EUROPEAN JOURNAL OF PSYCHOTRAUMATOLOGY
2025, VOL. 16, NO. 1, 2504841
https://doi.org/10.1080/20008066.2025.2504841

CLINICAL RESEARCH ARTICLE

OPEN ACCESS Check for updates

Feasibility of an interpreter-mediated neuropsychological test battery for trauma-affected refugees

Søren Kit Bothe [a,b], T. Rune Nielsen [b,c], Linda Nordin [a,d], Sabina Palic [a,e] and Marie Høgh Thøgersen [b,f]

[a]DIGNITY: Danish Institute Against Torture, Copenhagen, Denmark; [b]Department of Psychology, University of Copenhagen, Copenhagen, Denmark; [c]Danish Dementia Research Centre, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark; [d]Department of Psychology, Lund University, Lund, Sweden; [e]Mental Health Services, Psychiatric Centre Glostrup, Glostrup, Denmark; [f]Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

ABSTRACT

Background: Cognitive impairment in trauma-affected refugees is often unassessed due to the absence of suitable cross-cultural neuropsychological measures, the high prevalence of psychiatric symptoms, and the need for interpreter mediation. This poses both a clinical challenge and a health inequality for trauma-affected refugees who are left without access to much needed rehabilitation. Untreated cognitive impairment is likely to reduce the effectiveness of mental health interventions.

Aim: To assess the operational and clinical feasibility of a neuropsychological test battery specifically developed for trauma-affected refugees from Syria.

Method: A neuropsychological test battery was developed to assess executive function, mental speed, attention, and memory. The test battery was administered to 27 refugees from Syria recruited after being referred for specialized trauma treatment. Operational feasibility was assessed by examining completion time and completion rate. Clinical feasibility was assessed through performance validity rates, skewness, and floor effects.

Results: The test battery included Supermarket Fluency, Color Trails Test, Symbol Digit Modalities Test, World Health Organization/University of California, Los Angeles – Auditory Verbal Learning Test, and Digit Span. The mean completion time was 54 min, with 62% of participants completing the full battery and 70% of all tests being completed. Concerning performance validity, 11% of the participants failed two performance validity tests, while 26% failed one. Tests scores were generally skewed, and one subtest, Color Trails 2, showed evidence of a floor effect.

Conclusions: Despite high levels of cognitive impairment in trauma-affected refugee populations, to our knowledge, the feasibility of a neuropsychological test battery has not previously been established for this group. This study supports the operational and clinical feasibility of the proposed interpreter-mediated neuropsychological test battery to trauma-affected refugees, provided that additional time is allocated to complete tests with time limits. The study highlights the need for cross-cultural validation of neuropsychological tests in trauma-affected refugees.

Viabilidad de una batería de pruebas neuropsicológicas mediada por intérprete para refugiados afectados por trauma

Antecedentes: El deterioro cognitivo en los refugiados afectados por trauma suele no ser evaluado debido a la ausencia de pruebas neuropsicológicas transculturales adecuadas, la alta prevalencia de síntomas psiquiátricos, y la necesidad de mediación por intérprete. Esto plantea tanto un desafío clínico como una desigualdad en salud para los refugiados afectados por trauma, quienes quedan sin acceso a una rehabilitación muy necesaria. El deterioro cognitivo no tratado probablemente reduce la eficacia de las intervenciones en salud mental.

Objetivo: Evaluar la viabilidad operativa y clínica de una batería de pruebas neuropsicológicas desarrolladas específicamente para refugiados sirios afectados por trauma.

Método: Se desarrolló una batería de pruebas neuropsicológicas para evaluar la función ejecutiva, velocidad mental, atención y memoria. Esta batería fue administrada a 27 refugiados sirios reclutados luego de ser derivados a tratamiento especializado para trauma. La viabilidad operativa se evaluó examinando el tiempo de aplicación y la tasa de finalización. La viabilidad clínica se evaluó mediante los índices de validez del rendimiento, asimetría y el efecto suelo.

