugees as a social group of increasing social, political, and academic relevance, especially in refugee-receiving countries.

Attitudes are commonly defined as evaluations of a target or stimulus (Ajzen, 2001). Although many classical and contemporary conceptualizations of attitudes exist (e.g., Allport, 1954; Brown, 2010; Cuddy et al., 2007; Katz & Stotland, 1959; Rosenberg & Hovland, 1960), the most prominent conceptualizations converge on the postulation of a three-component structure of attitudes, meaning that attitudes consist of a cognitive (thoughts), affective (emotions), and/or conative (behavioural/behavioural intentional) component.[1]

Despite the prominence of the three-component conceptualization, which is often referred to in introductory psychology books, this conceptualization is less often applied in empirical research, including our own. Indeed, we do not know many researchers interested in attitudes towards social groups that have not used the popular feeling thermometer (Nelson, 2008) or its derivatives in their research, capturing overall positivity or negativity towards the attitude target; or the Bogardus scale (Bogardus, 1947) and its derivatives, capturing the felt social distance to outgroup members. Besides disregarding the three-component structure, these popular attitude scales are frequently faced with other measurement-related drawbacks, such as that they consist of one or very few items that are reduced to a single data point using the Guttman scale logic (Stouffer et al., 1950), limiting the assessment of the measures' reliability and validity (Brown, 2015).

Other prominent scales have been developed to explicitly measure cognitive, affective, and behavioural intentional aspects (e.g., Cuddy et al., 2007; Haddock et al., 1993; Pettigrew & Meertens, 1995). However, these scales either do not consider all three attitude components (Haddock et al., 1993), some measurement properties have been questioned (Friehs et al., 2022; Kotzur et al., 2020), or they have not yet been formally validated. Indeed, the quality of measures and the formal validation of scales has rarely been a top priority in leading social psychological journals in the past (Wetzel & Roberts, 2020), disincentivizing authors of existing scales to invest time and effort into such endeavours.

Some of the established scales are relatively long (e.g., Pettigrew & Meertens, 1995). The usage of long scales can imply practical problems, such as burden on participants and monetary costs when included in large-scale surveys. Relatedly, some participants may switch from optimizing to satisficing when they are confronted with longer scales (Krosnick, 1991, 1999). Whereas optimizing describes high engagement with the questions that are posed to the participant, satisficing participants 'interpret each question only superficially and select what they believe *appear* to be a reasonable answer to each question without referring to any internal psychological cues specifically relevant to the attitude' (Krosnick, 1991, p. 215, italics in original). Satisficing is explained by the number of

[1]But see, e.g., Fishbein & Ajzen, 2010, who view cognitive and affective components as defining elements of attitudes and intention as a dependent construct explained by attitudes, or Fishbein (1967), who considers the affective component as the essential element.

used items, endangering data quality (Krosnick, 1991, 1999). Based on these postulations, the first goal of this article has been to develop and validate a short-scale that measures attitudes towards refugees using comprehensive theoretical considerations and state-of-the-art procedures to overcome these limitations.

Although many theorists and researchers subscribe to the idea that attitudes have a cognitive, affective, and conative component, it is heavily debated how these components relate to one another. A prominent conceptualization of (negative) attitudes in Allport's (1954) tradition is put forward by Brown (2010), who advocates no 'firm distinctions between biased attitudes, hostile feelings and discriminatory behaviour' (p. 7), since negative attitudes 'can engage our emotions, as well as finding expression in behaviour' (Brown, 2010, p. 7). Such a conceptualization implies that we should attempt to develop a homogenous scale across all three components, in which all items load on one common attitude factor (one-factor model; see Figure 1a).

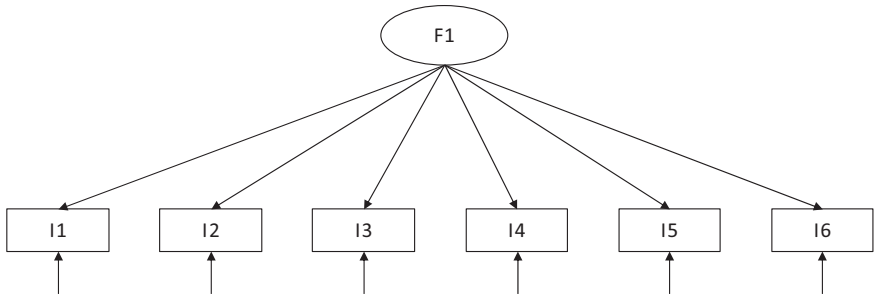
Theorists from other research traditions, however, do draw a firm distinction, suggesting that attitudes consist of three separate components (Cuddy et al., 2007; Eagly & Chaiken, 1993; Fishbein & Ajzen, 2010; Rosenberg & Hovland, 1960; Stephan & Stephan, 2000), which, in some models, may even be causally related to one another. According to such conceptualizations, cognitive, affective, and behavioural intentional components of a scale should be empirically separable into clear sub-factors. This suggests that we should attempt to develop a scale that distinguishes between the three components, and in which items of each attitude component load on a separate subcomponent factor (three-factor model; see Figure 1b).

Finally, if we see both conceptualizations above as two extreme ends of opposing views, there are ways to conceptualize the relationship between the attitude components somewhere in-between. Based on ideas by Eagly and Chaiken (1993), we propose a hierarchical conceptualization of attitudes that integrates both opposing views by acknowledging that the cognitive, affective, and behavioural intentional components are separate attitude dimensions (e.g., Rosenberg & Hovland, 1960), while also acknowledging that these components have common roots (e.g., Allport, 1954; Brown, 2010). This conceptualization takes into account that the three components stem from one common construct, a general attitude towards a target. At the same time, it defines cognitions, affect, and behavioural intentions as related but qualitatively distinct from one another, meaning that, for instance, different aspects of cognitions share more commonalities than, say, aspects of cognitions and affect. The three attitude components are conceptualized to be jointly predicted by a common, overarching, higher-order attitude construct that is theorized to jointly predict people's cognitions, affect, and behaviour towards the outgroup (second-order-factor model; see Figure 1c)[2]. Although Eagly and Chaiken (1993) conceptualized attitudes in this way (e.g., see Eagly & Chaiken, 1993, p. 10, figure 1.2), we are not aware of any formal test of attitudes towards outgroups using a second-order conceptualization of attitudes towards social groups.

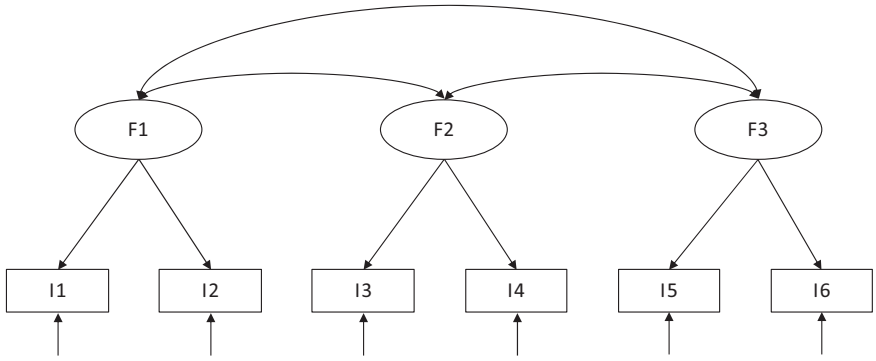
It remains an open question which theoretical conceptualization fits best to the empirical reality when developing a scale to measure attitudes towards refugees. Moreover, researchers have repeatedly called to formally test alternative and/or equivalent model conceptualizations (Bentler & Satorra 2010; Jöreskog, 1993; MacCallum & Austin, 2000), which is especially important in the present context when the relationship between components is highly contested. Thus, our second goal has been to assess the performance of the presented attitudes towards refugees scale with respect to the attitude conceptualizations above (Model 1: one-factor model, Model 2: three-factor model, and/or Model 3: second-order factor model). To our knowledge, this is the first attempt to develop a scale considering cognitions, emotions, and behavioural intentions towards an outgroup that acknowledges different theoretical conceptualizations of how these components should relate to one another.

[2]Although we do acknowledge that other kinds of conceptualizations of attitudes are potentially possible from a methodological perspective (Brown, 2015; Dalege et al., 2016), we limited ourselves here to the three conceptualizations that we deemed most promising from a theoretical perspective and that we felt would be most attractive for attitude researchers who would like to apply the measure in their own research.

(a) *Visual representation of one-factor conceptualization of attitudes*



(b) *Visual representation of three-factor conceptualization of attitudes*



(c) *Visual representation of second-order conceptualization of attitudes*

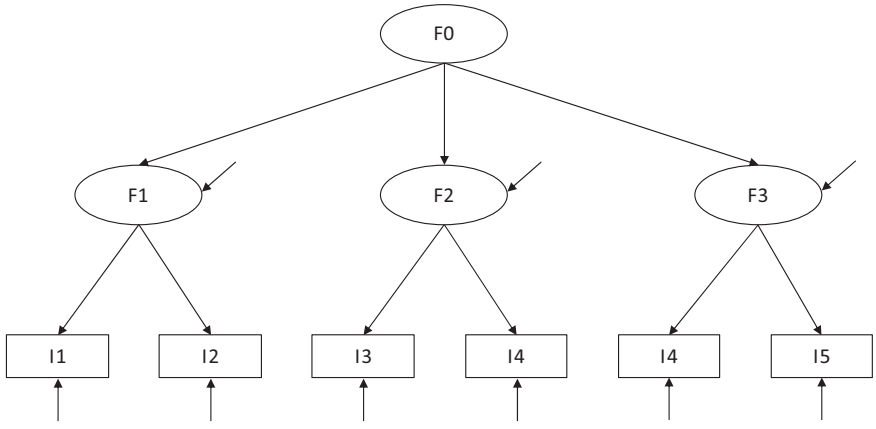


FIGURE 1 (a) Visual representation of one-factor conceptualization of attitudes. *Note*. Cognitive, affective, and behavioural intentional components are not distinguished from one another and instead homogenously contribute to one joint attitude construct (F1 = attitude construct). (b) Visual representation of three-factor conceptualization of attitudes. *Notes*. Cognitive, affective, and behavioural intentional components are firmly distinguished from one another. Although correlated, these components form distinct subfactors (F1 = cognitive component, F2 = emotional component, F3 = behavioral intentional component). (c) Visual representation of second-order conceptualization of attitudes. *Notes*. Although cognitive, affective, and behavioural intentional components are qualitatively distinct from one another, all components contribute to a higher-order, joint attitude construct (F0 = higher order attitude construct; F1 = cognitive component, F2 = emotional component, F3 = behavioral intentional component). Apart from the higher-order attitude factor, there are no interrelations between the cognitive, affective, and behavioural intentional attitude components.

