Comprehensive Psychoneuroendocrinology 19 (2024) 100253

Contents lists available at ScienceDirect

Comprehensive Psychoneuroendocrinology

journal homepage: www.sciencedirect.com/journal/comprehensive-psychoneuroendocrinology

Reflections on the study of empathy in a sample of refugees and migrants from Arabic-speaking countries with diverse experiences of war-related trauma

Christiane Wesarg-Menzel [a,b,*], Mathilde Gallistl [a], Michael Niconchuk [c], Veronika Engert [a,b,d,e]

[a] Social Stress and Family Health Research Group, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[b] Institute for Psychosocial Medicine, Psychotherapy and Psychooncology, Jena University Hospital, Friedrich-Schiller University, Jena, Germany
[c] Beyond Conflict, Boston, MA, USA
[d] German Center for Mental Health (DZPG), Germany
[e] Center for Intervention and Research in Adaptive and Maladaptive Brain Circuits Underlying Mental Health (C-I-R-C), Halle-Jena-Magdeburg, Germany

ARTICLE INFO

Keywords:
Empathy
Compassion
Assessment
Study design
Trauma
War
Refugees

ABSTRACT

Empathic abilities are proposed to affect the trajectory from trauma exposure to psychopathology. Yet, studies addressing the role of empathy in refugees with diverse experiences of war-related trauma are lacking. This may relate to missing recommendations on aspects to consider in the planning and execution of such a study. In the present methodological paper, we hence share our experiences in designing and implementing a study on the interrelations of war-related trauma, post-traumatic stress disorder, and empathy in individuals from Arabic-speaking countries who had entered Germany as refugees or migrants. In specific, we reflect on decisions related to the choice of experimental groups and measures of empathy, and describe unanticipated problems encountered during recruitment, screening and testing. Overall, we recommend applying a multi-method approach (i.e., a combination of questionnaire, behavioral and biological measures) to gain a comprehensive picture of the different facets of empathy. Further, we stress the importance to consider that not only refugees, but also migrants may have experienced war-related trauma. Beyond that, we advise to consult individuals of the study population of interest for the translation of instruments, realization of effective recruitment strategies, and to ensure that the testing procedures are sensitive to participants' past experiences and current needs. We hope that sharing these insights will benefit researchers interested in conducting basic and intervention research aimed at improving the mental health of individuals exposed to war-related trauma.

Empathic abilities have been suggested to influence an individual's risk to develop mental health problems in the context of stress [1]. Despite a surge in research on the interrelations of empathy, trauma and health, studies addressing the role of empathy in refugees with diverse experiences of war-related trauma are currently lacking. This may be due to several challenges that come along with the planning and implementation of such a study, including the choice of experimental groups, the assessment of empathy, and recruitment of participants. As of yet, no recommendations exist on how to deal with these issues. Beyond that, the existing studies on empathy and trauma have primarily been conducted with individuals from Western societies or China, with trauma contexts comprising natural catastrophes [1,2] or childhood trauma ([3,4]; for meta-analysis, see [5]). However, the methodology employed in these studies has limited transferability to a study with individuals living in or fleeing from areas of war. In this methodological paper, we hence share our experiences in designing and implementing a study on the interrelations of empathy, war-related trauma, and post-traumatic stress disorder in individuals from Arabic-speaking countries who had entered Germany as refugees or migrants. We hope that sharing these insights will benefit researchers interested in conducting basic and intervention research aimed at improving the mental health of individuals exposed to war-related trauma. Although we engaged a very specific population of refugees and migrants, we are convinced that offering insight into the background factors of our study will be useful for future studies with other refugee populations.

In the first part of this manuscript, we will define the highly complex

* Corresponding author. Institute for Psychosocial Medicine, Psychotherapy and Psychooncology, Jena University Hospital, Friedrich-Schiller University, Stoystr. 3, 07743 Jena, Germany.
E-mail address: christiane.wesarg-menzel@med.uni-jena.de (C. Wesarg-Menzel).

https://doi.org/10.1016/j.cpnec.2024.100253
Received 19 December 2023; Received in revised form 19 July 2024; Accepted 22 July 2024
Available online 23 July 2024

construct of "empathy" and describe its relevance for health. Then, we will provide a brief overview of how empathy has been assessed in the past, spanning questionnaires, behavioral tasks, and biological measures. Readers interested in a more detailed account of these assessment methods are referred elsewhere [6–8]. In the second part of this manuscript, we will draw on experience from our recent study to discuss challenges in designing and implementing the study, supplemented by ways on how to overcome said challenges. In specific, we will reflect on decisions related to the choice of the experimental groups and measures of empathy, and describe unanticipated problems encountered during recruitment, screening and testing.

1. Definition and relevance of empathy

Broadly speaking, empathy is a multi-faceted concept that refers to the ability to share and understand the internal states of others [9]. Although many definitions of empathy exist [10,11], there is the widely-shared view that it contains at least three components. First, empathy involves a cognitive component, also referred to as **theory of mind (ToM)**, mentalizing, or perspective taking, which describe the capacity to understand others' thoughts, intentions or emotions [12,13]. Second, empathy involves an affective component, known as **experience sharing** or emotion contagion, whereby individuals vicariously feel others' emotional states [14,15]. Third, empathy involves a motivational component, which has also been termed **compassion** or empathic concern, describing a warm sense of care involving the wish for the other's well-being [13].