Resultados: La batería de pruebas incluyó Fluidez de Supermercado (*Supermarket Fluency)*, Prueba de Colores y Sendas (*Color Trails Test)*, Prueba de Simbolos y Digitos (*Symbol Digit Modalities Test)*, Prueba de Aprendizaje Verbal Auditivo de la Organización Mundial de la Salud/Universidad de California, Los Angeles (*World Health Organization/University of*

ARTICLE HISTORY

Received 31 January 2025
Revised 24 April 2025
Accepted 6 May 2025

KEYWORDS

PTSD; refugees; torture; cross-cultural; neuropsychology; feasibility

PALABRAS CLAVES

TEPT; refugiados; tortura; transcultural; neuropsicología; viabilidad

HIGHLIGHTS

- Cognitive impairment in trauma-affected refugees is often overlooked due to a lack of suitable neuropsychological tests.
- An interpreter-mediated neuropsychological test battery was found to be both operationally and clinically feasible for trauma-affected refugees from Syria.
- The findings highlight the critical need to validate neuropsychological tests for trauma-affected refugee populations.

CONTACT Søren Kit Bothe sobo@dignity.dk DIGNITY: Danish Institute Against Torture, Bryggervangen 55, 2100 Copenhagen, Denmark; Department of Psychology, University of Copenhagen, Øster Farimagsgade 5 1353 Copenhagen, Denmark

California, Los Angeles Auditory Verbal Learning Test), y Apertura de Digitos (*Digit Span)*. La media del tiempo de aplicación fue de 54 minutos, con un 62 % de los participantes completando toda la batería y un 70 % de todas las pruebas individuales completadas. En cuanto a la validez del rendimiento, el 11 % de los participantes falló en dos pruebas de validez del rendimiento, mientras que el 26 % falló en una. Los puntajes de las pruebas fueron generalmente asimétricos y un subtest, Color Trails 2, mostro evidencia de un efecto suelo.

Conclusión: A pesar de los altos niveles de deterioro cognitivo en poblaciones de refugiados afectados por trauma, hasta donde sabemos, la viabilidad de una batería de pruebas neuropsicológicas no había sido previamente establecida para este grupo. Este estudio respalda la viabilidad operativa y clínica de la batería de pruebas neuropsicológicas mediada por intérprete propuesta para refugiados afectados por trauma, siempre que se asigne tiempo adicional para completar las pruebas con límite de tiempo. El estudio subraya la necesidad de validación transcultural de las pruebas neuropsicológicas en refugiados afectados por trauma.

1. Introduction

Never before have so many people been displaced as today, with one in every 69 people globally being forced to flee their homes (UNHCR, 2023). Refugees face a heightened risk of developing mental disorders and experiencing cognitive impairment (Mirabolfathi et al., 2022; Steel et al., 2009). However, cognitive impairments can be difficult to assess due to the lack of suitable tools for this culturally and linguistically diverse group. As the global refugee population continues to grow, there is an urgent need for cross-culturally sensitive approaches to effectively assess and address cognitive impairment in this vulnerable group.

Refugees who have experienced war, conflict, and flight are at an increased risk of developing post-traumatic stress disorder (PTSD) (Steel et al., 2009) and have a high prevalence of cognitive impairment measured in the cognitive domains attention (Mirabolfathi et al., 2022), memory, and executive functions (Ainamani et al., 2017; Díaz & Wallin Lundell, 2023; Nordin, 2020; Nordin et al., 2024; Stergiopoulos et al., 2015). This substantially reduces wellbeing, level of functioning, and aggravates comorbid psychiatric conditions, such as depression or anxiety (Esterman et al., 2019; Riley et al., 2019). Cognitive impairment is an important clinical challenge as it may impact patients' ability to profit from first line mental health treatment. Neuropsychological testing enables the identification of both the type and severity of cognitive impairment, and adjustment of treatment strategies accordingly.

