What is a Good Way to Ask About Racist Experiences? Exploring the Difference Between Direct and Indirect Measures

Zühal Arikan[1] · Jörg Dollmann[1,2] · Jannes Jacobsen[1]
[1]Deutsches Zentrum für Integrations- und Migrationsforschung, e.V.
[2]University of Mannheim, MZES

Surveying racist attitudes and actions has shifted from using direct to using indirect measures. However, investigating the different forms of racist attitudes and actions has yet to translate into the measurement of racist experiences. This poses a challenge for capturing the prevalence of racist experiences amidst divergent patterns of racist attitudes and actions. To address this gap, we conducted a comparative analysis of two instruments: a direct and an indirect measure of racist experiences. We aimed at determining the prevalence of racist experiences using both of these instruments, as well as identifying differences in reporting among respondents. Our findings indicate that the instruments yield different prevalence rates, with the indirect instrument reporting a higher prevalence. Additionally, based on the indirect measure, men, younger respondents and immigrants are more likely to report racism, while the level of respondents' education had no discernible effect on the measurement.

Keywords: racism; discrimination; measurement error; DeZIM.panel; survey instrument

1 Introduction

The study of racist attitudes and actions has undergone a shift from using direct measures towards indirect measures (for an overview see Hamidou-Schmidt & Elis, 2023). Instead of asking directly about overt preferences—for example white people versus black people, often referred to as 'old-fashioned racism'—researchers have developed more indirect self-reporting measures of racial attitudes and actions, such as 'symbolic racism' (Sears, 1988), 'modern racism' (McConahay, 1986), and 'new racism' (Jacobson, 1985, for a summary see Axt, 2018). The main reason for this shift is the abolition of racial segregation in North America, which has led to a superficial delegitimisation of racist practices (e.g., Clair & Denis, 2015; Elis & Hamidou-Schmidt, 2023; Sears et al., 2000). Nowadays, racism manifests in more subtle forms, which are not always ex-

plicitly verbalised as such (Axt, 2018; Paulhus, 1984). It is striking, however, that the debate on the different forms of racist *attitudes and actions*—direct vs. indirect—has not been translated into the measurement of racist *experiences*. However, this would be necessary to capture the prevalence of racist experiences against the backdrop of diverging patterns of racist attitudes and actions.

To translate conceptual differences between direct and indirect measures of racist attitudes into survey instruments, researchers rely on three approaches (for a comprehensive overview see Quillian, 2006). One set of studies estimates racism as a residual (e.g., Fernández-Reino et al., 2022). Here, researchers compare outcomes, such as hiring practices, and estimate whether racialised people—for example, black people, or women wearing hijabs—are hired less often, in spite of having the same qualifications as a white person. Such approaches have often been criticised, as they rely on the assumption that the identification strategy is not affected by unobserved heterogeneity. A second set of studies captures direct forms of racism, sometimes also referred to as 'biological racism' (e.g., Forrest et al., 2021). Direct forms of racism are racist attitudes and actions that devaluate individuals due to their phenotype. A third set of studies focuses on implicit, subtle, rather indirect forms of racism (e.g., Gran-Ruaz et al., 2022). These studies argue that racism today is not only visible through the lens of bi-

Corresponding author: Zühal Arikan, Deutsches Zentrum für Integrations- und Migrationsforschung, e.V., Berlin, Deutschland (Email: arikan@dezim-institut.de)

ological racism but also through more subtle forms such as, for instance, the devaluation of a 'black culture', or through unconscious bias.

In this article, we seek to address the question of how the measurement of racist experiences can be best translated and operationalised against the change in the measurement of attitudes and actions. To this end, we test two instruments against each other—one direct and one indirect measure of racist experiences—and then compare the prevalence rates of racist experiences captured by each instrument. The direct measure of racist experience involves directly querying the experience of racism, whereas in the indirect measure, the term 'racism' is explicitly excluded and replaced by a normative list of racist events. Besides comparing prevalence rates, we also focus on respondents who report racist experiences on both scales and compare them to those who report only on the direct measurement but not on the indirect measurement, and vice versa.