Empathy plays a fundamental role in social functioning, as it facilitates interpersonal understanding, thereby enabling helping behavior and cooperation [16,17]. Although one may believe that high levels of empathy are generally beneficial for an individual, a more differentiated view on the components of empathy reveals rather distinct associations with behavioral and socioemotional outcomes, suggesting that empathy can have potential costs [9,18,19]. To illustrate, whereas compassion appears to be a highly reliable predictor of helping [20], experience sharing can also inhibit helping behavior by inducing personal distress [13,21]. With regards to health, lower levels of experience sharing and higher levels of compassion have been identified as protective factors against developing mental health problems in the context of stress [1,22,23].

Previous research has shown that experiences of empathy can be changed through trainings such as meditation practices, yielding long-term benefits for health [24]. To exemplify, the systematic training of compassion was shown to lead to reductions in acute stress reactivity [25] and chronic stress load in healthy participants [26–28]. The malleability of empathic processes has recently been used in the service of interventions aimed at improving the mental health of individuals with a history of trauma exposure [29,30]). Specifically adapted to the needs of refugees, a 9-week mindfulness-based, trauma-sensitive group intervention involving compassion practices was shown to lead to reduced rates and symptom severity of PTSD, depression, anxiety and multi-morbidity in Eritrean asylum-seekers residing in Israel [31].

2. Measurement of empathy

2.1. Questionnaires

Self-report questionnaires have the longest tradition in empathy research, dating back to the 1940s (e.g., [32]). Since then, there has been increasing interest in the use and development of new empathy questionnaires. Commonly employed examples that measure dispositional trait empathy are the "Interpersonal Reactivity Index" (IRI; [33]), the "Empathy Quotient" [34], the "Toronto Empathy Questionnaire" [35], and the "Questionnaire of Cognitive and Affective Empathy" [36], for all of which several translations and validations exist (see [6]). To do justice to the multi-faceted nature of empathy, different scales commonly reflect the different components of empathy. For example, in the IRI [33], the scale "perspective taking" relates to ToM, "personal distress" to experience sharing (with a focus on negative experiences), and "empathic concern" to compassion. While a major advantage of the use of questionnaires is their ease of application, criticism has been raised regarding their proneness to social desirability bias [37]. Moreover, studies have revealed low convergent validity between different empathy questionnaires, and no instrument can currently be regarded as the gold standard [6].

2.2. Behavioral measures

With the aim of gaining more objective and ecologically valid measures of empathy, researchers have developed behavioral tasks, which include the performance on tests and evaluation of experimental stimuli [7]. Several tasks focus on ToM specifically. For instance, the "Faux-Pas Test" [38] requires participants to detect if a person made a "faux-pas" in a conversation (i.e., the person says something he or she should not have said, not realizing the words' inappropriateness, which could hurt the listener's feelings). In "Strange Stories" [39], participants need to provide context-appropriate mental state explanations for a character's behaviors. The specificity of many of the commonly employed ToM tasks has recently been questioned (see [40]), as they may not actually require the participant to represent another's mental state. This is of particular concern regarding the frequently used "Reading the Mind in the Eyes Test" [41], in which participants have to infer the emotional state of others from their facial expression or gaze, yielding a measure of emotion recognition rather than ToM [40,42].

There are few behavioral tasks that combine measures of ToM with experience sharing and compassion. To exemplify, in a paradigm called the "EmpaToM" [43], participants view videos of actors reporting shortly on their emotionally negative or neutral experiences. After each video, participants are asked to rate their own affect as a measure of experience sharing, and the extent to which they felt compassion for the person in the video. They are further prompted to answer a question requiring either a ToM-inference or factual reasoning on the content of the previous video. Other computational tasks yielding separate measures of the different empathy components are the "Multifaceted Empathy Test" [44] and the "Tübinger Empathy Test" [45]. It should be noted that studies have found poor convergence between self-reported and behavioral measures of empathy [46,47]—a problem that appeared most severe for ToM measures [48]. Critically, these findings raise serious concerns about the common view that self-reported empathy can be seen as a proxy for empathic abilities and behaviors shown in daily life. Yet, behavioral measures of empathy suffer from limitations as well. Although behavioral measures of acute experience sharing and compassion are likely more ecologically valid than questionnaire measures focusing on trait empathy, they cannot be considered completely objective given their continued reliance on self-report.

2.3. Biological measures

Encountering the limitations of both questionnaire and behavioral measures of empathy, cognitive neuroscience has brought up new paradigms that allow for a more ecologically valid and objective assessment of empathy. A common experimental setting to elicit empathy are computational tasks in which participants view pictures or videos of emotional content, such as a person in pain or in conflict with someone else. Although these paradigms make use of stimuli that approximate real-life settings, the extent to which they elicit empathy may be limited because the shown events are not actually happening in real-life. Hence, a stronger trigger of empathy can be achieved by asking the participant to observe a distressed person in real-life. The "Empathic Trier Social Stress Test" [49] makes use of this principle. In this setting, one participant undergoes the classic "Trier Social Stress Test" (TSST; [50]), which requires giving a speech and performing mathematical operations

in front of an emotionally neutral evaluation committee. In the E-TSST, a second participant observes this procedure. An abundant number of studies has shown that the TSST reliably elicits stress on a self-report and physiological level in the "target" [51]. Moreover, a growing body of literature suggests that a considerable number of "observers" demonstrate a stress response as well, apparent through increases in the release of cortisol [49,52] and autonomic arousal [49]. This second-hand stress experience has been termed "empathic stress" [49] or "empathic resonance of stress" [52]. Besides cortisol as a measure of the hypothalamic-pituitary-adrenal (HPA) axis, as well as salivary alpha-amylase (sAA), heart rate, and heart rate variability as measures of the autonomic nervous system, a diversity of other approaches has been applied to capture physiological correlates of empathy. These approaches include functional brain imaging, electroencephalography (EEG), facial electromyography (EMG), and electrocardiography (ECG; for summary, see [7]).