However, a lack of suitable neuropsychological tests for trauma-affected refugee populations makes cognitive assessment highly challenging. This is due to most tests being developed and normed in a western context and therefore representative of a western cultural, educational, and linguistic context. Further, a range of factors are known to substantially influence neuropsychological test performance, including linguistic factors (e.g. age of second language acquisition), education and literacy, and culture (Ardila et al., 2010; Franzen et al., 2020, 2022). The cultural factor includes issues like being unfamiliar with testing situations (test naivety) or a cultural weight on thoroughness rather than speed (Ardila, 2005). Studies further shows that test results differ between ethnic groups (Nielsen et al., 2018) and that refugees have been unfairly accused of underperforming (malingering) on neuropsychological tests, as evidenced by failing performance validity tests when using (western) criterion-based cut-offs (Gates et al., 2023). Together, these factors make most conventional tests unsuitable for assessment of heterogenic trauma-affected refugee populations. A biased assessment of cognitive impairments can have substantial negative consequences for refugees, who are often forced to pursue new careers, establish new homes, and seek out somatic and mental health treatment. Further, a valid assessment can be quite the ordeal for the psychologist tasked with assessing refugees, with current guidelines leaving it to the individual psychologist to describe and judge strengths and limitations when doing so (Nielsen et al., 2024).

Cognitive impairment has long been suspected of impacting treatment of mental illness in refugees. Treatment outcomes in trauma-affected refugees in routine care in Denmark and other Nordic countries are generally quite limited, with studies indicating that a large subgroup of patients do not experience substantial treatment gains (Carlsson et al., 2018; Nordbrandt et al., 2020; Sandahl et al., 2017; Thøgersen et al., 2023). In addition, trauma-affected refugees have reported high levels of head injury (Nordin, 2020) and severe cognitive difficulties. This led to the hypothesis that a subgroup of trauma-affected refugees might have unaddressed cognitive impairments, leading to poor treatment outcomes. Thus, this undocumented impairment could lead to a significant healthcare inequality for trauma-affected refugees. To address this challenge first the Symbol Digit Modalities Test was included in routine monitoring

in a specialized clinic for trauma-affected refugees, between 2013 and 2014 (Nordin et al., 2024) laying the groundwork for the development of a broader test battery presented in this study. After over a decade of conflict, Syria represents the world's largest refugee crisis and largest refugee group in Denmark (Bendixen, 2023). Refugees from Syria have experienced a high degree of trauma exposure, including torture, violence, or other organized crime (Peconga & Høgh Thøgersen, 2020).

The aim of this study was to investigate the feasibility of a newly developed neuropsychological test battery in recently arrived refugees from Syria, assessed in a specialized trauma clinic setting. Feasibility was evaluated based on the ability to deliver and complete the tests within a set timeframe and the extent to which the results indicated floor effects or invalid performance.

2. Method

2.1. Participants

All participants were newly arrived refugees from Syria in Copenhagen, Denmark, referred for treatment in DIGNITYs national rehabilitation clinic (Thøgersen et al., 2023). Inclusion criteria for receiving treatment are: > 18 years, refugee status, exposure to torture or other organized violence, ability to self-finance transportation to the rehabilitation clinic, presence of primary psychiatric symptoms and somatic symptoms requiring treatment, no current alcohol or drug-abuse, and not currently suffering from psychosis or acutely suicidal. Between 2016 and 2018, a total of 71 patients were eligible for participation in the project. Of these, 33 patients accepted participation (46%). A key reason for non-participation was prioritizing treatment. Of the 33 patients who initially accepted participation, six were excluded (three did not originate from Syria, one failed to provide written consent, one had previously been exposed to the test battery, and one later withdrew participation) resulting in a final sample of 27 participants.

2.2. The interpreter-mediated neuropsychological test battery for trauma-affected refugees

The Danish National Research Ethics Committee exempted the study in accordance with national law. Data collection was performed according to the Helsinki Declaration and informed consent was obtained at the neuropsychological testing session.

The neuropsychological test battery was developed over a period of 15 months. This process included a review of the existing literature, discussions in an expert group comprising psychotraumatologists and neuropsychologists, standardization of the administration procedure in a manual, training of clinicians and interpreters, and clinical implementation.

Tests were selected to assess core cognitive domains commonly affected in PTSD and Traumatic Brain Injury (TBI): Mental Speed, Attention, Memory, and Executive Functions.

Additional selection criteria included cultural and educational fairness – meaning the tests could be administered across diverse ethnic backgrounds and varying levels of formal education without requiring content adaptation. Finally, all tests needed to be feasible to conduct with an interpreter.

The test battery consists of five tests presented in Table 1. The administration manual included procedures for test administration, scoring, norming,

Table 1. Description of neuropsychological test battery.