The present research

Many researchers argue that cognitions, affect, and behavioural intentions are facets of attitudes, yet current scales either focus on specific attitude components or suffer other kinds of limitations that question their validity or reliability. With the present research, we aim to develop, empirically test, and independently validate a short-scale that explicitly considers the three components of attitudes towards refugees, a group of increasing relevance in many parts of the world. Despite the agreement among many scholars about the composition of attitudes, the relationship between attitude components is debated, allowing for at least three theoretical models differing in the formalization of the relationship between their components. Consequently, we pursue the following sub-goals aiming at developing and validating a new short-scale across three studies, while acknowledging and testing different theoretical conceptualizations how these components relate to one another:

a. Determining which theoretical model of attitudes is best captured empirically by the factorial structure of the short-scale (Model 1: one-factor model, Model 2: three-factor model, and/or Model 3: second-order factor model). This is to see which formalization of the theoretical relationship between attitude components fits best to the empirical reality (see Study 1);
b. Assessing the scale's internal consistency (ω) and τ-equivalence. The former is to assess reliability, the latter a prerequisite to allow for the application of the resulting scale on the level of observed values by simply computing unweighted mean values or sum scores (e.g., Brown, 2015; see Study 1, Study 2a, Study 2b);
c. Replicating the factorial structure using independent samples. This is to cross-validate the factorial structure (see Study 2a, Study 2b);
d. Testing for measurement equivalence over time. This is to ensure that the scale is equally conceptualized over time, and that correlational, regression-based, or mean-value-based analyses over time are valid and meaningful (e.g., Vandenberg & Lance, 2000; see Study 2b);
e. Testing the convergent/criterion and discriminant validity of the short-scale. This is to see whether our scale is distinguishable enough from other measures, while also ensuring that it relates to other measures as can be expected based on the literature (see Study 2b).

We conducted our research in Germany, a country that has seen a large rise in first-time applications for asylum (Bundesamt für Migration und Flüchtlinge, 2017) and where migration and refugees are topics in public debates (Green-Pedersen & Otjes, 2017). Thus, refugees are a salient and relevant attitude target, which provides the ideal setting for our research.

Item generation

The items for the scale were generated using expert discussions, in which social psychologists, political scientists, sociologists, and scale development experts knowledgeable in relevant theories and conceptualizations of attitudes and in the assessment of scale quality participated. In these discussions, we drew inspiration from the attitude literature, existing scales, and media contents. Further information on the expert discussions can be found in the Supporting Information. Based on the rationale elaborated on above, we aimed for two items per attitude component. Items on which our expert discussions converged are summarized in Table 1.

From all possible *cognitions* that are plausible to include in an attitude towards refugees scale, one particularly important domain is threat. Threat is the expectation that something aversive is going to happen (Fritsche et al., 2011). These expectations can be linked to social groups (Cottrell & Neuberg, 2005; Stephan & Stephan, 2000). As such, threat is a cognitive appraisal (Stephan & Renfro, 2002; Landmann et al., 2019). Researchers oftentimes distinguish between symbolic and realistic threats. Symbolic threats are threats to the ingroup's values and worldviews, whereas realistic threats are

TABLE 1 Items generated in expert discussions to tap into cognitive, affective, and behavioural intentional components of attitudes towards refugees

No.	Items	Scale anchors
Cognitive component		
I1	Stellen Flüchtlinge eher eine Bedrohung oder eine Bereicherung für die Werte in Deutschland dar? *Do refugees rather threaten or enrich the values in [country]?*	[1 – Bedrohung; 5 – Bereicherung] *[1 – threaten; 5 – enrich]*
I2	Stellen Flüchtlinge eher eine Bedrohung oder eine Bereicherung für den Wohlstand in Deutschland dar? *Do refugees rather threaten or enrich the prosperity in [country]?*	[1 – Bedrohung; 5 – Bereicherung] *[1 – threaten; 5 – enrich]*
Affective component		
I3	Wie stark fühlen Sie mit Flüchtlingen mit? *How strongly do you sympathize with refugees?*	[1 – überhaupt nicht; 5 – sehr stark] *[1 – not at all; 5 – very strongly]*
I4	Wie sympathisch sind Ihnen Flüchtlinge? *How likeable are refugees to you?*	[1 überhaupt nicht – 5 sehr] *[1 – not at all; 5 – very much]*
Behavioural intentional component		
I5	Ich kann mir vorstellen mit Flüchtlingen zusammenzuarbeiten. Bitte denken Sie bei Ihrer Antwort an jegliche Art von Zusammenarbeit, z.B. beruflich, ehrenamtlich, in der Freizeit, wie z.B. in Vereinen. *I can imagine myself collaborating with refugees. When answering, please think of any type of collaboration, e.g., in professional and voluntary settings, in your leisure time, e.g., in clubs.*	[1 – trifft überhaupt nicht zu; 5 – trifft vollkommen zu] *[1 – strongly disagree; 5 – strongly agree]*
I6	Ich hätte nichts gegen Flüchtlinge in meinem Wohnumfeld. *I would not mind refugees in my residential area.*	[1 – trifft überhaupt nicht zu; 5 – trifft vollkommen zu] *[1 – strongly disagree; 5 – strongly agree]*

Note: I = item. Non-italic represents original formulation in German, italic represents English translation using the back-translation procedure (Brislin, 1990).

threats to the ingroup's resources and economic power (Stephan & Renfro, 2002). Decades of research established these threats as a cornerstone of intergroup research (Riek et al., 2006), whereas media portrayals of refugees endangering refugee-receiving societies (Couliaraki & Stolic, 2017) stress the need to cover this aspect in the current scale. Moreover, recent research showcases the relevance of threats associated with refugees in contemporary societies (Landmann et al., 2019). Thus, we included items that tapped into a general *cognitive* assessment of threat for society associated with refugees; one item into symbolic threat (threat of values), and one into realistic threat (threat of wealth).

From all possible *emotions* that are plausible to include, sympathy and likeability are particularly important in this context. Sympathy is an emotion that comprises both compassion and sadness (Cuddy et al., 2007), and involves 'concern for another in light of apparent threats to her well-being or good' (Darwall, 1998, p. 275). According to the UN refugee agency, refugees are people who have fled threats to their well-being and good, including wars, violence, conflicts, or persecution. Thus, refugees are by definition exposed to threats to their well-being or good, making sympathy towards them an emotion of utmost relevance. This is supported by prior literature (Kotzur et al., 2017, 2019). Moreover, sympathy plays a vital role in contemporary models of social perception (Cuddy et al., 2007). Lastly, European media frequently cover apparent threats to refugees' well-being (Geogriou & Zaboroski, 2017), which stresses the need to cover the corresponding emotion in the current scale.

Many researchers conceive likeability as a central element of an attitude. As it captures the general favourability with which a perceiver evaluates an attitude target, likeability can be seen as the very essence of an attitude (Eagly & Chaiken, 1993). One-item scales, like the feeling thermometer (Nelson, 2008), capitalize on this. Meta-analyses on attitudes and its determinants show that much research has been

conducted to investigate to what extent attitude targets are perceived as likeable (Paluck et al., 2021). Finally, contemporary theories build on the idea that likeability is a core aspect of social perception (Abele et al., 2021). Thus, we included items that tapped into two *emotional* assessments associated with refugees; one item into sympathy, and one into likeability.

From all possible *behavioural intentions* that are plausible to include, we incorporated approach (versus avoidance) intentions in the housing setting and settings requiring cooperation in professional and leisure contexts. Approach versus avoidance intentions are frequently used to capture the behavioural intentional component of the favourability of the evaluation of social groups (Cuddy et al., 2007). They are fundamental to the study of human behaviour (Elliot & Thrash, 2002), have been studied since the inception of psychology (James, 1890), and are paramount to theorizing in this domain to the present day (Cuddy et al., 2007).

Recent work distinguishes between different intensities or effort of intentions, which vary on an active versus passive dimension (Cuddy et al., 2007). Whereas active intentions entail directed efforts to overtly affect the target group, passive intentions affect the group less directly and openly. The contexts of housing and the context of collaboration in work and other settings have historically been of high relevance in the study of such intentions. Already the classic Bogardus scale (1947) asked people to what extent people would be accepting outgroup members in these contexts, and many studies continued to study these aspects (e.g, Bohrer et al., 2019; Veit & Yemane, 2020). Whereas the acceptance of refugees in one's residential area maps on the passive approach versus avoidance facet (passive non-segregation/segregation), collaborations with refugees in work or recreational settings is an active approach versus avoidance facet (active non-collaboration/collaboration). Consequently, we dedicated one item each to these *behavioural intentions*.

STUDY 1: PRETEST

We conducted a pretest to address research sub-goal A, to examine the underlying factorial structure of the short-scale, and research sub-goal B, to test the internal consistency and τ-equivalence. We omitted exploratory factor analysis, as we had theoretically based, precise competing expectations concerning the factorial structure (Brown, 2015). Instead, we examined the data using confirmatory factor analyses and model-fit comparisons between the models to determine the most suitable factorial structure.

Methods

Sample description

Boomsma (1982, 1985) recommended to base latent variable modelling on samples consisting of at least 100 to 200 individuals. We recruited $N = 330$ participants ($M_{age} = 47.7$, $SD_{age} = 20.31$, female = 52.7%, male = 47.0%, 0.3% missing; 50% with a lower secondary degree or below, 49.4% with a upper secondary degree or above, 0.6% other[3]; 71.2% active in workforce or education; 90% without migration background[4]) via a professional recruiting agency in December 2016 who completed a short online survey containing the six attitude items in return for a small monetary incentive.

[3]Lower secondary level corresponds to a Realschule (i.e., year 10) degree or below, upper secondary level to a Gymnasium (i.e., A-levels/year 12/13) degree, (Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, 2015, pp. 5–6).

[4]People without migration background denotes people who themselves *and* whose parents are non-foreign nationals.

Analysis

We used Mplus Version 8.0 or above (Muthén & Muthén, 1998–2017) in all analyses, using the robust maximum likelihood estimator (MLR) to account for multivariate non-normality. We transformed our a-priori theoretical assumptions about the possible relationships between the attitude components into three measurement models (Figure 1a-c), examined each of them for goodness of fit, and subsequently compared the model-fit of the models.