3. Measurement of empathy in refugees

In contrast to a vast literature on empathy *towards* refugees, only little attention has been spent on empathy *in* refugees. To our knowledge, only one study by Aragona and colleagues (2020) assessed different facets of empathy with the IRI [33] in 40 male participants who came from different African countries and lived in Italy. The authors compared two groups: A clinical group of 20 asylum seekers and refugees who had received a diagnosis of PTSD, and a non-clinical control group of 20 students of theology. While no group differences were found in *Perspective Taking* and *Empathic Concern*, asylum seekers and refugees with PTSD reported higher levels of *Personal Distress* compared to the control group [53].

4. Challenges in studying empathy in individuals with war-related trauma: an experience-based report

In 2018, we conceptualized a study that aimed to investigate the influences of traumatic war experiences on empathic processes, including empathic stress resonance. We here briefly outline the study design, followed by a detailed discussion of the challenges that were encountered while designing and implementing the study, placing a particular focus on the assessment of empathy.

4.1. Study design

Our study was approved by the Ethics Board of the medical faculty of Leipzig University, Germany (ethics number: 405/18-ek). The research design envisaged to recruit refugees from Arabic-speaking countries with war-related trauma, and migrants from Arabic-speaking countries with no trauma experiences as a control group. Following the definition released by the United Nations High Commissioner for Refugees [54], which is based on the 1951 Refugee Convention [55], individuals who indicated that they had been *forced* to flee their home country and seek safety in another country were considered as refugees. We also included asylum seekers in our operationalization of the refugee group, who by definition had not yet been legally recognized as refugees but were waiting on a decision regarding their asylum claim [54]. In contrast, individuals who indicated that they had *chosen* to leave their home to work, study, or join family in a new country were considered as migrants [54]. While refugees are unable to return to their own country because of conflict or feared persecution, migrants can return home without risking their life or freedom [54].[1] Further inclusion criteria for both groups were living in Germany for at least six months, speaking Arabic as native language and German at an intermediate (B1) level, and being aged between 20 and 40 years. This restricted age range was chosen due to the influence of age and female hormonal status (e.g., menopause) on HPA axis activity [56], which we aimed to assess in one of our experimental paradigms (the E-TSST). Refugees were included if they reported a war-related trauma (e.g., exposure to violence, persecution).[2] Exclusion criteria for refugees were 1) exposure to any major non-war-related trauma (including maltreatment, severe accident, and natural disaster), 2) the presence of a diagnosed psychiatric disorder during the last two years except for PTSD and depression, and 3) the presence of severe depressive symptoms during the last four weeks as confirmed by scores $\geq$5 on the depression section of the Structured Clinical Interview for DSM-IV Personality Disorders [57]. Migrants were included if they were never exposed to major trauma, and did not have a diagnosed psychiatric disorder within the past two years. These criteria were assessed in a structured telephone screening by Arabic-speaking student assistants.

Eligible participants were invited twice to the lab. In a first computer-based testing session, they were asked to fill out questionnaires measuring trauma exposure, mental health, and self-reported empathy. They also completed the EmpaToM [43] as a behavioral measure of different facets of empathy (ToM, experience sharing, and compassion). In a second session, participants took part in a real-life empathy-eliciting paradigm, the E-TSST [49], during which biological measures reflective of empathic processes were obtained. To this end, German-speaking participants were invited to undergo the stress task as targets of firsthand stress, while a refugee or migrant observed. All participants received financial compensation in relation to the time spent for study participation.

4.2. Challenges in study conceptualization

In the phase of study conceptualization, we encountered several issues that required extensive discussions, for which knowledge of Southwest Asian cultures (often called 'the Arab world'), as well as refugees' experiences and needs were critical. In all of the decisions made, we heavily relied on the support of several Arabic-speaking student assistants who were recruited specifically for the study. We also reached out to a local refugee center in Leipzig for additional advice.

Choice of trauma group. From 2014 to 2021, Syria was in first place among the nationalities with the largest number of asylum applicants in Germany [58]. This forced displacement is attributable to the Syrian Civil War, which has been ongoing since 2011, and caused exposure to extreme traumatic events to millions of Syrian civilians [59]. We chose to focus on this specific refugee population whose native language is commonly Arabic (with several Arabic dialects being used in everyday life). To facilitate recruitment, we decided to also include refugees from other predominantly Arabic-speaking countries (e.g., from Iraq as the country of origin with the second largest volume of asylum applicants in Germany in 2018 [60]). We did not broaden inclusion to other native languages spoken by refugees to ensure feasibility of the study (taking the translation of instruments and experimenters' language skills into account). We further did not screen for ethnicity or cultural affiliation. It is important to note that the recruitment of Arabic-speaking refugees

[1] We assigned participants to either the refugee or migrant group based on the reason they indicated for leaving their home country at that respective time. Please note that an individual who left the country as a migrant can become a refugee, because it becomes unsafe to return to the home country. In that specific case, a participant of our study would have nevertheless been assigned to the migrant group.