Our results show that, first, both instruments produce largely dissimilar prevalence rates of racist experiences, with the indirect instrument yielding a higher prevalence. Additionally, according to the indirect measure, men, younger respondents, and non-natives (immigrants) are

more likely to report racism. Second, we find no evidence that respondents are more likely to report a racist experience on the indirect measure if they consider the experience to be significant. In summary, this research report provides strong evidence that direct and indirect measures of racist experiences work in fundamentally different ways, and that a direct measurement tends to underestimate racist experiences. Furthermore, the combination of both a direct and an indirect scale to measure racist experiences provides a potential for new research. For example: How can we explain that some respondents report theoretically classified instances of indirect racist experiences, but they themselves would not frame those instances as racist?

1.1 Scale Development

This study is based on two instruments. One instrument is designed to observe direct racist experiences and the other to observe indirect racist experiences. The direct instrument is derived from the German National Discrimination and Racism Monitor (NaDiRa; DeZIM, 2022). It is direct in the sense that it explicitly asks for *racist* experiences in the

Table 1

Items (events) of the indirect instrument and the frequency of experience. (Source: own development by the authors' (Dollmann et al., 2023) own calculation; not weighted; translation can be found in the supplementary material)

	N	%
Verbal hostilities		
1. Others made fun of my ethnic origin and/or my ethnic culture of origin in front of me.	580	20
2. Often, I was treated like I wouldn't understand the native language.	367	12
3. I was insulted with swear words related to my ethnic origin and/or my ethnic culture of origin.	216	7
4. I was abused because of my ethnic origin, e.g., I was told to go back to my homeland.	191	6
Exclusion		
5. Persons of the native-born ethnic majority reacted as they were afraid of me, because of my ethnic origin.	159	5
6. I was treated unfairly or harshly by individuals or groups of the native-born ethnic majority because of my accent and/or my foreign appearance.	284	10
7. I was excluded by society because of my accent and/or my foreign appearance, e.g., I was denied entry to a nightclub.	113	4
8. The police stopped and/or questioned me unfairly because of my ethnic origin.	91	3
Physical violence		
9. I was threatened with violence by individuals or groups of the native-born ethnic majority because of my ethnic origin.	67	2
10. I was frisked and/or threatened unfairly by the police because of my ethnic origin.	59	2
11. I was spat at, pushed or accosted by individuals or groups of native-born ethnic majority because of my ethnic origin and/or my foreign appearance.	42	1
12. I got hurt by individuals or groups of the native-born ethnic majority because of my ethnic origin.	27	1
N (distinct individuals)	825	28

past: 'Have you experienced racism in the last 12 months?' The answer options are either 'yes' or 'no' (DeZIM, 2022, 31). The indirect measure was developed by the authors and includes three dimensions: the experiences of (i) verbal hostilities, (ii) exclusion/discrimination, and (iii) physical violence. The dimensions and items were selected and constructed according to Stuart Hall and Eduardo Bonilla-Silva, who define racism as the discrimination of people based on cultural or ethnic characteristics (Bonilla-Silva 2003; Hall, 1997) or discrimination that is linked to positions of power in society (Bonilla-Silva, 2003). Furthermore, racism can manifest itself in three areas (Essed, 1990; Jones, 1972): (i) individual racism, which is experienced on a personal level, (ii) institutional racism, which emanates from the institutions of a society, such as police, authorities, or schools, and (iii) cultural racism, which refers to the cultural practices of a group (Hall, 1989). Any action from a person who discriminates, insults, threatens, or excludes people based on group-related physical characteristics (such as skin colour), ethnic or national origin, or certain cultural characteristics (such as language, religion, or names) can be described as racist.

For each dimension of our scale, four items were constructed listing racist experiences, though not explicitly labelled as such (see Table 1). Respondents were asked whether they had encountered these experiences in the past 12 months, and if so, to what extent were they upset by them. Responses were recorded using a five-point Likert scale (1 = This has never happened to me; 2 = This event happened, but did not bother me; 3 = This event happened and I was slightly upset; 4 = This event happened and I was upset; 5 = This event happened and I was extremely upset). This rating scale is consistent with that used for the Index of Race-Related Stress by Utsey and Ponterotto (1996). Table 1 shows the frequency for each experience (scale rating 2–5).

Both the direct and the indirect scales were placed at the end of the questionnaire of the fourth wave of the DeZIM.panel. This wave focused on the topics of education, labour-market integration, and discrimination. The questions on the indirect scale were asked after the direct measurement. Before both scales, several other questions on experiences with discrimination in different situations—for example, with authorities, at work, during a job search etc.—were asked. The respondents were therefore already sensitised to the topic before the two scales were asked.