Model 1 reflected the formalization of the assumptions of a one-factor model, that is, that one factor (F1) is sufficient to explain the covariances between all six items; Model 2 formalized the postulation that three separate subcomponent-specific latent variables (F1/F2/F3) are more adequate to explain the covariances between the six items; Model 3 reflected the formalization of a hierarchical or second-order factor model, where the second-order latent variable (F0) represents the general attitude towards refugees, which explains the covariances between the first-order latent subcomponent-specific variables (F1/F2/F3).

In all models, no covariations between indicator residuals were assumed. Factors of the same order were allowed to covary in absence of a higher-order factor (Model 2), as they reflected attitudes towards the same attitude target.

We examined the resulting models for standardized factor loadings of $|\lambda| \geq 0.5$ (implying that at least 25% of the indicators' variance was explained by the attitude factor; Brown, 2015), factor correlations of $|r| < .80$ (as higher factor correlations imply high conceptual overlap between constructs, indicating the possible existence of a more parsimonious solution; Brown, 2015), and the absence of parameter estimates with out-of-range values (Heywood cases). We determined model-fit to be adequate if all criteria formulated by Schermelleh-Engel et al. (2003) were met: $\chi^2/df \leq 3$; root mean standard error of approximation (RMSEA) ≤ .08; standardized root mean square residual (SRMR) ≤ .10; comparative fit index (CFI) ≥ .95. The one-factor model can be regarded as a more constrained version of the three-factor model (Model 2); therefore, the model-fit of these models can be compared directly. However, the fit of the higher-order factor model with three first-order factors (Model 3) is empirically indistinguishable from the fit of a three-factor model with correlated first-order factors (Model 2) in the absence of additional model constraints. They are therefore equivalent (Brown, 2015). Since the distinction between the Models 2 and 3 has important and meaningful theoretical implications, as well as implications for the potential use of the scale, we followed Brown's (2015) recommendation to include and report results for the second-order factor model in our analyses and comparisons regardless of the equivalence. In all cases of model comparison, we used the Akaike Information Criterion (AIC) as a criterion, which accounts for model complexity (favouring parsimonious models) when evaluating the overall model-fit of the competing models (Brown, 2015). A lower AIC value indicates better model-fit.

For all models indicating acceptable model-fit, we report the internal consistency, meaning that we computed ωs for the latent factors to test the measure's reliability, and τ-equivalence tests, meaning that we restricted the factor loadings to be equal for all indicators per factor to examine whether we could assume that the indicators relate equally to the latent factor (Brown, 2015). We applied the Satorra-Bentler scaled χ^2-difference test adapted for the use of the robust maximum likelihood estimator (MLR) to test for τ-equivalence (Bentler & Satorra, 2010). The Satorra-Bentler scaled χ^2-difference test examines whether a more restricted model (i.e., a model in which τ-equivalence is assumed) presents a non-substantially worse model-fit than a more liberal model (i.e., a model in which factor loadings are freed). Consequently, a non-significant χ^2-difference test value would indicate an equally good fit of the more restricted model, while a significant χ^2-difference test value would indicate that the restricted model shows significantly worse goodness of fit. τ-equivalence would allow for the application of the resulting scale on the level of observed values by computing mean or sum score values, while non-equivalence would advise for weighting each indicator when computing scales on the level of observed values. In case of τ-non-equivalence, we assessed the magnitude of the bias when using unweighted manifest mean scores by correlating unweighted mean scores with weighted factor scores. The larger the correlation, the lower the bias (Bobko et al., 2007).

TABLE 2 Study 1: Goodness of fit information of the three competing models.

Model	AIC	χ^2_{cor}	*df*	*p*	χ^2_{cor}/df	RMSEA	CFI	SRMR
Model 1: One-factor model	4825.156	96.287	9	<.001	10.699	.171	0.928	.039
Model 2: Three-factor model	4723.626	6.660	6	.353	1.110	.018	0.999	.008
Model 3: Second-order factor model	4723.626	6.661	6	.353	1.110	.018	0.999	.008

Abbreviations: AIC, Akaike Information Criterion; CFI, comparative fit index; df, degrees of freedom; *p*, probability value; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

Results

Descriptive information is shown in Table S1. The model-fit information of the three competing models is given in Table 2. The standardized factor loadings, residual variances, factor correlations, and internal consistency indicators are presented in Table S2. Results of τ-equivalence testing are depicted in Table S3, and correlation coefficients between unweighted observed scale mean values and factor are shown in Table S4.

In all three models, all standardized factor loadings were above $|\lambda| > 0.5$, indicating that a large portion of indicator variance can be explained by the specified factors. The second-order model produced an extraordinary, though plausible parameter (estimated standardized factor loading of the second-order factor on the emotional component factor was $\lambda = 1.00$)[5]. For the three-factor model, the factor correlations of the cognitive, emotional, and behavioural intentional factor all exceeded $|r| < .80$.

The one-factor model did not show acceptable model-fit. Additionally, the AIC value and all other criteria of model-fit were less favourable in the one-factor model compared to the competing models. The three-factor model and the second-order model both demonstrated acceptable and – apart from small deviations due to rounding – identical fit parameters. Thus, we subsequently examined the internal consistency and τ-equivalence of the models with superior fit.

We found high internal consistencies of all factors in the three-factor model/first-order factors of the second-order model (ωs > .82), and the second-order factor ($\omega = .96$). When we tested for τ-equivalence, significant Satorra-Bentler-scaled χ^2-differences suggested equivalence only for the emotional component factor in the three-factor model, and consequently, also for the first-order factors in the second-order model. Correlation coefficients between the unadjusted mean scores and corresponding factor scores were $r = .948–.989$.

Discussion

We found significant and high standardized factor loadings for all indicators in all three competing models. Overall, the one-factor model fitted less well to the data, whereas the three-factor model and the second-order model fitted well. This indicates that the factorial structure of the indicators is best represented by explicitly modelling the three subcomponents of attitudes. Two of the three factors were not τ-equivalent, suggesting that weighted observed mean scores are advisable, although the bias of using unweighted scores was very small. Moreover, internal consistencies of all components were high, indicating that the items are apt to assess subcomponents of attitudes towards refugees reliably. When

[5]$\lambda > 1.00$ would be an out-of-range parameter (i.e. Heywood case), which sometimes occur when samples are small (Kolenikov & Bollen, 2012), like in this pretest. However, since $\lambda = 1.00$, and thus not out-of-range, model results are trustworthy.

considering the second-order factor model, the subcomponents formed a highly reliable, τ-equivalent superordinate attitude measure.

As stated previously, the findings did not allow us to decide for or against either the three-factor model or the second-order model, as they were equivalent (Brown, 2015). However, the need to explicitly model subcomponents as well as very high factor correlations in the three-factor model indicating high conceptual overlap between the three components give reason to believe that a second-order factor model may be the best representation of the data.

It remains to be seen whether the results of Study 1 were unique to the sample on which our analyses were based, or replicable. Moreover, a borderline-range parameter estimate in the second-order model calls for replications in larger samples. Thus, in Study 2a, one of our aims was to replicate the factorial structure in a large, independent, cross-sectional data set.

STUDY 2A: CROSS-SECTIONAL LARGE-SCALE APPLICATION

We conducted Study 2a to address research sub-goal C, to replicate the factorial structure in a larger, independent sample to cross-validate the factorial structure, and to revisit research sub-goal B, to test the internal consistency and τ-equivalence.

To pursue these questions, we included our items in the GESIS-Panel, a probability-based mixed-mode access panel of German-speaking adults with permanent residence in Germany. Detailed information on the panel are provided in Bosnjak et al. (2018).

Methods

Sample description

Overall, our scale was administered in the GESIS-Panel in two waves in 2017; in wave EA to the entire sample (N = 4,521, collected February to April 2017) and wave EC to a random half split (N = 2,183, collected June to August 2017). For study 2a, we used data from those participants of wave EA who were presented the items of our scale in the following wave (N = 2,083, M_{age} = 51.6, SD_{age} = 14.8, female = 52.2%; 41.2% with a lower secondary degree or less, 47.4% with a upper secondary degree or above, 11.4% with other degrees; 70.6% active in workforce or education; 9.7% of the respondents were born abroad and 8.4% have at least one foreign-born parent).

Analysis

We replicated all procedures explained in Study 1, applying the same analyses and cut-off criteria.

Results

Descriptive statistics can be found in Table S1. Model-fit information of all three competing models are provided in Table 3. Standardized factor loadings, residual variances, factor correlations, and internal consistency indicators are depicted in Table S5. Results of τ-equivalence testing are depicted in Table S6. Correlation coefficients between unweighted observed scale mean values and factor scores are shown in Table S7.

All models converged without computational issues. Replicating the basic findings of Study 1, we found high standardized factor loadings ($|\lambda| > .5$) in all three competing models, very high ($|r| > .80$) latent factor correlations in the three-factor model (exception: $r_{F1\,F3} = .796$), and comparable results concerning model-fit. The model-fit of the one-factor model was not acceptable; the three-factor model and the second-order

TABLE 3 Study 2a: Goodness of fit information of the three competing models

Model	AIC	χ^2_{cor}	*df*	*p*	χ^2_{cor}/df	RMSEA	CFI	SRMR
Model 1: One-factor model	28618.419	546.061	9	<.001	60.673	.169	0.905	.042
Model 2: Three-factor model	27884.791	17.452	6	.0078	2.909	.030	0.998	.008
Model 3: Second-order factor model	27884.791	17.452	6	.0078	2.909	.030	0.998	.008

Abbreviations: AIC, Akaike Information Criterion; CFI, comparative fit index; df, degrees of freedom; *p*, probability value; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; χ^2_{cor}, Corrected chi-square.

model showed identical acceptable model-fit and equal expected variance covariance matrices, indicating model equivalence. This led us to examine the internal consistency and τ-equivalence of Models 2 and 3.

Internal consistencies were high for all factors in the three-factor model and the first-order factors of the second-order model (ωs > .85), as well as for the second-order factor (ω = .95). Testing for τ-equivalence of the three-factor model, all Satorra-Bentler-scaled χ^2-differences were significant, indicating non-equivalence for all factors, and consequently, also for the first-order factors in the second-order model. Correlation coefficients between the unadjusted mean scores and corresponding factor scores ranged between r = .968 and .995. τ-equivalence of the second-order factor in the respective model could be established.