[2] Note that exposure to war-related trauma is no prerequisite to be considered a refugee.

may imply cultural homogeneity amongst our population of interest. This is, in fact, not the case, since countries in which Arabic is the official language are multi-ethnic, multi-language, and often religiously diverse.

Choice of control group. Since we intended to identify the distinct influence of war-related trauma on empathic processes, we planned to compare a group of individuals with war-related trauma who had not been exposed to any other trauma (e.g., maltreatment, severe accident, natural disaster) to trauma-free individuals. We first considered to recruit German participants for the trauma-free control group, but this would have limited comparability between both groups due to various dissimilarities. Beyond that, we feared that under specific circumstances, research findings might be misinterpreted with political relevance. To be specific, what if we would have observed that trauma-exposed individuals from Arabic-speaking countries demonstrated lower levels of empathy as compared to Germans without trauma exposure? Although this finding could have been interpreted as an effect of trauma exposure, it could have also been misinterpreted to suggest that individuals from certain cultural backgrounds are less empathic, potentially fostering xenophobia. Due to these reasons, we decided to have two groups with similar cultural background, one composed of refugees with war-related trauma, and one of trauma-free migrants also from Arabic-speaking countries.

Choice of empathy measures. We decided to apply the IRI [33] as a multi-faceted questionnaire measuring empathy, because it is the most common one in the field of trauma research, and hence enables comparability with former studies. Since no validated Arabic translation of the IRI existed at the time of study planning, it was translated into Modern Standard Arabic by a bilingual, native Arabic speaker and then back translated by a second bilingual individual to ensure linguistic equivalence [61]. Yet, the use of translated, unvalidated scales turned out to be a major limitation of our study, since we found low reliabilities in all four IRI scales. More research will need to be devoted to the translation and validation of questionnaires to enable the widespread use among individuals exposed to war-related trauma.

In a similar vein, at the time of study conceptualization, the EmpaToM [43] had only recently been developed and validated for German-speaking participants, and no other language version was available. Although it would have been the gold standard to rely on a translated and culturally adapted version of the EmpaToM in our study, the development of such would have been a distinct and time-consuming endeavor. Hence, the translation of the already existent videos as well as the EmpaToM instructions were displayed as written Arabic subtitles. Again, two bilingual Arabic-German speakers were involved in the translation and back-translation process [61]. We acknowledge that one limitation of this procedure is the unknown reliability and validity of the adapted task version. Further, empathic responses elicited in our participants might have been influenced by cultural dissimilarity with the German-speaking individuals seen in the videos (see e.g., [62]) and potentially, given the diversity within spoken and written Arabic between different regions, unfamiliarity with the words chosen for subtitles. Intriguingly, a cross-study comparison including data from one of our other lab studies [63] revealed significantly higher levels of EmpaToM-reported compassion in our participants from Arabic-speaking countries ($M = 63.22$, $SD = 12.28$) relative to healthy German participants of comparable age range ($M = 57.14$, $SD = 13.89$; $t = 3.58$; $p < 0.001$). This finding is in line with previous reports on higher self-reported compassion in traumatized individuals than in non-traumatized controls [3,64](, and suggests that cultural dissimilarity was not interfering with our participants' basic ability to feel compassion for others.

Conceptualization of the E-TSST. Our participants also took part in the E-TSST: refugees and migrants from Arabic-speaking countries as passive observers, and German native speakers as directly stressed targets. We had extensive debates on the composition of the target group.

First, when considering inclusion criteria for the target group, we again assumed that cultural similarity might enhance empathic responses in the observers [62]. Based on this reasoning, a logical step would have been to recruit non-traumatized individuals from Arabic-speaking countries as targets. However, these individuals would have additionally qualified for the migrant control group, for which we expected recruitment difficulties; specifically that it would be challenging to find a sufficient number of trauma-free individuals (which indeed, it was).

Second, we anticipated language issues: If the TSST had been conducted in German, targets from Arabic-speaking countries would have been required to have proficient German skills. On the other hand, if the TSST had been conducted in Arabic, we would have needed a much higher number of student research assistants with proficient Arabic skills to conduct the testing sessions. Furthermore, spoken Arabic differs not only from country to country, but between regions within most Arabic-speaking countries. Due to pragmatic reasons, we hence decided to conduct the TSST in German with German native speakers, for whom we did not expect or encounter recruitment difficulties. Of note, there were no specific criteria related to the targets' physical appearance, meaning that ethnicity did not play a role in the screening of targets.

4.3. Challenges in study implementation

Recruitment of Arabic-speaking participants. Our Arabic-speaking student assistants were responsible for the recruitment of refugees and migrants. The most effective recruitment strategies were the use of the student assistants' personal networks, and word of mouth among participants' personal networks. Slightly less effective were posting advertisements in social media groups, and talking to people around university campus and in Arabic community areas while handing over flyers. Almost ineffective were our usual recruitment strategies, such as hanging flyers without personal contact, or postings on electronic community billboards. Without a detailed knowledge of the Arabic-speaking refugee community in Leipzig, we would have not succeeded to recruit our study sample.

Overall, it turned out to be more difficult to recruit female Arabic-speaking participants as compared to males, yielding a ratio of around 1:4. We can only speculate that reasons for this imbalance are the higher number of male than female refugees applying for asylum in Germany [65], and potential disproportionate level of childcare and home responsibilities among females which limits their ability to spend time away from home. In general, it was more efficient if our female (rather than male) Arabic-speaking student assistant approached other women for study participation.