2   Materials and Methods

To test both scales in the same population, we used data from the DeZIM.panel, a randomly offline-recruited panel of the population in Germany aged 18–67 that is conducted four times per year (for further information, see Dollmann et al., 2022). The offline recruitment of respondents in 2021 followed a two-stage stratified procedure. The first stage involved regional registration offices ($N = 57$), and the second stage listed individuals (gross sample, $N = 37{,}583$; net sample, $N = 9168$; AAPOR RR2 = 24%). Regional registration offices were assigned sampling probabilities proportional to their size and were stratified by federal state (*Bundesland*) and by rural/urban regions. In the second stage (individual respondents) the DeZIM.panel employed an oversampling of specific immigrant groups. In detail, the oversampling approach focused on the four most important immigrant cohorts in Germany: immigrants from Turkey, immigrants from other majority Muslim countries, immigrants from states with guest-worker agreements, and re-settlers from Eastern Europe. Furthermore, the sample included all other immigrant groups as well as the native-born ethnic majority population. Using adjusted design weights, the data can be inferred to the population living in Germany. For our study, we relied on data from the fourth wave, conducted in autumn 2022 ($N = 2741$). Of all participants in our analytical sample, 46% were men and 54% were women; 87% were German citizens and 13% were non-German citizens; 72% belonged to the native-born ethnic majority and 28% were immigrants or first-generation descendants of immigrants. All respondents provided informed consent prior to the study. More details can be obtained from Dollmann et al. (2022).

2.1   Identification Strategy

To compare the prevalence between the direct and indirect measures, the direct measurement scale was recoded as a dichotomous variable. Respondents who reported at least one of the 12 events listed in Table 1 were coded as 1, otherwise they were coded as 0. Then, to examine the functional differences between the direct and the indirect measurement of racist experiences, we proceeded in three steps. First, using univariate distributions, we showed the prevalence of racist experiences for each instrument separately. Second, we cross-tabulated both scales to identify the number of matches and mismatches between the two measurements, i.e., those respondents who reported racist experiences only in the direct measurement but not in the indirect measurement, and vice versa. We included a McNemar test to test the differences of both proportions. And third, we ran logistic regression models to explain the match and mismatch constellations. To this end, we employed seven variables.

As independent variables, we tested age (metric) and education (1 = 'Basic [*Hauptschule*]'; 2 = 'Intermediate [*Realschule*]'; 3 = 'Entrance qualification for university [*Abitur*]', and 4 = 'University/college degree'). We tested

age because concepts and definitions of racism change over time, and what is classified as racism today could have been called something else in the past. We therefore assumed that age groups would differ in how they answered the direct measure, thus making a mismatch between a direct measure and an indirect measure age-dependent. We tested education, because we anticipated the retrieval to be cognitively more demanding for the direct than for the indirect instrument. Higher educated respondents should therefore be less likely to have a mismatch between the two instruments.

Additionally, we employed an indicator for whether the respondent rated the reported experience with the indirect measure as being significant, in the sense that it led to anger (index on the five-point Likert scale: 1 = no anger, no racial experiences to 5 = extreme anger coded to 0 = no anger, no racial experiences to 4 = extreme anger due to programming reasons). The reasoning behind this was the assumption that insignificant events are more often forgotten and therefore do not come to mind when answering a very general question about racist experiences. In contrast, we believed that asking for specific events is more likely to activate the memory. We tested gender (1 = male, 0 = female) to capture differences in gendered experiences of racism, and non-native status (0 = native, 1 = non-native, i.e., immigrants or first-generation descendants of immigrants) as we believed that immigrants should be more aware that they are at risk of experiencing racism and are therefore more likely to remember and recall specific events that happened to them.

Since the comparison of two scales taps into the concept of measurement error, we additionally controlled for the interview language (1 = German, 0 = otherwise) as well as the total interview duration in minutes. We hence controlled for interview particularities that could confound the sociodemographic effects—such as non-natives using a different language—which could potentially mitigate cross-cultural comparability, as well as the duration of the interview, which is an indicator of response burden. People who experienced a higher response burden might be less detail-oriented when answering the questions and are therefore

Table 2

Frequencies of racist experience for the direct and the indirect instruments. (Source: Dollmann et al. (2023) own calculation, not weighted)

	Direct		Indirect	
	N	*%*	*N*	*%*
Yes	297	10.8	751	27.4
No	2444	89.2	1990	72.6
N	2741	100	2741	100

Table 3

Cross-tabulation of the direct and indirect instrument for measuring racist experiences.