Discussion

This replication of Study 1 yielded similar results based on a large independent sample. Again, all indicators showed high and significant standardized factor loadings. Based on model-fit information, we discarded the one-factor model as the other models provide an better representation of the empirical relationships between the items. Consequently, this model was not considered in further analyses. The two remaining models showed equally good model-fit and other psychometric properties, including excellent internal consistencies, suggesting that the items measured attitudes towards refugees reliably. None of the three factors of the three-factor model, and by implication the first-order factors of the second-order factor model, were τ-equivalent. However, the bias for using unadjusted observed mean scores remained small again. Given the high factor correlations in the three-factor model and the τ-equivalence of the factor loadings from the global second-order factor on the first-order factors, the second-order model might be preferred from a theoretical and methodological viewpoint.

Despite the scale's excellent performance on the examined dimensions thus far, it remains to be seen whether the scale can be used in longitudinal research, as well as whether the scale has convergent, discriminant, and criterion validity. We conducted Study 2b to address these points.

STUDY 2B: LONGITUDINAL LARGE-SCALE APPLICATION

We conducted Study 2b to readdress research sub-goal B, to test the internal consistency and τ-equivalence, and research sub-goal C, to replicate the factorial structure in a larger, independent sample to cross-validate the factorial structure. Moreover, we intended to address research sub-goal D, to test measurement equivalence over time, and research sub-goal E, to examine the convergent/criterion and discriminant validity[6] of the short-scale. Given the performance of the one-factor

[6]Since we used cross-sectional data to address this research question, we did not distinguish between convergent and criterion validity in the present study.

model in previous analyses, we only considered the three-factor and second-order factor model in this study.

Methods

Sample description

We used data of those participants of the GESIS-Panel who were presented the items of our scale in both wave EA (N = 2,086) and EC (N = 2,167; in contrast to study 2a, where only those participants were included that were presented our items at wave EA). Across both waves, we included data of N = 2,174 respondents in our analyses, M_{age} = 52.4, SD_{age} = 14.5, female = 49.8%; 41.2% with a lower secondary degree or below, 47.1% with a upper secondary degree or above, 11.7% with other degrees; 69.1% active in workforce or education; 8.8% of the respondents were born abroad and an additional 9.9% have at least one foreign born parent).

Analysis

We used the same procedures as in our previous studies, except that we made use of the 2-wave panel design to fit our models longitudinally to address research sub-goals B-D. To account for item-specific variance (specificity), measurement errors of the equally worded items were correlated across waves. The same applied to the disturbances of the equivalent subfactors in the second-order factor model. Additionally, we examined longitudinal measurement equivalence to assess whether our proposed scale 'measures the same concept in the same way' over time (Davidov et al., 2014). This psychometric property is an important precondition for the valid, unambiguous interpretation of longitudinal processes, regression-based analyses, or mean value comparisons (Horn & McArdle, 1992). Various levels of measurement equivalence exist and are usually tested in a hierarchical stepwise process of increasing restrictions in a longitudinal confirmatory factor analysis. The least restrictive model is configural equivalence, which restricts the factor-indicator relations (i.e., the number of factors, the pattern by which items load onto each factor, and the existence of covariances between factors) to be equal over time, thus ensuring that the scale is equally conceptualized over time (Vandenberg & Lance, 2000). Metric equivalence can be established by restricting factor loadings of similar indicators to be equal over time, thus testing the equality of scaling units across groups. Metric measurement equivalence assures valid and meaningful correlational or regression-based analyses over time. Additionally, scalar measurement equivalence can be introduced by restricting indicator intercepts of similar indicators to be equal over time, thus ensuring that no systematic response bias is given across groups and allowing for valid and meaningful comparison of (latent) mean values.

We tested for metric and scalar measurement equivalence of the three-factor (Model 2) and second-order factor models (Model 3) by introducing restrictions of equality first for factor loadings of same-worded indicators/factors over time (i.e., metric equivalence), and second by restricting similar indicator intercepts to be equal over time (i.e., scalar equivalence). For the evaluation of measurement equivalence, we used Chen's (2007) criteria. To test for measurement equivalence in the second-order factor model, we followed Rudnev et al. (2018).

Whereas all previously described analyses are based on a two-wave panel, we examined convergent/criterion and discriminant validity of the short-scale based on data of Wave 1 only. We did so by correlating the three attitude subfactors (Model 2) and the second-order attitude factor (Model 3) with intergroup constructs that were included in the GESIS-Panel. Said constructs should be distinct from, yet theoretically correlated moderately to highly with attitudes towards refugees in a hypothesized direction to evidence convergent/criterion validity (expected $|rs|$ = .30–.80; Cohen, 1988; Hemphill,

2003; Brown, 2015). We also correlated the scale with intergroup constructs specific to another group expected to be related to attitudes towards refugees to a smaller extent than corresponding refugee-specific constructs to evidence discriminant validity. Lastly, we correlated non-intergroup constructs expected to be related to attitudes towards refugees to a small extent, if any, with the factors to provide further evidence for discriminant validity (expected $|rs| < .30$; Cohen, 1988; Hemphill, 2003). To do so, we took advantage of a GESIS-Panel submodule by Wagner et al. (2018, Wave EB). In this module, a subgroup of GESIS-Panel participants provided their assessment of refugee-related intergroup constructs that perfectly fit the purpose of evidencing convergent/criterion validity. Thus, we based some of our analyses on the subsample that was assigned to this submodule. Constructs to test convergent/criterion validity are: Positive contact with refugees (Wagner et al., 2002), negative contact with refugees (Wagner et al., 2002), diversity beliefs (Asbrock et al., 2011), allophilia (Pittinsky et al., 2011), identification with national ingroup (i.e., Germans; Haddock et al., 1993), and fraternal relative deprivation (refugees vs. Germans; Pettigrew et al., 2008).

A further subgroup of GESIS Panel-participants answered intergroup items targeted at Sinti and Roma which fit the purpose of evidencing discriminant validity. Attitudes towards refugees should be associated with them less than with the corresponding refugee-specific constructs. These constructs are: Positive contact with Sinti and Roma (Wagner et al., 2002), negative contact with Sinti and Roma (Wagner et al., 2002), and fraternal relative deprivation (Sinti and Roma vs. Germans; Pettigrew et al., 2008).

Finally, we also tested whether attitudes towards refugees are as expected weakly, if at all, related to Big Five personality traits (usually $|rs| < .30$; see Sibley & Duckitt, 2008), and demographics, to provide further evidence for discriminant validity (expected $|rs| < .30$). The constructs we included from this and other relevant (sub-) modules for these analyses, item verbatims, scale ranges, and anchors, as well as expected direction of the relationship with attitudes towards refugees are summarized in Table 4. Whenever constructs were measured with more than one item, we modelled them latently. To examine convergent/criterion and discriminant validity, we fit four models (a three-factor and a second-order factor model of the refugee submodule, and the same for the Sinti and Roma submodule) in which we correlated all constructs listed above with one another and inspected whether correlations between these constructs and the attitude factor(s) were indeed in the expected direction and of expected strength.

Results

Descriptive statistics of both waves of measurement are presented in Table S1. Model-fit indicators for both measurement waves are depicted in Table 5. Standardized factor loadings, residual variances, factor correlations, and internal consistency indicators of the two models at both measurement waves are depicted in Table S8. Results of τ-equivalence testing are shown in Table S9, and the assessment of the amount of bias introduced if observed unweighted mean scores were used is shown in Table S10. Information on the longitudinal measurement equivalence testing is provided in Table S11. Correlations of the three subfactors and the second-order attitude factor of Wave 1 with relevant items to assess convergent/criterion and discriminant validity are depicted in Table 6.

Both two-wave models showed good model-fit. Once again, the standardized factor loadings were high ($|\lambda| > 0.5$). For the three-factor model, again, the factor correlations within each wave of measurement were above $|r| < .80$ (exception: Wave 2, $r_{F1\ F3} = .797$), indicating a low discriminant validity between the three attitude component factors. Auto-correlations of each of the factors between Waves 1 and 2 (e.g., $F1_{Wave\ 1}$ with $F1_{Wave\ 2}$, etc.) were $rs = .854–.925$, indicating that the construct vis-à-vis construct facets remained highly stable over time.

Like in the previous studies, internal consistencies of all factors in the three-factor model/first-order factors of the second-order model ($\omega s > .85$), and the second-order factor ($\omega = .96$) were high in both waves. Also, like in the previous studies, τ-equivalence could not be established (exception: for the behavioural intentional factor of Wave 2), and subsequently, also not for the first-order factors in the

TABLE 4 Study 2b: Constructs, including wave of assessment, verbatim, scale range, and anchors, as well as expected direction of relationship with attitudes towards refugees, for the assessment of convergent/criterion and discriminant validity

Constructs	Included in Wave	Verbatim of items	Scale anchors	Expected direction of relationship with attitudes towards refugees
Intergroup-related constructs expected to be related to attitudes towards refugees to a moderate to high extent to evidence convergent/criterion validity (expected $\|rs\|$ = .30–.80)				
Positive contact with refugees	EB (04-06/2017)	Wie häufig haben Sie positiven oder erfreulichen Kontakt mit Flüchtlingen in Ihrer Nachbarschaft? *How frequently do you have positive or good contact with refugees in your neighbourhood?* Wie häufig haben Sie positiven oder erfreulichen Kontakt mit Flüchtlingen an Ihrem Arbeits- oder Ausbildungsplatz? *How frequently do you have positive or good contact with refugees at your place of employment or apprenticeship?*	[1 – nie; 4 – häufig] *[1 – never; 4 – frequently]*	Positive (e.g., Barlow et al, 2012; Wagner et al., 2002)
Negative contact with refugees	EB (04-06/2017)	Wie häufig haben Sie negativen oder unerfreulichen Kontakt mit Flüchtlingen in Ihrer Nachbarschaft? *How frequently do you have positive or good contact with refugees in your neighborhood?* Wie häufig haben Sie negativen oder unerfreulichen Kontakt mit Flüchtlingen an Ihrem Arbeits- oder Ausbildungsplatz? *How frequently do you have negative or bad contact with refugees at your place of employment or apprenticeship?*	[1 – nie; 4 – häufig] *[1 – never; 4 – frequently]*	Negative (e.g., (e.g., Barlow et al, 2012; Wagner et al., 2002)
Diversity beliefs	EB (04-06/2017)	Ich schätze die kulturelle Vielfalt in Deutschland, weil sie dem Land einen Nutzen bringt. *I value cultural diversity in Germany because it is useful for the country.* Eine Gesellschaft mit einem hohen Ausmaß an kultureller Vielfalt ist eher befähigt, neue Probleme in Angriff zu nehmen. *A society with a high degree of cultural diversity is more capable of tackling new problems.*	[1 – stimme überhaupt nicht zu; 4 – stimme vollkommen zu] *[1 – strongly disagree; 4 – strongly agree]*	Positive (e.gAsbrock et al., 2011; Kauff & Wagner, 2012)