Screening of participants from Arabic-speaking countries. Our Arabic-speaking student assistants encountered several challenges during the telephone screening of migrants and refugees. For instance, it was difficult to assess the presence of former psychiatric disorders, given that refugees and migrants had often not received a diagnosis or treatment despite experiencing mental health problems. Further, in a considerable number of instances, the telephone screening was burdensome for student assistants as potential participants reported on suicidality in the past. In these instances, we referred to a local psychological counselling center. In some cases, there were also communication difficulties between our student assistants and participants due to different dialects of the Arabic language.

Since we aimed to assess cortisol levels in the E-TSST, an initial exclusion criterion had been regular smoking for all participants. After seeing that many potential participants from Arabic-speaking countries reported daily cigarette use, we dropped this criterion for the sake of practicality. Although we did not anticipate this challenge in recruitment, it is perhaps not surprising given that prevalence rates of smoking are generally high in many Middle Eastern countries [66], and exposure to traumatic life events was shown to increase the risk of smoking [67].

Trauma-freeness in the migrant control group. To ensure trauma-freeness in the migrant control group, we explicitly asked for life-time trauma exposure in the lengthy telephone interview conducted before

study inclusion. However, we needed to discover later, based on analyses of a more detailed trauma instrument, the Harvard Trauma Questionnaire [68], that also migrants were exposed to considerable war-related traumatic experiences. We hence learned that trauma exposure is also prevalent among migrants, although to a less severe extent than among refugees, and that the recruitment of a trauma-free Arabic-speaking migrant group of considerable size is extremely difficult.

We assume that migrants did not voluntarily withhold information on war-related trauma. Rather, the telephone interview may have been too unspecific and unprompted to trigger correct responses. Seeing events spelled out in the Harvard Trauma Questionnaire [68] may have made it much more obvious for affected individuals to identify their traumatic experiences as such, or have activated memories that were not accessible spontaneously.

Testing specificities. While testing procedures ran smoothly, two organizational challenges deserve mentioning. One, because we wanted participants to feel at ease when at the laboratory, we assured that there was always a native Arabic speaker of the same sex as our refugee or migrant participant present for the duration of the testing session. This was of particular relevance when the experimenter placed the ECG belt around the participant's chest. Clearly, given this setup, we would have greatly profited from a larger pool of Arabic-speaking research assistants to allow for more flexibility in scheduling the testing sessions. Especially for smaller labs, this is an important additional expense needing prior planning and recruitment of financial resources. Also, whenever a participant with disclosed trauma experience was present, we assured that a psychiatrist was on call to professionally deal with potential recurrence of traumatic stress triggered by completion of the detailed trauma questionnaire. In the end, psychiatric assistance was never needed. Future studies may also consider allocating a specific room as private space for self-care activities of participants if desired, or having a culturally-relevant supportive person on site who is not a mental health professional (e.g., a priest or imam).

There were further noteworthy challenges that should be taken into consideration in future study set-ups. A recurrent concern from our refugee participants was whether we would be in contact with the refugee registration office and forward information from questionnaire assessments. During informed consent, it was hence of utmost importance to explain confidentiality and that study participation had no influence on participants' residency status. Nevertheless, many participants reported back to our Arabic-speaking student assistants that they were afraid of indicating "politically incorrect" information in the questionnaires, which could later be used against them.

During the testing session, participants had many questions about the interpretation of questionnaire items, which needed additional explanation from our research assistants. These enquiries were also related to the complexity of the Arabic language, and particularly migrants, who had been living in Germany for longer time periods, experienced comprehension difficulties. Another concern of our refugee participants related to the payment for study participation. Many preferred to receive financial compensation as cash, as any income into their bank account might have been deducted from their monthly assistance payments from the government.

5. Conclusion

Studying empathy in the context of war and trauma is an important yet challenging endeavor. Although several instruments exist for the assessment of empathy at the self-report and behavioral level, these instruments have mainly been developed in Western societies, and validated translations for other languages, including Arabic, are still lacking. Biological measures of empathy elicited by the observation of another person in distress not only circumvent such language barriers, but also enable a more objective and ecologically valid assessment of empathy. Given the lack of convergence among various empathy measures, it remains advisable to apply a multi-method approach to provide a comprehensive picture of the different facets of empathy. Further, it is indispensable to consult native Arabic speakers in the phases of study conceptualization and implementation. This will not only facilitate the translation of instruments and realization of effective recruitment strategies, but also ensure that the study design and procedures are sensitive to past experiences and current needs of the specific study population.

Funding

This study was funded by the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, and Beyond Conflict.

CRediT authorship contribution statement

Christiane Wesarg-Menzel: Writing – original draft, Conceptualization. **Mathilde Gallistl:** Writing – review & editing, Conceptualization. **Michael Niconchuk:** Writing – review & editing, Conceptualization. **Veronika Engert:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

none.

Acknowledgements

We cordially thank our Arabic-speaking student assistants (Abdel Bakr, Ghada Mohammad, Ahmad Shaban, and Karim Zagha) for their indispensable support with conducting the study. Additionally, we thank Ahmad Shaban for his conceptual input to this paper.