	Direct instrument		
Indirect instrument	*Yes*	*No*	*Total*
Yes	262	489	751
Row %	35	65	100
Cell %	10	18	
No	35	1955	1990
Row %	2	98	100
Cell %	1	73	
Total	296	2444	2741
Row %	11	89	100
Cell %			100

Source: Dollmann et al. (2023) own calculations, not weighted

less likely to report an incident from the rather long list of the indirect measure in Table 1.

2.2 Robustness Checks

We employed three robustness checks. First, to identify whether certain racist hostilities are more likely to occur, we showed the prevalence of each dimension of the indirect scale separately: verbal hostilities, exclusion, and physical violence. Comparing the prevalence of these distributions with the overall distribution allowed us to identify whether some incidents were more likely to happen (or to be reported).

Similarly, in the second robustness check, we calculated the mismatch between the direct and the indirect measure for each dimension separately. This was to identify whether certain dimensions of the indirect measure were more likely to produce a mismatch. And third, we did the same for the multivariate analyses. All models for the robustness checks can be found in the Appendix (see Tables A1–A10).

3 Results

Table 2 presents the frequencies of racist experiences for both the direct and indirect instruments. As can be seen, the proportion of racist experiences is 11% using the direct instrument, compared to 27% using the indirect instrument. The indirect instrument thus leads to a 2.5 times higher prevalence than the direct instrument.

Table 3 cross-tabulates the responses of both instruments, direct and indirect. Regarding the general overlap of the measures, we first focus on the cell percentages, as dis-

played at the bottom of each cell. As can be seen, most cases (73%) report no racist experiences on either the direct or on the indirect measure.

Another 10% of respondents report racist experiences on both measurements. Therefore, 82% of respondents show a match on both direct and indirect measures. However, just under 20% do not show a match. When looking at these mismatches (Table 3, cell percentages), it becomes evident that most mismatches are due to respondents reporting no racist experiences on the direct measure while indicating racist experiences on the indirect measure (18%). In contrast, the mismatch the other way around (direct: yes; indirect: no) is much less likely (1%). When taking the indirect measure as a baseline (row percentages), almost 35% of respondents who responded positively to this measure also responded positively to the direct measures (row percentages). For 65% of respondents there appears to be a mismatch. Turning to those who reported no racist experience on the indirect measure, only 2% reported a racist experience on the direct instrument, i.e., 98% responded negatively on the direct instrument. Furthermore, we applied a McNemar test to test the differences of these proportions. In our analysis, the McNemar chi-squared value is 393.35 and is highly significant, which additionally highlights the discrepancy of the two measurements.

In a next step, we focus on the characteristics of those respondents whose responses agree and those whose responses disagree with the two measurements. For this pur-pose, we first combine the respondents who indicated either a 'yes' ($n = 262$) or a 'no' ($n = 1996$) for both instruments (match). Second, we focus on those who indicated a 'yes' for the indirect measurement and a 'no' for the direct measurement ($n = 489$, mismatch).

Since Table 3 shows a negligible mismatch where respondents reported an experience on the direct measure but not on the indirect measure (1%, $n = 35$), the logistic regression analysis focuses on the mismatch where respondents reported an experience on the indirect measure but not on the direct measure. In the Appendix (Table A11), we present a sensitivity analysis by using both mismatches—reporting a racist experience on the direct measurement but none on the indirect measurement, and reporting an experience on the indirect but none on the direct measurement—as the dependent variable (1 = mismatch, 0 = match). The results remain nearly the same.

Table 4 displays the results of the logistic regression analysis with the mismatch between the instruments as the dependent variable (1 = reporting an experience on the indirect and none on the direct measurement, 0 = match). The descriptive statistics of the determinants in the main model can be found in the Appendix (Table A12).