TABLE 4 (Continued)

Constructs	Included in Wave	Verbatim of items	Scale anchors	Expected direction of relationship with attitudes towards refugees
Allophilia	EB (04-06/2017)	Ich habe grundsätzlich eine positive Einstellung gegenüber Menschen aus anderen Kulturkreisen. *In general, I have positive attitudes about people from different cultural backgrounds.* Ich mag Menschen aus anderen Kulturkreisen. *I like people from different cultural backgrounds.*	[1 – stimme überhaupt nicht zu; 4 – stimme vollkommen zu] *[1 – strongly disagree; 4 – strongly agree]*	Positive (e.g., Pittinsky et al., 2011)
Ingroup identification	EB (04-06/2017)	Ich bin stolz, Deutscher zu sein. *I am proud to be German.* Deutsch zu sein macht einen bedeutenden Teil meiner Persönlichkeit aus. *To be German is an important part of my personality.*	[1 – stimme überhaupt nicht zu; 4 – stimme vollkommen zu] *[1 – strongly disagree; 4 – strongly agree]*	Negative (e.gBecker et al., 2007; Haddock et al., 1993)
Fraternal relative deprivation (refugee)	EB (04-06/2017)	Wenn Sie die wirtschaftliche Lage der Deutschen mit der wirtschaftlichen Lage der in Deutschland lebenden Flüchtlinge vergleichen, wie geht es den Deutschen im Vergleich zu Flüchtlingen? *If you compare the economic situation of Germans with the economic situation of refugees in Germany, how are Germans doing compared to refugees?*	[1 – sehr viel besser; 5 – sehr viel schlechter][a] [1 – very much better; 5 – *very much worse*]	Negative (e.g Pettigrew et al., 2008; Smith et al., 2012)
Political right-wing orientation	EB (04-06/2017)	In der Politik spricht man manchmal von"links" und"rechts". Wo auf der Skala von 0 bis 10 würden Sie sich selbst einstufen, wenn 0 für"links" steht und 10 für"rechts"? *In politics people sometimes talk of"left" and"right". Where would you place yourself on this scale, where 0 means"left" and 10 means"right"?*	[0 – links; 10 – rechts] *[0 – left; 10 – right]*	Negative (e.g., Cowling et al., 2019)
Probability to vote for AfD (right-wing party)	EC (06-08/2017)	Wie wahrscheinlich ist es, dass Sie jemals die AfD wählen werden? *How probable is it that you will ever vote for the AfD?*	[1 – sehr unwahrscheinlich; 7 – Sehr wahrscheinlich] *[1- Very unlikely; 7 – very likely]*	Negative (e.g., Cowling et al., 2019)

(Continues)

TABLE 4 (Continued)

Constructs	Included in Wave	Verbatim of items	Scale anchors	Expected direction of relationship with attitudes towards refugees
Perceived conflict between foreigners and Germans	EC (06-08/2017)	Sind die Konflikte Ihrer Meinung nach bei folgenden Gruppen sehr stark, ziemlich stark, eher schwach, oder gibt es da gar keine Konflikte? Zwischen Ausländern und Deutschen. *Please tell me whether in your view these conflicts are very strong, fairly strong, quite weak, or not conflicts at all. Between foreigners and Germans.*	[1 – Gibt gar keine; 4 – sehr stark][a] *[1 – Don't exist; 4 – Very strong]*	Negative (Sherif et al. (1954/1988)
Intergroup constructs expected to be related to attitudes towards refugees to a smaller extent than the corresponding refugee-specific constructs to evidence divergent validity				
Positive contact with Sinti and Roma	EB (04-06/2017)	Wie häufig haben Sie positiven oder erfreulichen Kontakt mit Sinti und Roma in Ihrer Nachbarschaft? *How frequently do you have positive or good contact with Sinti and Romani in your neighborhood?* Wie häufig haben Sie positiven oder erfreulichen Kontakt mit Sinti und Roma an Ihrem Arbeits- oder Ausbildungsplatz? *How frequently do you have positive or good contact with Sinti and Roma at your place of employment or apprenticeship?*	[1 – nie; 4 – häufig] *[1 – never; 4 – frequently]*	Positive (e.g., Barlow et al, 2012; Wagner et al., 2002)
Negative contact with Sinti and Roma	EB (04-06/2017)	Wie häufig haben Sie negativen oder unerfreulichen Kontakt mit Sinti and Roma in Ihrer Nachbarschaft? *How frequently do you have positive or good contact with Sinti and Roma in your neighborhood?* Wie häufig haben Sie negativen oder unerfreulichen Kontakt mit Sinti and Roma an Ihrem Arbeits- oder Ausbildungsplatz? *How frequently do you have negative or bad contact with Sinti and Roma at your place of employment or apprenticeship?*	[1 – nie; 4 – häufig] *[1 – never; 4 – frequently]*	Negative (e.g., Barlow et al, 2012; Wagner et al., 2002)

TABLE 4 (Continued)

Constructs	Included in Wave	Verbatim of items	Scale anchors	Expected direction of relationship with attitudes towards refugees
Relative deprivation (Sinti and Roma)	EB (04-06/2017)	Wenn Sie die wirtschaftliche Lage der Deutschen mit der wirtschaftlichen Lage der in Deutschland lebenden Sinti und Roma vergleichen, wie geht es den Deutschen im Vergleich zu Sinti und Roma? *If you compare the economic situation of Sinti and Roma with the economic situation of refugees in Germany, how are Germans doing compared to Sinti and Roma?*	[1 – sehr viel besser; 5 – sehr viel schlechter][a] *[1 – very much better; 5 – very much worse]*	Negative (e.gPettigrew et al., 2008; Smith et al., 2012)
Non-intergroup constructs expected to be related to attitudes towards refugees to a small extent, if any, to evidence divergent validity (expected $\|rs\| < .30$)				
Big Five: Openness	EC (06-08/2017)	Ich bin nicht sonderlich kunstinteressiert (r). *I have few artistic interests (r).* Ich kann mich für Kunst, Musik und Literatur begeistern. *I am fascinated by art, music, or literature.* Es macht mir Spaß, gründlich über komplexe Dinge nachzudenken und sie zu verstehen. *I am complex, a deep thinker.* Mich interessieren abstrakte Überlegungen wenig (r). *I have little interest in abstract ideas (r).* Ich bin nicht besonders einfallsreich (r). *I have little creativity (r).* Ich bin originell, entwickle neue Ideen. *I am original, come up with new ideas.*	[1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu] *[1 disagree completely; 5 = agree completely]*	Positive (Sibley & Duckitt, 2008)
Big Five: Conscientiousness	EC (06-08/2017)	Ich bin eher unordentlich (r). *I tend to be disorganized (r).* Ich mag es sauber und aufgeräumt. *I keep things neat and tidy.* Ich neige dazu, Aufgaben vor mir herzuschieben (r). *I have difficulty getting started on tasks.* Ich bleibe an einer Aufgabe dran, bis sie erledigt ist. *I am persistent, work until the task is finished.* Ich bin manchmal ziemlich nachlässig (r). *I can be somewhat careless (r).* Ich bin verlässlich, auf mich kann man zählen. *I am reliable, can always be counted on.*	[1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu] *[1 disagree completely; 5 = agree completely]*	/ (Sibley & Duckitt, 2008)

(Continues)

TABLE 4 (Continued)

Constructs	Included in Wave	Verbatim of items	Scale anchors	Expected direction of relationship with attitudes towards refugees
Big Five: Extraversion	EC (06-08/2017)	Ich gehe aus mir heraus, bin gesellig. *I am outgoing and a sociable person.* Ich bin eher ruhig (r). *I tend to be quiet (r).* Ich neige dazu, die Führung zu übernehmen. I am dominant, act as a leader. In einer Gruppe überlasse ich Lieber anderen die Entscheidung (r). *I prefer to have others take charge (r).* Ich bin weniger aktiv und unternehmenslustig als andere (r). *I am less active than other people (r).* Ich bin voller Energie und Tatendrang. *I am full of energy.*	[1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu] *[1 disagree completely; 5 = agree completely]*	Positive (Sibley & Duckitt, 2008)
Big Five: Agreeableness	EC (06-08/2017)	Ich bin einfühlsam, warmherzig. *I am compassionate, have a soft heart.* Andere sind mir eher gleichgültig, egal (r). *I can be cold and uncaring (r).* Ich begegne anderen mit Respekt. *I am respectful, treat others with respect.* Ich bin manchmal unhöflich und schroff (r). *I am sometimes rude to others (r).* Ich neige dazu, andere zu kritisieren (r). *I tend to find fault with others (r).* Ich schenke anderen leicht Vertrauen, glaube an das Gute im Menschen. *I assume the best about people.*	[1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu] *[1 disagree completely; 5 = agree completely]*	Positive (Sibley & Duckitt, 2008)

TABLE 4 (Continued)

Constructs	Included in Wave	Verbatim of items	Scale anchors	Expected direction of relationship with attitudes towards refugees
Big Five: Emotionality	EC (06-08/2017)	Ich bleibe auch in stressigen Situationen gelassen (r). *I am relaxed, handle stress well (r).* Ich mache mir oft Sorgen. *I worry a lot.* Ich bin selbstsicher, mit mir zufrieden (r). *I feel secure, comfortable with myself (r).* Ich bin oft deprimiert, niedergeschlagen. *I tend to feel depressed, blue.* Ich bin ausgeglichen, nicht leicht aus der Ruhe zu bringen (r). *I am emotionally stable, not easily upset (r).* Ich reagiere schnell gereizt und genervt. *I am temperamental, get emotional easily.*	[1 = stimme überhaupt nicht zu; 5 = stimme voll und ganz zu] *[1 disagree completely; 5 = agree completely]*	/ (Sibley & Duckitt, 2008)
Gender	DF (12/2016-02/2020)	Sind Sie männlich oder weiblich? *Are you male or female?*	[1 – männlich; 2 – weiblich] *[1 = male; 2 = female]*	/
Year of birth	DF (12/2016-02/2020)	In welchem Jahr sind Sie geboren? *When were you born?*	Open question	/
Personal income	DF (12/2016-02/2020)	Wie hoch ist Ihr eigenes durchschnittliches monatliches Nettoeinkommen? *How high is your average monthly net income?*	[1 – Unter 300 €; 15–5000 € und mehr] *[1 – Under 300 €; 15–5000 € and more]*	/

[a]Answering scale inverted, which means that higher values indicate higher levels of respective construct (e.g., more fraternal relative deprivation, or more perceived conflict). (*r*) = Single item on scale inverted. Non-italic represents original formulation in German, italic represents English translation as provided by the codebook of the GESIS Panel, available at https://www.gesis.org/gesis-panel/documentation. Items have the following labels in the GESIS Panel: positive contact with refugees = 'EBBD084A & EBBD086A'; negative contact with refugees = 'EBBD085A & EBBD087A'; diversity beliefs = 'EBBD124A & EBBD125A'; allophilia = 'EBBD126A & EBBD127A' ingroup identification = 'EBBD134A & EBBD135A'; fraternal relative deprivation (refugees) = 'EBBD128A'; political right-wing orientation = 'EBZC061A'; probability to vote for AfD (right-wing party) = 'ECZY123A'; perceived conflict between foreigners and Germans = "ECBO095A"; positive contact with Sinti and Roma = 'EBBD095A & EBBD097A'; negative contact with Sinti and Roma = 'EBBD096A & EBBD098A'; fraternal relative deprivation (Sinti and Roma) = 'EBBD133A'; Gender = 'DFZH037A'; Year of birth = 'DFZH038C', personal income = 'DFZH055B', Big Five = 'ECBQ051A – ECBQ071A'.