References

[1] I.A. Cristea, E. Legge, M. Prosperi, M. Guazzelli, D. David, C. Gentili, Moderating effects of empathic concern and personal distress on the emotional reactions of disaster volunteers, Disasters 38 (4) (2014) 740–752, https://doi.org/10.1111/disa.12075.
[2] M. Nagamine, J. Shigemura, T. Fujiwara, F. Waki, M. Tanichi, T. Saito, K. Shimizu, The relationship between dispositional empathy, psychological distress, and posttraumatic stress responses among Japanese uniformed disaster workers: a cross-sectional study, BMC Psychiatr. 18 (2018) 1–7.
[3] D.M. Greenberg, S. Baron-Cohen, N. Rosenberg, P. Fonagy, P.J. Rentfrow, Elevated empathy in adults following childhood trauma, PLoS One 13 (10) (2018) e0203886.
[4] M. Parlar, P. Frewen, A. Nazarov, C. Oremus, G. MacQueen, R. Lanius, M. C. McKinnon, Alterations in empathic responding among women with posttraumatic stress disorder associated with childhood trauma, Brain and behavior 4 (3) (2014) 381–389.
[5] H. Zhang, X. Gao, Y. Liang, Q. Yao, Q. Wei, Does child maltreatment reduce or increase empathy? A systematic review and meta-analysis, Trauma Violence Abuse 25 (1) (2024) 166–182.
[6] F.F. de Lima, F. de Lima Osório, Empathy: assessment instruments and Psychometric quality - a systematic literature review with a meta-analysis of the past ten years, Front. Psychol. 12 (2021) 781346, https://doi.org/10.3389/fpsyg.2021.781346.
[7] D.L. Neumann, R.C. Chan, G.J. Boyle, Y. Wang, H.R. Westbury, Measures of empathy: self-report, behavioral, and neuroscientific approaches, Measures of personality and social psychological constructs (2015) 257–289.
[8] G. Sesso, G.E. Brancati, P. Fantozzi, E. Inguaggiato, A. Milone, G. Masi, Measures of empathy in children and adolescents: a systematic review of questionnaires, World J Psychiatry 11 (10) (2021) 876–896, https://doi.org/10.5498/wjp.v11.i10.876.
[9] E. Weisz, M. Cikara, Strategic regulation of empathy, Trends Cogn Sci 25 (3) (2021) 213–227, https://doi.org/10.1016/j.tics.2020.12.002.
[10] C.D. Batson, These things called empathy: eight related but distinct phenomena, in: The Social Neuroscience of Empathy, The MIT Press, 2013, pp. 3–16.
[11] B.M.P. Cuff, S.J. Brown, L. Taylor, D.J. Howat, Empathy: a review of the concept, in: Emotion Review, vol. 8, SAGE Publications Ltd, 2016, pp. 144–153, https://doi.org/10.1177/1754073914558466, 2.
[12] C. Frith, U. Frith, Theory of mind, Curr. Biol. 15 (17) (2005) R644–R646, https://doi.org/10.1016/j.cub.2005.08.041.
[13] T. Singer, O.M. Klimecki, Empathy and compassion, Curr. Biol. 24 (18) (2014) R875–R878, https://doi.org/10.1016/j.cub.2014.06.054.