The results indicate that younger respondents are more likely to report a mismatch (0.3 percentage points [pp] increase per year of life; cf. column 'AME'), and that respondents with an entrance qualification for university are more like to report a mismatch than the reference group of basic

Table 4

Logistic regression analysis on mismatch between the direct and indirect instruments (main model). (Dollmann et al. (2023) own calculations, not weighted)

	Coefficient	S.E.	AME	S.E.
Male (Reference: female)	0.500***	0.110	0.063***	0.014
Age	–0.022***	0.005	–0.003***	0.001
Anger (Ref: no anger)	0.100	0.139	0.013	0.017
Education (Reference: Basic)				
Intermediate	0.342	0.266	0.038	0.028
Entrance qualification for university	0.566*	0.256	0.066*	0.027
University/college degree	0.435	0.247	0.049	0.025
Non-native	1.763***	0.126	0.279***	0.022
Interview in German	0.007	0.188	0.001	0.024
Interview duration	0.002	0.005	0.000	0.001
Constant	–2.032***	0.385		
N	2706			
Pseudo R2	0.1454			
LR chi2(9)	371.71***			

*p < 0.05, ** p < 0.01, *** p < 0.001

education level (6.6 pp difference). There are no significant differences between the other levels of education. Regarding the significance of the reported events, there is no effect on the likelihood of reporting a mismatch. Regarding gender, men are more likely only to respond positively to the indirect measure (6.3 pp difference). We find a strong and large positive effect regarding non-native status. It is more likely that non-native respondents report a mismatch, i.e., indicating discrimination using the indirect as compared to the direct measure (27.9 pp difference). We find no significant effect for the interview language or for the interview duration.

4   Results of Robustness Checks

Tables A7–A10 in the Appendix display the results of the three robustness checks outlined above, with the main results summarised in the following sections.

4.1   Univariate Distribution

For verbal hostilities, we find that 26% of all respondents report having experienced verbal hostilities, 13% report having experienced racialised exclusion, and 4% experienced physical violence. Experiencing verbal hostilities—or identifying the experience as such—seems to be more likely than racialised exclusion, or physical violence (in that order).

4.2   Mismatch

Regarding the mismatches between verbal hostilities and the direct measure, about 20% of respondents' reports do not align: approximately 16% reported an incident on the indirect measure but not on the direct measure, while 2% showed the opposite pattern. When comparing racialised exclusion with the indirect measure, we find that around 10% of the respondents' reports do not match. Only around 6% report an incident of racialised exclusion but do not report it using the direct scale, whereas for 4% it is the other way around.

When comparing physical violence with the indirect measure, we find that around 10% of the respondents' reports do not match. Around 2% report an incident of physical violence but do not report it using the direct scale, whereas for another 8% it is the other way around.

In summary, our robustness checks show that a mismatch between the direct and indirect measure is mostly driven by verbal hostilities, especially by the fact that respondents do not identify verbal hostilities as being racist.

4.3   Multivariate Analyses

When looking at the mismatch between the dimension 'verbal hostilities' and the direct measure, we also find that males, younger respondents, and non-natives, and those with a higher educational degree—i.e., entrance qualification for university and university/college degree than the reference group of basic education level—are more likely to report a mismatch.

Considering the model for the mismatch between the dimension 'exclusion' and the direct measure, we find two indicators that are associated with a mismatch—male respondents and people who identified the incident as not being significant—are more likely to report a mismatch. The mismatch on the dimension 'physical violence' again shows similar results, but additionally here we see a negative age effect, meaning that younger respondents are less likely to report a mismatch between physical violence and the direct measure. Lastly, as in the main model, the interview language and the interview duration do not significantly explain the mismatch in all robustness checks.

To conclude, our robustness checks all provide similar results when compared to the main model. Only the dimension of exclusion in combination with the direct measure seems to produce somewhat different results.

5   Discussion

In the current article, we test two scales to capture racist experiences—one direct and one indirect—and find significant differences in the prevalence of racist experiences recorded, with the indirect measure showing a 2.5 times higher prevalence than the direct measure. Additionally, we also observe that almost everyone who reported an experience on the direct measure also reported an experience on the indirect measure. This is striking, because the indirect measure offers only a limited set of incidents for respondents to react to. We expected the direct measure to report a higher prevalence, as it could capture any possible experience. We additionally expected that respondents would be triggered by the word 'racism' and would therefore report more incidences, although this is far from what we find. To the contrary, it is the other way around, and the indirect instrument reports higher prevalence rates.