TABLE 5 Study 2b: Goodness of fit information of the two remaining competing models

Model	AIC	χ^2_{cor}	*df*	*p*	χ^2_{cor}/df	RMSEA	CFI	SRMR
Model 2: Three-factor model	51937.098	77.268	33	<.001	2.341	.025	.997	.009
Model 3: Second-order factor model	51937.390	85.393	38	<.001	2.247	.024	.997	.010

Abbreviations: AIC, Akaike Information Criterion; CFI, comparative fit index; df, degrees of freedom; *p*, probability value; RMSEA, root mean square error of approximation; SMR, standardized root mean square residual.

second-order model. τ-equivalence could be established for the second-order factor of Wave 2, yet not for Wave 1. Correlation coefficients between the unweighted mean scores and corresponding factor scores ranged between $r = .968–.993$.

Examining longitudinal measurement equivalence, the configural models for both conceptualizations showed good fit. When testing metric and scalar measurement equivalence in both models, the changes in the relevant model-fit indices were below the cut-off criteria (Chen, 2007), evidencing metric and scalar measurement equivalence.

To examine convergent/criterion and discriminant validity, we correlated the three-factor model and the second-order factor model of Wave 1 with the constructs enumerated above. In both models, the attitude factor(s) correlated in the expected direction and were of expected strength with the constructs (intergroup constructs $|rs| = .30–80$; intergroup constructs related to Sinti and Roma are smaller than for those related to refugees; positive contact; $rs = .052–.136$ vs. $.403–.529$; negative contact: $rs = −.065$ to $−.098$ vs. $−.350$ to $−.417$; fraternal relative deprivation $rs = −.238–.306$ vs. $−.369$ to $−.434$; non-intergroup constructs $|rs| < .30$), evidencing convergent/criterion and discriminant validity (exception: ingroup identification correlated with the emotional factor $r = −.289$, $p < .001$, which was slightly below expected range ($|rs| = .30–80$).

Discussion

Study 2b replicated our findings of the previous studies concerning good overall model-fit and indicator functioning. Like in the studies before, internal consistencies were high across all factors. Again, τ-equivalence of the factors was rather the exception than the rule, although our follow-up analyses suggested that the bias that would be introduced if one would not account for this in observed-value analyses would be small.

Our test for measurement equivalence showed equally good results for both competing factorial structure models, whereby scalar equivalence of the indicators (in both models) at first-order and – for the second-order factor model – also the second-order level could be established. Thus, the proposed scale is apt for comparative or longitudinal correlational/regression-based analysis and can also be used for valid and meaningful (latent) mean value comparison. Lastly, the scale performed well in terms of convergent/criterion and discriminant validity, both when attitudes were conceptualized as three separate factors, as well as and an overarching second-order factor.

GENERAL DISCUSSION

In the present research, we developed, empirically tested, and independently validated a short-scale that explicitly considers cognitive, emotional, and behavioural intentional components of attitudes towards refugees. This was motivated by two factors: First, the need of researchers to draw on short, reliable, and validated measures to assess attitudes towards refugees as a social group of increasing social, political,

TABLE 6 Study 2b: Correlations of the three attitude components of the three-factor model (Wave 1) and the global attitude factor of the second-order factor model (Wave 1) with relevant constructs based on subsample

	Model 2: 3-Factor Model correlates[a]			Model 3: Second-Order Factor Model correlates[a]
	F1	F2	F3	F0
Intergroup-related constructs expected to be related to attitudes towards refugees to a moderate to high extent to evidence convergent/criterion validity (expected |*rs*| = .30–.80)				
Positive contact with refugees	.403***	.529***	.479***	.514***
Negative contact with refugees	−.417***	−.403***	−.350***	−.417***
Diversity beliefs	.715***	.682***	.670***	.729***
Allophilia	.613***	.648***	.656***	.680***
Ingroup identification	−.324***	−.289***	−.374***	−.343***
Relative deprivation (refugees)	−.379***	−.434***	−.369***	−.433***
Right-wing political orientation	−.355***	−.321***	−.299***	−.339***
Probability to vote for AfD (right-wing party)	−.510***	−.535***	−.524***	−.557***
Perceived conflict between foreigners and Germans	−.374***	−.338***	−.404***	−.386***
Intergroup constructs expected to be related to attitudes towards refugees to a smaller extent than the corresponding refugee-specific constructs to evidence divergent validity				
Positive contact with Sinti and Roma	.131	.052	.136*	.114
Negative contact with Sinti and Roma	−.065	−.098	−.073	−.088
Relative deprivation (Sinti and Roma)	−.302***	−.298***	−.238***	−.306***
Non-intergroup constructs expected to be related to attitudes towards refugees to a small extent, if any, to evidence divergent validity (expected |*rs*| < .30)				
Gender	.004	.061	.015	.035
Year of birth	.064	−.021	.096*	.036
Personal income	.047	.102	.073	.080
Big Five: Openness	.243***	.201**	.273***	.248***
Big Five: Conscientiousness	−.144*	−.017	−.016	−.054
Big Five: Extraversion	.078	.122*	.108	.114
Big Five: Agreeableness	.055	.199**	.130	.151*
Big Five: Emotionality	−.115*	−.139*	−.097	−.129*

Note: We examined the correlations between the newly developed scale and these constructs based on $N = 516$ participants that were assigned to the GESIS Panel submodule on refugees (Wagner et al., 2018; exception: positive and negative contact with Sinti and Roma and relative deprivation (Sinti and Roma)). We fit two models (one for the 3-factor model, one for the second-order factor model for the submodule on refugee) in which we included multi-item constructs latently, and single-item constructs manifestly. We fit separate models to examine the correlations between positive and negative contact with Sinti and Roma based on $N = 510$ participants that were assigned to the GESIS Panel submodule on Sinti and Roma (Wagner et al., 2018). Only correlations with attitude components (three-factor model)/attitudes (second-order factor model) are shown. F1 = cognitive component; F2 = emotional component; F3 = behavioural intentional component; F0 = general second-order attitude factor. We dropped ECBQ042A in the analyses because of a very low factor loading on the extraversion factor, and allowed correlations of residual variances of Big Five items to improve model fit; openness: 'ECBQ061A' and 'ECBQ046A'; extraversion: 'ECBQ062A' and 'ECBQ047A'; emotionality: 'ECBQ060A' and 'ECBQ055A'.

[a]Model fit of model 2, refugee submodule: χ^2_{cor} (1089) = 369.535, $p < .001$, RMSEA = 048, CFI = 855, SRMR = .060, AIC = 64932.099; global model fit mostly determined by Big Five scale requiring further correlations between residuals; since this was not our focus we refrained from further model adjustments; Sinti and Roma submodule: χ^2_{cor} (30) = 84.553, $p < .001$, RMSEA = .060, CFI = .967, SRMR = .035, AIC = 11268.223.

[b]Model fit of model 3, refugee submodule: χ^2_{cor} (1123) = 2456.569, $p < .001$, RMSEA = .048, CFI = .850, SRMR = .060, AIC = 64951.261; Sinti and Roma submodule: χ^2_{cor} (36) = 90.930, $p < .001$, RMSEA = .055, CFI = .967, SRMR = .037, AIC = 11267.407.

and academic relevance; and second, claims of scholars that argue that cognitions, affect, and behavioural intentions are important facets of attitudes (e.g., Brown, 2010; Cuddy et al., 2007; Rosenberg & Hovland, 1960), while previous scales either focus on specific attitude components, or suffer other kinds of limitations that question their applicability, validity, or reliability.

In expert discussions among a diverse team of researchers, we used various sources to create a short, six-item scale that considers all three attitude components. Acknowledging and testing three different theoretical conceptualizations how these components relate to one another (one-factor model, three-factor model, second-order factor model), we validated the new scale across three studies based on heterogeneous adult samples.

Studies 1 and 2a allowed us to determine and cross-validate that a one-factor conceptualization of attitudes was ill-equipped to represent the relationship between the attitude components. Instead, a three-factor and second-order factor conceptualization emerged as similarly well qualified to represent the empirical reality of associations between items. Across all three studies we conducted (Studies 1, 2a, and 2b), model-fits were identical (Studies 1 and 2a) or very similar (Study 2b) due to model equivalence, allowing for little empirical grounds to recommend one conceptualization over the other. Factor loadings were consistently high using both conceptualizations, indicating that large proportions of variance could be explained by the attitude factors, which was also reflected in the high internal consistencies (ωs). The consistently high correlations between the three attitude components provides some evidence that a global, second-order factor conceptualization of attitudes may be preferred. Thus, overall, our analyses showed that researchers can unreservedly use the present scale using both a three-component and second-order conceptualizations of attitudes.