Measure	Cognitive domain	Description
Wechsler Adult Intelligence Scale 3rd edition – Digit Span (WAIS-III – Digit Span)	*Memory and attention*	Digit Span is a subtest from WAIS-III. An increasing number of digits is read out loud, which the test taker should repeat. Working memory is assessed by the number of correctly repeated numbers, forwards then backwards.
WHO/UCLA – Auditory Verbal Learning Test (WHO/UCLA AVLT)	*Memory and attention*	WHO/UCLA AVLT is a culture sensitive test version of Rey Auditory Verbal Learning Test adapted by WHO. Immediate recall (ALVT 1-5) is the number correctly repeated after a list of 15 words is repeated over five trials. Immediate recall after distractor (AVLT 6) is the number correctly recalled from the first list after presentation and repetition of words from a distractor wordlist. Recognition (AVLT7) is in this test battery administered after Supermarket Fluency, and is the number of correctly recognized words among 15 distractor words after approx. 20 min.
Color Trails Test Part A and Part B (CTT)	*Executive Function.*	CTT is a time-based language – and culture sensitive test comparable with Trail Making Test. It consists of yellow or red numbered (1-25) circles. Executive function is assessed by the time it takes to connect the circles in correct order on part 1 (CTT1) and the time it takes to shift between the yellow and red coloured circles while still connecting in the correct numbered order on part 2 (CTT2).
Symbol Digit Modalities test (SDMT)	*Executive functions and Mental Speed*	SDMT is a time-based test comparable to WAIS-IV-Coding subtest and described as culture free. Participants are presented with a symbol key corresponding to numbers from 1-9. Executive function is assessed by the number of correctly completed items within 90 s after 10 practising items.
Supermarket Fluency	*Executive Functions*	Supermarket Fluency is a time-based verbal fluency test. Fluency is assessed by is the number of named words in the semantic category 'supermarket' within 60 s.

and interpretation of results and included an introduction to the test battery for participants, as well as a structured questionnaire to collect demographic and medical information.

A revised test battery and accompanying manual are currently being validated and are available free of charge to qualified healthcare professionals upon request or via DIGNITY's website. Please note that some tests are subject to copyright and must be purchased from their respective publishers.

2.3. Procedures

All participants completed the neuropsychological test battery at the DIGNITY clinic with the psychologist assigned for treatment. Participants were also assessed for depressive and anxiety symptoms with the 25-item Hopkins Symptom Checklist (HSCL-25) (Wind et al., 2017), for PTSD-symptoms with the Harvard Trauma Questionnaire, part 4 (HTQ-4) (Mollica et al., 1992), and level of functioning with the WHO Disability Assessment Scale (WHODAS-20) (Üstün et al., 2010) as part of DIGNITY's routine clinical assessment (Thøgersen et al., 2023). At the neuropsychological test session, a structured questionnaire was used to collect demographic and medical information, including years of formal schooling, favourite subject in school, history of head trauma (y/n, loss of consciousness), subjective level of pain (on a five-point scale), and other factors potentially affecting test performance (open question). Also, the psychologist noted the time to complete the test battery.

2.4. Implementation

The four administering psychologists had 5–12 years of clinical experience and received specialized training in neuropsychological assessment. Four experienced Arabic-Danish interpreters were trained to interpret the tests and adopt specific roles, depending on the nature of the task (verbal or non-verbal). For non-verbal tests (SDMT and CTT), the interpreter provided verbatim interpretation of the instructions given by the psychologist. For verbal tests, the interpreter first provided verbatim interpretation and then took on different roles. In Digit Span, the digit strings were read out, and responses noted, by the interpreter. In WHO/UCLA AVLT, a translated Arabic version of the word list was read out, and responses noted, by the interpreter. In Supermarket Fluency, the interpreter noted the number of words produced by the patient on the score sheet, while the psychologist kept track of time.

2.5. Feasibility measures

Feasibility was assessed both operationally and clinically with criteria presented in Table 2.

Operational feasibility was assessed by examining completion time and completion rate for the test battery. The test battery was considered feasible if: Data collection for the entire test battery, including the structured questionnaire, could be administered in 90 min or less. More than 50% of participants completed the entire test battery (Kalmar et al., 2008).