We offer three explanations for why the indirect measure produces a higher prevalence. First, classifying an experience as 'racist' is not straightforward. It requires knowing what specifically qualifies as being racist. Second, being racialised and devalued is distressing, and as a coping strategy, respondents might avoid labelling these experiences as racist. Our robustness checks support both assumptions, showing that mismatches are mainly driven by verbal hos-

tilities. This suggests that more subtle incidents, such as verbal hostilities, are often not identified as racist. Additionally, the mismatch involving physical violence and the direct measure supports the idea that avoiding reporting such incidents may also serve as a coping strategy to avoid reliving the experience. Third, particularly in Germany, scholars observe the tendency that racism is externalised in the sense that racist attitudes and actions are predominantly viewed as particularly right-wing phenomena. Experiences that could be classified as racist might therefore not be viewed as such if the actor is not perceived as belonging to the far right.

The second important finding is that our data shows that two out of three people who reported an incident on the indirect measure did not deem this experience noteworthy in the direct measure. To shed light on this, we employed a logistic regression analysis with this mismatch as the dependent variable. Our results indicate that younger respondents' responses especially are more likely to indicate a mismatch. This is interesting in so far as we expected it to be the other way around. While a mismatch could reflect that respondents have difficulties understanding and applying the term 'racism', we would have expected this to be more likely with older respondents. Considering our results, we now assume that younger respondents seem to be more triggered by such specific events listed in the indirect measure. It is not plausible that every person is equally offended by the provided incidences; young people could be more sensitive than older respondents in this regard.

Third, contrary to our expectations, we do not find a clear education effect. Although we find that respondents with a university entrance qualification are more likely to report a mismatch, respondents with a university degree do not differ statistically from the reference group of those with basic education—although the direction of the non-significant effect again points in the same direction as those with a university entrance qualification. The results therefore do not necessarily provide an indication that one instrument is cognitively more burdensome than the other. Another finding is that non-natives are more likely to report a mismatch than German natives. This finding could again reflect the sensitive nature of racist experiences and, as a coping strategy, affected people might not want to classify incidents as racist.

Fourth, our robustness checks reveal that the reported mismatch between the direct and indirect measure is predominantly driven by the dimension 'verbal hostilities' in the way that people report having experienced verbal hostilities, but they do not report an incident on the direct measure. This is a strong indication that in daily life, verbal hostilities are not marked as racist experiences. This finding is a promising starting point for more nuanced analyses of what people recognise as racism, and how that stands

in contrast to theoretical developments and academic concepts.

In summary, our study supports the assumption that in research on (subjective) racist experiences it is far from trivial to design a valid instrument that speaks to people's everyday lives. Our data points to the conclusion that there is, as yet, no consensus about what types of experiences qualify as racist, and what types do not. As a result, employing a direct instrument increases the risk of under-reporting racist experiences in a population. Additionally, our study points to age-dependent and gendered effects of racist experiences and comes to the conclusion that the underlying target population of the affected respondents differs depending on whether they choose a direct or an indirect instrument. When employing a direct instrument, the distribution will likely overestimate the affectedness of women and older people. Survey practitioners need to take this into consideration when developing survey instruments, otherwise they risk underestimating the real prevalence of racist experiences in a population.

5.1 Limitations and Outlook

Some limitations of this study should be noted. First, we only employed two variables. We therefore have no variation in the wording of the direct and indirect measurements and cannot assess how much of the found differences is due to the wording of the items and how much is due to the concept. Second, we carried out our analyses using a sample of the German population, and it is well known that racist attitudes and actions—and how they are voiced and acted out—differ between societal contexts. We therefore cannot make well-informed statements about how our findings might translate to different contexts. Third, in the academic debate, there is as yet no consensus about the core elements of racism. In the operationalisation of the indirect measure, we rely on a definition by Hall and Bonilla-Silva (Bonilla-Silva, 2003; Hall, 1997), which could be called into question. It is astonishing however, that even when working with Hall's and Bonilla-Silva's narrow approach we find a higher prevalence in the indirect measure.