While Studies 1 and 2a showed that the scale behaved well cross-sectionally, Study 2b demonstrated that the scale also performs well longitudinally. As can be expected by an attitude, stability was high for both the three-factor and second-order model. Additionally, scalar measurement equivalence could be established using both conceptualizations of attitudes, indicating that the scale is not only equally conceptualized over time, but also that valid and meaningful (latent) correlational/regression-based as well as mean value-based analyses over time are possible. Lastly, Study 2b illustrated that the scale is correlated in the expected direction and to an expected extent with other measures, evidencing convergent/criterion and discriminant validity of the short-scale with both conceptualizations of attitudes. Thus, as the scale has passed all these tests successfully, we can recommend the usage of the scale in research on attitudes towards refugees from the majority's perspective. Interested researchers are also welcome to use the data of the waves we have just presented free of charge, as well as later waves fielded this scale in the GESIS-Panel (GESIS, 2017).

We recommend to use this scale applying the second-order conceptualizations of attitudes, especially to researchers who follow the rationale we built based on Eagly and Chaiken (1993), and who are interested in a general attitude and/or how this relates to third variables. However, a three-component conceptualization fits equally well, justifying the use of the three-component conceptualization and modelling each of the scale components separately, if the researcher follows the rationale of scholars who see the components as separate, maybe even be causally related to one another, they aim to explicitly examine these components and/or how they relate to one another or third variables.

Using the scale *as is* only makes sense if the researchers' conceptualization of attitudes matches the respective conceptualization we elaborated on (second-order or three-factor model), and if they do not include constructs that conceptually overlap with scale components. If researchers aim to predict attitudes towards refugees, for example, using threat (which some researchers conceive as separate constructs that are causally linked; Stephan & Stephan, 2000), it may make sense to exclude the cognitive component of this scale, which taps into related cognitions. Thus, based on theoretical and empirical considerations, we encourage researchers to flexibly adjust the scale and which components to include in their analyses based on their research needs.

Despite the overall pleasing results, we need to caution researchers who intend to use the scale based on observed values by simply computing mean or sum scores. We could establish τ-equivalence

only rarely, which would be a requirement for such applications. Having said that, our follow-up analyses suggest that using unweighted scores would bias results only very little. Indeed, weighted factor scores and unweighted mean scores consistently correlated very highly. Thus, we argue that it may be justifiable, especially in settings where the costs outweigh the benefits, to forego this adjustment step and use manifest mean scores instead. However, especially when addressing sophisticated research questions, we recommend weighting each indicator when computing scales on the level of observed values, for instance by using factor scores instead manifest means (for a hands-on example in SPSS, see IBM Support, 2020, for Mplus, see our syntaxes in the OSM, or Muthén & Muthén, 1998–2017, Example 13.15).

All scale items tap into dimensions that we believe to be universally relevant to assess attitudes towards refugees regardless of the country they originate from and the country in which the scale is administered. We also believe that the scale can be easily translated and adapted to other languages. Nonetheless, the presented scale has been developed and tested in the German context only, and cross-cultural psychology suggests that invariance of scale properties across countries should not be assumed, but tested (van de Schoot et al., 2015). Although we drew on large, heterogeneous adult samples, some of them even German probability samples (Study 2a and 2b), which provides confidence that the scale works well in Germany, only further empirical work can show to what extent our findings are generalizable to other country-contexts.

Additionally, we acknowledge that one of the core strengths of this scale – its shortness – is also a weakness. As pointed out by an anonymous reviewer, the scale does not capture other important cognitions, including people's knowledge about some specific characteristics of individuals engaging in forced migration, such as knowledge about the involuntary nature of leaving one's home, encountered hardships, possibly severe risks and threats to one's life before or during migration, distress, and trauma. People can harbor many more relevant emotions towards refugees, including empathy, fear/anxiety, anger, admiration, disgust, and guilt; Cottrell & Neuberg, 2005; Cuddy et al., 2007; Mackie et al., 2000; Thomas et al., 2019; Pettigrew & Tropp, 2008). There are also numerous additional ways people could intend to behave towards refugees, both pro- and anti-socially, such as direct helping, engaging in solidarity-based collective action, ignoring, discriminating, or committing hate crimes (Cuddy et al., 2007; Kotzur et al., 2019; Pager & Shepherd, 2008; Penner et al., 2005; Tajfel & Turner, 1979; Thomas et al., 2019; van Zomeren et al., 2008; Wagner et al., 2020). To ensure the scale's brevity while still covering three attitude components, we had to make choices which aspects to cover, which was not an easy task. We encourage researchers to explore these and further additional aspects we did not cover in our attempt to provide a very short scale.

CONCLUSION

We developed and validated a six-item short-scale to measure attitudes towards refugees explicitly considering cognitive, emotional, and behavioral intentional components. The scale passed all psychometric tests successfully, if a three-factor or second-order factor conceptualization of attitudes in a latent variable framework were used. We hope that we have contributed to a new wave of highly valid and reliable research of social, political, and academic relevance on attitudes towards refugees in Germany and beyond, and that this research inspires researchers to explicitly consider, debate, and formally test different conceptualization of attitudes towards refugees or other attitude targets.

ACKNOWLEDGEMENT

The authors thank Simon Greipl for running preliminary analyses for Study 1. All authors declare no conflict of interest. The research was conducted in accordance with the relevant APA and DGPS ethical guidelines. Data of Study 1 can be requested from the corresponding author, since participants of Study 1 have been informed that data accessibility will be controlled by the authors at all times. Data of Study 2a and 2b can be accessed from the GESIS data repository.

CONFLICT OF INTEREST

All authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data of Study 1 can be requested from the corresponding author, since participants of Study 1 have been informed that data accessibility will be controlled by the authors at all times. Data of Study 2a and 2b can be accessed from the GESIS data repository.

AUTHOR CONTRIBUTION

Patrick Ferdinand Kotzur: Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Visualization; Writing – original draft; Writing – review & editing. **Maria-Therese Friehs:** Conceptualization; Validation; Visualization; Writing – original draft; Writing – review & editing. **Peter Schmidt:** Conceptualization; Methodology; Supervision; Writing – review & editing. **Ulrich Wagner:** Conceptualization; Methodology; Resources; Supervision; Writing – review & editing. **Steffen Pötzschke:** Conceptualization; Data curation; Writing – review & editing. **Bernd Weiss:** Conceptualization; Data curation; Writing – review & editing.

ORCID

Patrick F. Kotzur https://orcid.org/0000-0002-5193-3359
Maria-Therese Friehs https://orcid.org/0000-0002-5897-8226

REFERENCES

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314. https://doi.org/10.1037/rev0000262

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, *52*(1), 27–58. https://doi.org/10.1146/annurev.psych.52.1.27

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Asbrock, F., Kauff, M., Issmer, C., Christ, O., Pettigrew, T. F., & Wagner, U. (2011). Kontakt hilft – Auch wenn es die Politik nicht immer einfach macht. In W. Heitmeyer (Ed.), *Deutsche Zustände. Folge 10* (pp. 199–219). Suhrkamp.

Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R., Harwood, J., Rubin, M., & Sibley, C. G. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin*, *38*(12), 1629–1643. https://doi.org/10.1177/0146167212457953

Becker, J., Wagner, U., & Christ, O. (2007). Nationalismus und Patriotismus als Ursache von Fremdenfeindlichkeit. In W. Heitmeyer (Ed.), *Deutsche Zustände. Folge 5* (pp. 131–149). Suhrkamp.

Bentler, P. M., & Satorra, A. (2010). Testing model nesting and equivalence. *Psychological Methods*, *15*(2), 111–123. https://doi.org/10.1037/a0019625

Berry, M., Garcia-Blanco, I., & Moore, K. (2015). *Press coverage of the refugee and migrant crisis in the EU: A content analysis of five European countries*. [Project Report]. United Nations High Commissioner for Refugees. http://www.unhcr.org/56bb369c9.html

Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: a literature review, application to content validity, and meta-analysis. *Organizational Research Methods*, *10*, 689–709. https://doi.org/10.1177/1094428106294734

Bogardus, E. S. (1947). Measurement of personal-group relations. *Sociometry*, *10*, 306–311. https://doi.org/10.2307/2785570

Bohrer, B., Friehs, M. T., Schmidt, P., & Weick, S. (2019). Contacts between natives and migrants in Germany: Perceptions of the native population since 1980 and an examination of the contact hypotheses. *Social Inclusion*, *7*, 320–331. https://doi.org/10.17645/si.v7i4.2429

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. G. Joreskog, & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction (Part I)*. North-Holland.

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika*, *50*, 229–242. https://doi.org/10.1007/BF02294248

Bosnjak, M., Dannwolf, T., Enderle, T., Schaurer, I., Struminskaya, B., Tanner, A., & Weyandt, K. (2018). Establishing an open probability-based mixed-mode panel of the general population in Germany: The GESIS Panel. *Social Science Computer Review*, *36*, 103–115. https://doi.org/10.1177/0894439317697949

Brislin, R. W. (1990). *Applied cross-cultural psychology*. Sage.

Brown, R. (2010). *Prejudice: Its social psychology*. Wiley-Blackwell.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*, 2nd ed. The Guilford Press.

Bundesamt für Migration und Flüchtlinge (2017). *Das Bundesamt in Zahlen 2016 – Asyl, Migration und Integration. [The federal bureau in numbers 2016 – Asylum, migration, and integration]*. Retrieved from: http://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/Broschueren/bundesamt-in-zahlen-2016.html?nn=9121126

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. Erlbaum.

Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to "prejudice". *Journal of Personality and Social Psychology*, *88*(5), 770–789. https://doi.org/10.1037/0022-3514.88.5.770

Couliaraki, L., & Stolic, T. (2017). Rethinking media responsibility in the refugee 'crisis': a visual typology of European news. *Media, Culture & Society*, *39*, 1162–1177. https://doi.org/10.1177/0163443717726163

Cowling, M. M., Anderson, J. R., & Ferguson, R. (2019). Prejudice-relevant correlates of attitudes towards refugees: A meta-analysis. *Journal of Refugee Studies*, *32*, 502–524. https://doi.org/10.1093/jrs/fey062

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*, 631–648. https://doi.org/10.1037/0022-3514.92.4.631

Dalege, J. B., Borsboom, D., Van Harreveld, D., Van den Berg, H., Conner, M., & Van der Maas, H. L. J. (2016). Toward a formalized account of attitudes: The causal attitude network (CAN) model. *Psychological Review*, *123*, 2–22. https://doi.org/10.1037/a0039802

Darwall, S. (1998). Empathy, sympathy, care. *Philosophical Studies*, *89*, 261–282.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*, 55–75. https://doi.org/10.1146/annurev-soc-071913-043137

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich.