Table 2. Operational and clinical criteria.

Feasibility domain	Feasibility measure	Criterion
Operational	Completion time	≤ 90 min
	Full battery completion	> 50% of participants
	Individual test completion	> 70% of participants
Clinical	Proportion under floor	< 20% per test
	Failed ≥2 performance validity tests (PVTs)	< 20% of participants

Each individual test was completed by 70% of the participants (Badawi et al., 2021). Clinical feasibility was assessed by examining the distribution of results in terms of skewness, and failure on performance validity tests (PVT) and proportion under floor. Failure on PVT measures was based on established cutoffs from previous research: WAIS-III Digit Span: ≤ 4 score for digits forward, ≤ 2 score for digits backward and/or ≤ 5 score in total (Daugherty et al., 2021; Lippa, 2018). AVLT recognition trial (AVLT7): ≤ 9 (Lippa, 2018; Reyes et al., 2019; Spreen & Strauss, 1998). Proportion under floor was determined using individual floor thresholds for each measure. Very low cut-offs were selected to investigate if the battery can capture the heterogeneity of the sample: Digit span forward/backward ≤ 2 (second-to-last possible score), ALVT1-7 ≤ 2 (second-to-last possible score), CTT1/2 ≥ 239 (10 s pr number), SDMT ≤ 5 (first row of the score sheet), and Fluency ≤ 5. The test battery was considered clinically feasible if.

The proportion of scores under floor was < 20%. The percentage of participants failing two PVT measures was < 20% (Gates et al., 2023; Lippa, 2018).

2.6. Data analysis

The collected data were analyzed with the statistical programme R, version 4.3.3, using packages dplyr (data manipulation and summary functions) and ggplot2 (data visualization). Operational feasibility was assessed through descriptive statistics, and analysis of missing data and test completion time. Clinical feasibility was examined by analyzing neuropsychological test raw scores, including skewness, number of participants scoring under PVT cut-offs or floor.

3. Results

3.1. Sociodemographic data

All participants scored over cutoff for PTSD, and the majority were exposed to torture. Participants

reported low functioning on the WHODAS-12, scoring above the 95th percentile for dysfunction (Üstün et al., 2010), including on the cognitive domain (items 3 and 6). The participants were educated but three participants reported very low level of education (<3 years). Specifically, one participant reported being both illiterate and innumerate and knowing only the numbers 1–6 (Figure 1).

3.2. Operational feasibility

A total of 27 refugees from Syria completed parts of (10 (37%) participants), or the entire (17 (63%) participants), neuropsychological test battery with a mean completion time of 54 min (SD = 11) and a maximum completion time of 75 min. AVLT had the lowest completion rate, with eight participants (30%) not completing all AVLT measures. Digit Span and Supermarket Fluency had the highest completion rates, with 26 participants (96%) completing all test measures. Completion rates for the individual tests in the battery are presented in Table 3. After assessment of the first three participants, the order of test administration was adjusted so that WAIS-III Digit Span was administered first rather than the AVLT to start with a shorter test. Further inclusion of confounding factors, e.g. pain or medication, on the structured questionnaire was added for inclusion into patient records.

Sociodemographic and psychiatric characteristics

Characteristic	M (SD)	N	N missing
N		27	
Female		17	
Age	40 (8)		
Education (years)	10.5 (3.1)		8
Years in Denmark	3 (1)		2
Exposed to torture		17	2
Psychiatric mean			
*HSCL**	*2.7 (0.49)*		9
*HTQ-4**	*2.69 (0.5)*		
WHODAS-12**	36 (8.86)		10
HTQ-4 subscale domains			
Re-experiencing	3.08 (0.6)		
Dissociation	2.15 (0.9)		
WHODAS-12 cognitive domain* **			
Problems learning	3.2 (1.3)**		10
Problems with concentration	3.1 (1.2)**		10

*HSCL and HTQ-4 range from 1 (low) to 4 (high) and a cut-off point of 1,75 on HTQ4 suggests a DSM IV-PTSD diagnosis (Mollica et al., 1992). **WHODAS-12 items range from 1 (low) to 5 (high). ***WHODAS-12 items 3 and 6.