However, since developing adequate measures for observing racist experiences is still a young and comparatively underdeveloped field, we consider our study to be a valuable starting point for further and more fundamental research. We particularly recommend including the different theories of racism in future research and empirically contrasting a wider range of operationalisations when working with indirect measures. Additionally, more qualitative and more cross-country research is needed to better understand what is classified, from a subjective point of view, as a 'racist experience'. Additionally, we propose that survey

practitioners who want to estimate the prevalence of racist experiences should not employ only one type of instrument but use a wide array of instruments to make sure that all aspects that could qualify as racist are captured.

References

Axt, J. R. (2018). The best way to measure explicit racial attitudes is to ask about them. *Social Psychological and Personality Science*, *9*(8), 896–906.

Bonilla-Silva, E. (2003). Racial attitudes or racial ideology? An alternative paradigm for examining actors' racial views. *Journal of Political Ideologies*, *8*(1), 63–82.

Clair, M., & Denis, J. S. (2015). Racism, sociology of. In *International encyclopedia of the social & behavioral sciences* (pp. 857–863).

DeZIM (2022). *Rassistische Realitäten. Wie setzt sich Deutschland mit Rassismus auseinander? Auftaktstudie zum Nationalen Diskriminierungs- und Rassismusmonitor (NaDiRa). Berlin*

Dollmann, J., Mayer, S. J., Lietz, A., Siegel, M., & Köhler, J. (2022). DeZIM.panel—data for Germany's post-migrant society. *Jahrbücher für Nationalökonomie und Statistik*, *243*(1), 93–108.

Dollmann, J., Jacobsen, J., Mayer, S. J., Lietz, A., Siegel, M., Köhler, J., & Kalter, F. (2023). *Gesellschaftliches Zusammenleben in Deutschland. DeZIM.panel Welle 4. Datensatz, Version: 4.0.0. SUF C*. Berlin: Deutsches Zentrum für Integrations- und Migrationsforschung (DeZIM). https://doi.org/10.34882/dezim.panel.bw1.c.2.0.0.

Elis, J., & Hamidou-Schmidt, H. (2023). Wie rassistische Einstellungen gemessen werden: Rassimuskonzepte und Messinstrumente in quantitativen Verfahren. In S. Polat (Ed.), *Rassismusforschung I* (pp. 129–168). Bielefeld: transcript.

Essed, P. (1990). *Everyday racism: reports from women of two cultures*. Claremont: Hunter House.

Fernández-Reino, M., Di Stasio, V., & Veit, S. (2022). Discrimination unveiled: a field experiment on the barriers faced by muslim women in Germany, the Netherlands, and Spain. *European Sociological Review*, *39*(3), 479–497.

Forrest, J., Blair, K., & Dunn, K. (2021). Racist attitudes, out-groups and the Australian experience. *Australian Journal of Social Issues*, *56*(1), 78–93.

Gran-Ruaz, S., Feliciano, J., Bartlett, A., & Williams, M. T. (2022). Implicit racial bias across ethnoracial groups in Canada and the United States and Black mental health. *Canadian Psychology / Psychologie canadienne*, *63*(4), 608–622.

Hall, S. (1989). Rassismus als ideologischer Diskurs. Le discours idéologique du racism. *Das Argument (Berlin)*, *31*(178), 913–921.

Hall, S. (1997). The spectacle of the 'other'. In S. Hall (Ed.), *Representation: cultural representations and signifying practices* (pp. 223–290). London: SAGE.

Jacobson, C. K. (1985). Resistance to affirmative action: self-interest or racism? *Journal of Conflict Resolution*, *29*(9), 306–329.

Jones, J. M. (1972). *Prejudice and racism*. Menlo Park: Addison-Wesley.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). San Diego: Academic Press.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*(3), 598–609.

Quillian, L. (2006). New approaches to understanding racial prejudice and discrimination. *Annual Review of Sociology*, *32*(1), 299–328.

Sears, D. O. (1988). Symbolic racism. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: profiles in controversy* (pp. 53–84). New York: Plenum Press.

Sears, D. O., & Kosterman, H. R. (2000). Egalitarian values and contemporary racial politics. In D. O. Sear, J. Sidanius & L. Bobo (Eds.), *Racialized politics: the debate about racism in america* (pp. 75–117). Chicago: University of Chicago Press.

Utsey, S. O., & Ponterotto, J. G. (1996). Development and validation of the Index of Race-Related Stress (IRRS). *Journal of Counseling Psychology*, *43*(4), 490–501.