Echterhoff, G., Hellmann, J. H., Back, M. D., Kärtner, J., Morina, N., & Hertel, G. (2020). Psychological antecedents of refugee integration (PARI). *Perspectives on Psychological Science*, *15*, 856–879. https://doi.org/10.1177/1745691619898838

Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, *82*, 804–818. https://doi.org/10.1037/0022-3514.82.5.804

Felser, D., & Bohrer, G. (2016). *Bilanz des Helfens 2016 [Record of helping 2016]*. Retrieved from: http://www.spendenrat.de/wp-content/uploads/2016/02/Bilanz_des_Helfens_2016.pdf

Fishbein, M., & Aizen, I. (2010). *Predicting and changing behaviour. The reasoned action approach*. Routledge.

Friehs, M.-T., Kotzur, P. F., Böttcher, J., Zöller, A.-K., Lüttmer, T., Wagner, U., Asbrock, F., & Van Zalk, M. H. W. (2022). Examining the structural validity of stereotype content scales – a preregistered re-analysis of published data and discussion of possible future directions. *International Review of Social Psychology*, *35*(1), 1–18. https://doi.org/10.5334/irsp.613

Fritsche, I., Jonas, E., & Kessler, T. (2011). Collective reactions to threat: Implications for intergroup conflict and for solving societal crises. *Social Issues and Policy Review*, *5*, 101–136. https://doi.org/10.1111/j.1751-2409.2011.01027.x

Geogriou, M., & Zaboroski, R. (2017). *Media coverage of the "refugee crisis": A cross-European perspective*. Council of Europe Report DG1(2017)03.

GESIS (2017). *GESIS panel standard edition*. GESIS Data Archive, Cologne. ZA5665 Datafile. Version 22.0.0. https://doi.org/10.4232/1.12903

Green-Pedersen, C., & Otjes, S. (2017). A hot topic? Immigration on the agenda in Western Europe. *Party Politics*, *25*(3), 424–434. https://doi.org/10.1177/1354068817728211

Haddock, G., Zanna, M. P., & Esses, V. M. (1993). Assessing the structure of prejudicial attitudes: The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology*, *65*, 1105–1118. https://doi.org/10.1037/0022-3514.65.6.1105

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*(1), 78–79. https://doi.org/10.1037/0003-066X.58.1.78

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*, 117–144. https://doi.org/10.1080/03610739208253916

James, W. (1890). *The principles of psychology*, Vol. *II*. Henry Holt & Co.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Sage.

Karakayali, S., & Kleist, O. (2015). *Ehrenamtliche Flüchtlingsarbeitsstudie: Strukturen und Motive der ehrenamtlichen Flüchtlingsarbeit in Deutschland [Voluntary refugee aid study: structures and motives of voluntary refugee aid in Germany]*. Retrieved from: https://www.bim.hu-berlin.de/media/2015-05-16_EFA-Forschungsbericht_Endfassung.pdf

Katz, D., & Stotland, E. (1959). A preliminary statement to a theory of attitude structure and change. In S. Koch (Ed.), *Psychology: A study of a science*, Vol. *3* (pp. 423–475). McGraw-Hill.

Kauff, M., & Wagner, U. (2012). Valuable therefore not threatening: The influence of diversity beliefs on discrimination against immigrants. *Social Psychological and Personality Science*, *3*, 714–721. https://doi.org/10.1177/1948550611435942

Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances: Is a Heywood case a symptom of misspecification? *Sociological Methods and Research*, *41*(1), 124–167. https://doi.org/10.1177/0049124112442138

Kotzur, P. F., Forsbach, N., & Wagner, U. (2017). Choose your words wisely: Stereotypes, emotions, and action tendencies toward fled people as a function of the group label. *Social Psychology*, *48*, 226–241. https://doi.org/10.1027/1864-9335/a000312

Kotzur, P. F., Schäfer, S. J., & Wagner, U. (2019). Meeting a nice asylum seeker: Intergroup contact changes stereotype content perceptions and associated emotional prejudices, and encourages solidarity-based collective action intentions. *British Journal of Social Psychology*, *58*, 668–690. https://doi.org/10.1111/bjso.12304

Kotzur, P. F., Veit, S., Namyslo, A., Holthausen, M.-A., Wagner, U., & Yemane, R. (2020). "Society thinks they are cold and/or incompetent, but I do not": Stereotype content ratings depend on instructions and the social group's location in the stereotype content space. *British Journal of Social Psychology*, *59*, 1018–1042. https://doi.org/10.1111/bjso.12375

Kotzur, P. F., & Wagner, U. (2021). The dynamic relationship between contact opportunities, positive and negative intergroup contact, and prejudice: A longitudinal investigation. *Journal of Personality and Social Psychology*, *120*, 418–442. https://doi.org/10.1037/pspi0000258

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*, 213–236. https://doi.org/10.1002/acp.2350050305

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*, 537–567. https://doi.org/10.1146/annurev.psych.50.1.537

Landmann, H., Gaschler, R., & Rohmann, A. (2019). What is threatening about refugees? Identifying different types of threat and their association with emotional responses and attitudes towards refugee migration. *European Journal of Social Psychology*, *49*(7), 1401–1420. https://doi.org/10.1002/ejsp.2593

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*, 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology*, *79*(4), 602–616. https://doi.org/10.10371/0022-3514.79.4.602

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide*, 8th ed. Muthén & Muthén.

Nelson, S. C. (2008). Feeling thermometer. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 275–277). Sage Publications. https://doi.org/10.4135/9781412963947.n183

Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual Review of Sociology*, *34*, 181–209. https://doi.org/10.1146/annurev.soc.33.040406.131740

Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*, 533–560. https://doi.org/10.1146/annurev-psych-071620-030619

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, *56*, 365–392. https://doi.org/10.1146/annurev.psych.56.091103.070141

Pettigrew, T. F., Christ, O., Wagner, U., Meertens, R. W., van Dick, R., & Zick, A. (2008). Relative deprivation and intergroup prejudice. *Journal of Social Issues*, *64*, 385–401. https://doi.org/10.1111/j.1540-4560.2008.00567.x

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice. *European Journal of Social Psychology*, *25*, 57–75.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922–934. https://doi.org/10.1002/ejsp.504

Pittinsky, T. L., Rosenthal, S. A., & Montoya, R. M. (2011). Measuring positive attitudes toward outgroups: Development and validation of the Allophilia Scale. In L. R. Tropp, & R. K. Mallett (Eds.), *Moving beyond prejudice reduction: Pathways to positive intergroup relations* (pp. 41–60). American Psychological Association.

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *10*, 336–353. https://doi.org/10.1207/s15327957pspr1004_4

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland, & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). Yale University Press.

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., & Zick, A. (2018). Testing measurement invariance for a second-order factor: A cross-national test of the alienation scale. *Methods, data, analyses: A journal for quantitative methods and survey methodology (mda)*, *12*(1), 47–76. https://doi.org/10.12758/mda.2017.11

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*, 23–74. https://doi.org/10.1016/j.ssresearch.2003.11.003

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1954/1988). *Intergroup conflict and cooperation: The robbers cave experiment*. Wesleyan University Press

Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, *12*(3), 248–279. https://doi.org/10.1177/1088868308319226

Smith, H. J., Pettigrew, T. F., Pippin, G. M., & Bialosiewicz, S. (2012). Relative deprivation: A theoretical and meta-analytic review. *Personality and Social Psychology Review*, *16*, 203–232. https://doi.org/10.1177/1088868311430825

Stephan, W. G., & Renfro, C. L. (2002). The role of threat in intergroup relations. In D. M. Mackie, & E. R. Smith (Eds.), *From prejudice to intergroup relations* (pp. 191–208). Psychology Press.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–46). Lawrence Erlbaum Associates.

Stouffer, S. A., Guttman, L., Suchman, E. A., Lazarsfeld, P. F., Star, S. A., & Clausen, J. A. (1950). *Measurement and prediction. [Studies in social psychology in World War II. Vol.4]*. Princeton University Press.

Support, I. B. M. (2020). *Applying factor analysis results to save factor scores for new cases in SPSS*. Retrieved from: https://www.ibm.com/support/pages/applying-factor-analysis-results-save-factor-scores-new-cases-spss

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In M. J. Hatch, & M. Schultz (Eds.), *Organisational identity: A reader* (pp. 33–47). Oxford University Press.

Thomas, E. F., Smith, L. G. E., McGarty, C., Reese, G., Kende, A., Bliuc, A.-M., Curtin, N., & Spears, R. (2019). When and how social movements mobilize action within and across nations to promote solidarity with refugees. *European Journal of Social Psychology*, *49*(2), 213–229. https://doi.org/10.1002/ejsp.2380

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Measurement invariance. *Frontiers in Psychology*, *6*, 1064. https://doi.org/10.3389/fpsyg.2015.01064

Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, *134*(4), 504–535. https://doi.org/10.1037/0033-2909.134.4.504

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

Veit, S., & Yemane, R. (2020). *Judging without knowing: how people evaluate others based on phenotype and country of origin (Technical Report)*. WZB Discussion Paper, SP VI 2020-101, WZB.

Wagner, U., Christ, O., & Pettigrew, T. F. (2008). Prejudice and group related behaviour in Germany. *Journal of Social Issues*, *64*, 403–416. https://doi.org/10.1111/j.1540-4560.2008.00568.x

Wagner, U., & Greipl, S. (2017). *Psychological research in the recent immigration movement*. Paper presented at the 18th General Meeting of the European Association of Social Psychology, Granada, Spain

Wagner, U., Tachtsoglou, S., Kotzur, P. F., Friehs, M. T., & Kemmesies, U. (2020). Proportion of foreigners negatively predicts the prevalence of xenophobic hate crimes within German districts. *Social Psychology Quarterly*, *83*(2), 195–205. https://doi.org/10.1177/0190272519887719

Wagner, U., van Dick, R., & Endrikat, K. (2002). Interkulturelle Kontakte. In W. Heitmeyer (Ed.), *Deutsche Zustände. Folge 1* (pp. 96-122). Suhrkamp.

Wetzel, E., & Roberts, B. W. (2020). Commentary on Hussey and Hughes (2020): Hidden invalidity among 15 commonly used measures in social and personal psychology. *Advances in Methods and Practices in Psychological Science*, *3*(4), 505–508. https://doi.org/10.1177/2515245920957618

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kotzur, P. F., Friehs, M.-T., Schmidt, P., Wagner, U., Pötzschke, S., & Weiß, B. (2022). Attitudes towards refugees: Introducing a short three-dimensional scale. *British Journal of Social Psychology*, *61*, 1305–1331. https://doi.org/10.1111/bjso.12538