[14] F. de Vignemont, T. Singer, The empathic brain: how, when and why? Trends Cogn Sci 10 (10) (2006) 435–441, https://doi.org/10.1016/j.tics.2006.08.008.
[15] N. Eisenberg, R.A. Fabes, Empathy: conceptualization, measurement, and relation to prosocial behavior, Motiv. Emot. 14 (2) (1990) 131–149, https://doi.org/10.1007/BF00991640, 1990 14:2.
[16] M. Allemand, A.E. Steiger, H.A. Fend, Empathy development in adolescence predicts social competencies in adulthood, J. Pers. 83 (2) (2015) 229–241, https://doi.org/10.1111/jopy.12098.
[17] N.M. McDonald, D.S. Messinger, The development of empathy: how, when, and why. Moral Behavior and Free Will: A Neurobiological and Philosophical Approach, 2011, pp. 333–359.
[18] P. Bloom, Against Empathy: the Case for Rational Compassion, Random House, 2017.
[19] C.D. Cameron, C.A. Hutcherson, A.M. Ferguson, J.A. Scheffer, E. Hadjiandreou, M. Inzlicht, Empathy is hard work: people choose to avoid empathy because of its cognitive costs, J. Exp. Psychol. Gen. 148 (6) (2019) 962.
[20] H.Y. Weng, A.S. Fox, H.C. Hessenthaler, D.E. Stodola, R.J. Davidson, The role of compassion in altruistic helping and punishment behavior, PLoS One 10 (12) (2015) e0143794.
[21] C.D. Batson, J. Fultz, P.A. Schoenrade, Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences, J. Pers. 55 (1) (1987) 19–39, https://doi.org/10.1111/j.1467-6494.1987.tb00426.x.
[22] H. von Harscher, N. Desmarais, R. Dollinger, S. Grossman, S. Aldana, The impact of empathy on burnout in medical students: new findings, Psychol. Health Med. 23 (3) (2018) 295–303, https://doi.org/10.1080/13548506.2017.1374545.
[23] M. Zhang, S. Wang, Z. Wang, X. Peng, W. Fei, Y. Geng, T. Zhang, Associations of affective and cognitive empathy with depressive symptoms among a sample of Chinese college freshmen, J. Affect. Disord. 292 (2021) 652–659, https://doi.org/10.1016/j.jad.2021.05.111.
[24] E. Weisz, J. Zaki, Empathy building interventions: a review of existing work and suggestions for future directions, The Oxford handbook of compassion science (2017) 205–217.
[25] V. Engert, B.E. Kok, I. Papassotiriou, G.P. Chrousos, T. Singer, Specific reduction in cortisol stress reactivity after social but not attention-based mental training, Sci. Adv. 3 (10) (2017) e1700495, https://doi.org/10.1126/sciadv.1700495.
[26] V. Engert, K. Hoehne, T. Singer, Specific reduction in the cortisol awakening response after socio-affective mental training, Mindfulness 14 (3) (2023) 681–694.
[27] L.M.C. Puhlmann, S.L. Valk, V. Engert, B.C. Bernhardt, J. Lin, E.S. Epel, P. Vrticka, T. Singer, Association of short-term change in leukocyte telomere length with cortical thickness and outcomes of mental training among healthy adults: a randomized clinical trial, JAMA Netw. Open 2 (9) (2019) e199687, https://doi.org/10.1001/jamanetworkopen.2019.9687.
[28] L.M.C. Puhlmann, P. Vrticka, R. Linz, T. Stalder, C. Kirschbaum, V. Engert, T. Singer, Contemplative mental training reduces hair glucocorticoid levels in a randomized clinical trial, Psychosom. Med. 83 (8) (2021) 894–905, https://doi.org/10.1097/PSY.0000000000000970.
[29] D.E. Hinton, R.A. Ojserkis, B. Jalal, S. Peou, S.G. Hofmann, Loving-kindness in the treatment of traumatized refugees and minority groups: A typology of mindfulness and the nodal network model of affect and affect regulation, Journal of Clinical Psychology 69 (8) (2013) 817–828.
[30] M. Müller-Engelmann, C. Schreiber, S. Kümmerle, T. Heidenreich, U. Stangier, R. Steil, A trauma-adapted mindfulness and loving-kindness intervention for patients with PTSD after interpersonal violence: A multiple-baseline study, Mindfulness 10 (6) (2019) 1105–1123.
[31] A. Aizik-Reebs, K. Yuval, Y. Hadash, S. Gebreyohans Gebremariam, A. Bernstein, Mindfulness-based trauma recovery for refugees (MBTR-R): randomized waitlist-control evidence of efficacy and safety, Clin. Psychol. Sci. 9 (6) (2021) 1164–1184.
[32] R.F. Dymond, A scale for the measurement of empathic ability, J. Consult. Psychol. 13 (2) (1949) 127–133, https://doi.org/10.1037/h0061728.
[33] M.H. Davis, A multidimensional approach to individual differences in empathy, JSAS Catalog of Selected Documents in Psychology 10 (85) (1980).
[34] S. Baron-Cohen, S. Wheelwright, The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences, J. Autism Dev. Disord. 34 (2) (2004) 163–175, https://doi.org/10.1023/b:jadd.0000022607.19833.00.
[35] R.N. Spreng, M.C. McKinnon, R.A. Mar, B. Levine, The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures, J. Pers. Assess. 91 (1) (2009) 62–71, https://doi.org/10.1080/00223890802484381.
[36] R.L. Reniers, R. Corcoran, R. Drake, N.M. Shryane, B.A. Vollm, The QCAE: a questionnaire of cognitive and affective empathy, J. Pers. Assess. 93 (1) (2011) 84–95, https://doi.org/10.1080/00223891.2010.528484.
[37] B.J. Lovett, R.A. Sheffield, Affective empathy deficits in aggressive children and adolescents: a critical review, Clin. Psychol. Rev. 27 (1) (2007) 1–13, https://doi.org/10.1016/j.cpr.2006.03.003.
[38] V.E. Stone, S. Baron-Cohen, R.T. Knight, Frontal lobe contributions to theory of mind, J. Cognit. Neurosci. 10 (5) (1998) 640–656.
[39] F.G. Happé, An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults, J. Autism Dev. Disord. 24 (2) (1994) 129–154, https://doi.org/10.1007/BF02172093.
[40] F. Quesque, Y. Rossetti, What do theory-of-mind tasks actually measure? Theory and practice, Perspect. Psychol. Sci. 15 (2) (2020) 384–396, https://doi.org/10.1177/1745691619896607.
[41] S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, I. Plumb, The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism, J Child Psychol Psychiatry 42 (2) (2001) 241–251. https://www.ncbi.nlm.nih.gov/pubmed/11280420.
[42] B.F.M. Oakley, R. Brewer, G. Bird, C. Catmur, Theory of mind is not theory of emotion: a cautionary note on the Reading the Mind in the Eyes Test, J. Abnorm. Psychol. 125 (6) (2016) 818–823, https://doi.org/10.1037/abn0000182.
[43] P. Kanske, A. Böckler, F.M. Trautwein, T. Singer, Dissecting the social brain: introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and theory of mind, Neuroimage 122 (2015) 6–19, https://doi.org/10.1016/j.neuroimage.2015.07.082.
[44] J. Foell, S.J. Brislin, L.E. Drislane, I. Dziobek, C.J. Patrick, Creation and validation of an English-language version of the multifaceted empathy test (MET), J. Psychopathol. Behav. Assess. 40 (2018) 431–439.
[45] A.S. Kimmig, D. Wildgruber, S.U. Wendel, I. Sundstrom-Poromaa, B. Derntl, Friend vs. Foe: cognitive and affective empathy in women with different hormonal states, Front. Neurosci. 15 (2021) 608768, https://doi.org/10.3389/fnins.2021.608768.
[46] B.A. Murphy, S.O. Lilienfeld, Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance, Psychol. Assess. 31 (8) (2019) 1062.
[47] C.S. Sunahara, D. Rosenfield, T. Alvi, Z. Wallmark, J. Lee, D. Fulford, B.A. Tabak, Revisiting the association between self-reported empathy and behavioral assessments of social cognition, J. Exp. Psychol. Gen. 151 (12) (2022) 3304–3322, https://doi.org/10.1037/xge0001226.
[48] S.A. Grainger, K.T. McKay, J.C. Riches, R.J. Chander, R. Cleary, K.A. Mather, N. A. Kochan, P.S. Sachdev, J.D. Henry, Measuring empathy across the adult lifespan: a comparison of three assessment types, Assessment 30 (6) (2023) 1870–1883, https://doi.org/10.1177/10731911221127902.
[49] V. Engert, F. Plessow, R. Miller, C. Kirschbaum, T. Singer, Cortisol increase in empathic stress is modulated by emotional closeness and observation modality, Psychoneuroendocrinology 45 (2014) 192–201, https://doi.org/10.1016/j.psyneuen.2014.04.005.
[50] C. Kirschbaum, K.M. Pirke, D.H. Hellhammer, The 'Trier Social Stress Test' - a tool for investigating psychobiological stress responses in a laboratory setting, Neuropsychobiology 28 (1–2) (1993) 76–81, https://doi.org/10.1159/000119004.
[51] S.S. Dickerson, M.E. Kemeny, Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research, Psychol. Bull. 130 (3) (2004) 355–391, https://doi.org/10.1037/0033-2909.130.3.355.
[52] T.W. Buchanan, S.L. Bagley, R.B. Stansfield, S.D. Preston, The empathic, physiological resonance of stress, Soc. Neurosci. 7 (2) (2012) 191–201, https://doi.org/10.1080/17470919.2011.588723.
[53] M. Aragona, A. Petta, F. Kiaris, E. Begotaraj, C. Lai, G.F. Spitoni, The empathic migrant: empathy is preserved in African refugees with PTSD, Dialogues Philos., Ment. Neuro Sci. 13 (2) (2020).
[54] United Nations High Commissioner for Refugees (UNHCR), UNHCR global website, Who we protect: Refugees (2023, December 8). https://www.unhcr.org/about-unhcr/who-we-protect/refugees.
[55] United Nations, Convention Relating to the Status of Refugees, 1951, 189 U.N.T.S. 150, entered into force April 22, 1954).
[56] C. Kirschbaum, B.M. Kudielka, J. Gaab, N.C. Schommer, D.H. Hellhammer, Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis, Psychosom. Med. 61 (1999) 154–162.
[57] M.B. First, R.L. Spitzer, M. Gibbon, J.B.W. Williams. Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV), American Psychiatric Press, Inc, Washington, D.C, 1996.
[58] Bundesamt für Migration und Flüchtlinge (BAMF), *2021 Migration Report: Key Findings*. Bundesamt für Migration und Flüchtlinge, Nürnberg, Germany, 2022.
[59] L. Matos, M.J. Indart, C.L. Park, I. Leal, "That is not my country anymore": pre-and postdisplacement trauma, stressors, and distress in war-affected Syrian civilians, Psychological trauma: theory, research, practice, and policy 14 (1) (2022) 80.
[60] Bundesamt für Migration und Flüchtlinge (BAMF), 2018 migration report: key results. Bundesamt für Migration und Flüchtlinge, Nürnberg, Germany, 2019.
[61] R. Brislin, The wording and translation of research instruments, in: W. Lonner, J. Berry (Eds.), Field Methods in Cross-Cultural Research, Sage Publications, Inc., 1986, pp. 137–164.
[62] B.K. Cheon, V.A. Mathur, J.Y. Chiao, Empathy as cultural process: insights from the cultural neuroscience of empathy, World cultural psychiatry research review 5 (1) (2010) 32–42.
[63] M. Gallistl, M. Kungl, S. Gabler, P. Kanske, P. Vrticka, V. Engert, Attachment and inter-individual differences in empathy, compassion, and theory of mind abilities, Am. J. Bioeth. (2024) 1–16, https://doi.org/10.1080/14616734.2024.2376762.
[64] P. Frazier, C. Greer, S. Gabrielsen, H. Tennen, C. Park, P. Tomich, The relation between trauma exposure and prosocial behavior, Psychological Trauma: Theory, Research, Practice, and Policy 5 (3) (2013) 286.
[65] Bundesamt für Migration und Flüchtlinge (BAMF), *Aktuelle Zahlen (01/2023)*. Bundesamt für Migration und Flüchtlinge, 2023. https://www.bamf.de/SharedDocs/Anlagen/DE/Statistik/AsylinZahlen/aktuelle-zahlen-januar-2023.html?nn=284722.
[66] A.M. Nasser, Y. Geng, S.A. Al-Wesabi, The prevalence of smoking (cigarette and waterpipe) among university students in some Arab countries: a systematic review, Asian Pac. J. Cancer Prev. APJCP: Asian Pac. J. Cancer Prev. APJCP 21 (3) (2020) 583.
[67] M.E. Roberts, B.F. Fuemmeler, F.J. McClernon, J.C. Beckham, Association between trauma exposure and smoking in a population-based sample of young adults, J. Adolesc. Health 42 (3) (2008) 266–274.
[68] R.F. Mollica, Y. Caspi-Yavin, P. Bollini, T. Truong, S. Tor, J. Lavelle, The Harvard Trauma Questionnaire. Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees, J. Nerv.

Ment. Dis. 180 (2) (1992) 111–116. https://www.ncbi.nlm.nih.gov/pubmed/1737972.