Figure 1. Sociodemographic and psychiatric characteristics.

3.3. Clinical feasibility

Table 3 shows the completion rate and raw score data for the neuropsychological tests as well as their published administration times (Axelrod, 2001; D'Elia et al., 1999; Smith, 1982; Spreen & Strauss, 1998, p. 328).

3.4. Skewness and floor effects

Skewness is an indication of the symmetry of the distribution where a high score describes how far the mean is from the middle of the distribution. Some tests displayed slight left or right skewness. Notably, AVLT1 and CTT1 showed strong right skewness (>1) with a long tail towards higher scores indicating few individuals achieving high scores, while the majority achieved lower scores. This skewed distribution is also illustrated in Figure 2.

CTT2 failed the floor effect criteria with 27% of participants exceeding the 240-second time limit. The proportion scoring under floor on the remaining tests was below 20%, ranging from 0% to 13%.

3.5. Performance validity testing

Seven participants (26%) failed one performance validity test; three WAIS-III Digit Span, and four AVLT-recognition. Further, three participants (11%) failed both performance validity tests (WAIS-III Digit Span and AVLT-recognition).

4. Discussion

Overall, the study supports the operational and clinical feasibility of the developed test battery. Test were delivered within the allocated time with completion rates above the set threshold. Results were skewed and non-normally distributed with few participants scoring at the higher end, and the majority scoring at the lower end. However, floor effects were generally low with only one subtest, CTT2, showing evidence of a floor effect. This indicates that the test battery can cover the diverse range of scores in the group of trauma-affected refugees. Several considerations regarding feasibility emerged from the analysis which are important to highlight.

Table 3. Raw neuropsychological data.

Subtest measure	Missing	Reported administration time	Completion rate (5)	Range of scores	Raw score mean[a] (SD)	Raw score skewness[b]	Proportion under floor[c]
DSF[d]	1	4 min	96	1–10	6.12 (2.20)	−0.67	8%
DSB[d]	1		96	1–9	5.42 (2.06)	−0.23	8%
AVLT1[e]	4	15 min	85	2–14	6.43 (2.46)	1.16	4%
AVLT5[e]	7		74	4–15	10 (3.2)	−0.12	0%
AVLT6[e]	5		81	2–13	7.4 (3.6)	0.32	4%
AVLT7[e]	8		70	5–15	11 (3.8)	−0.34	0%
CTT1[f]	2	8 min	93	31–240	106 (61.8)	1.09	13%
CTT2[f]	4		85	83–240	159 (60.3)	0.25	27%
SDMT	4	5 min	85	6–59	29.4 (14.2)	0.24	0%
Fluency	1	2 min	96	3–43	18.3 (11.2)	0.66	4%

Note. [a]A higher raw score indicates better performance except for CTT where the score represents time, so the opposite is true.
[b]A higher skewness indicates an asymmetrical distribution with a long tail on the right side e.g. few higher neuropsychological scores.
[c]Indicates the percentage of scores under the individually defined floor.
[d]DSF is WAIS-III Digit Span Forward and DSB is WAIS-III Digit Span Backward.
[e]Auditory Verbal Learning Test trial 1) immediate recall after first reading. Trial 5) recall after 5th reading, Trial 6) recall after distractor, Trial 7) 20 min recognition.
[f]Color Trails Test Part 1 and Part 2.

First, consistent with earlier studies (Veliu & Leathem, 2017) the findings suggest that refugees should be allowed more time to complete neuropsychological tests than the reported time limits. While the study supports the overall clinical feasibility of the test battery, the CTT2 failed the floor effect criterion, with 27% of participants reaching the floor. The CTT2 is a time-based test and uses a manual-specified 10-second prompt if participants fail to identify the next target, resulting in a maximum score of 240. As suggested in previous cross-cultural studies (Nielsen et al., 2018), it may be valuable to extend the time before a prompt is given to 20 s, with a corresponding maximum score of 480 to ensure the CTT effectively captures the range of scores in trauma-affected refugees. The need to extend the time limit can stem from multiple factors, such as a cultural weighting on thoroughness over speed (Ardila, 2005), low educational level (Rabelo et al., 2010), or impaired mental speed, a known

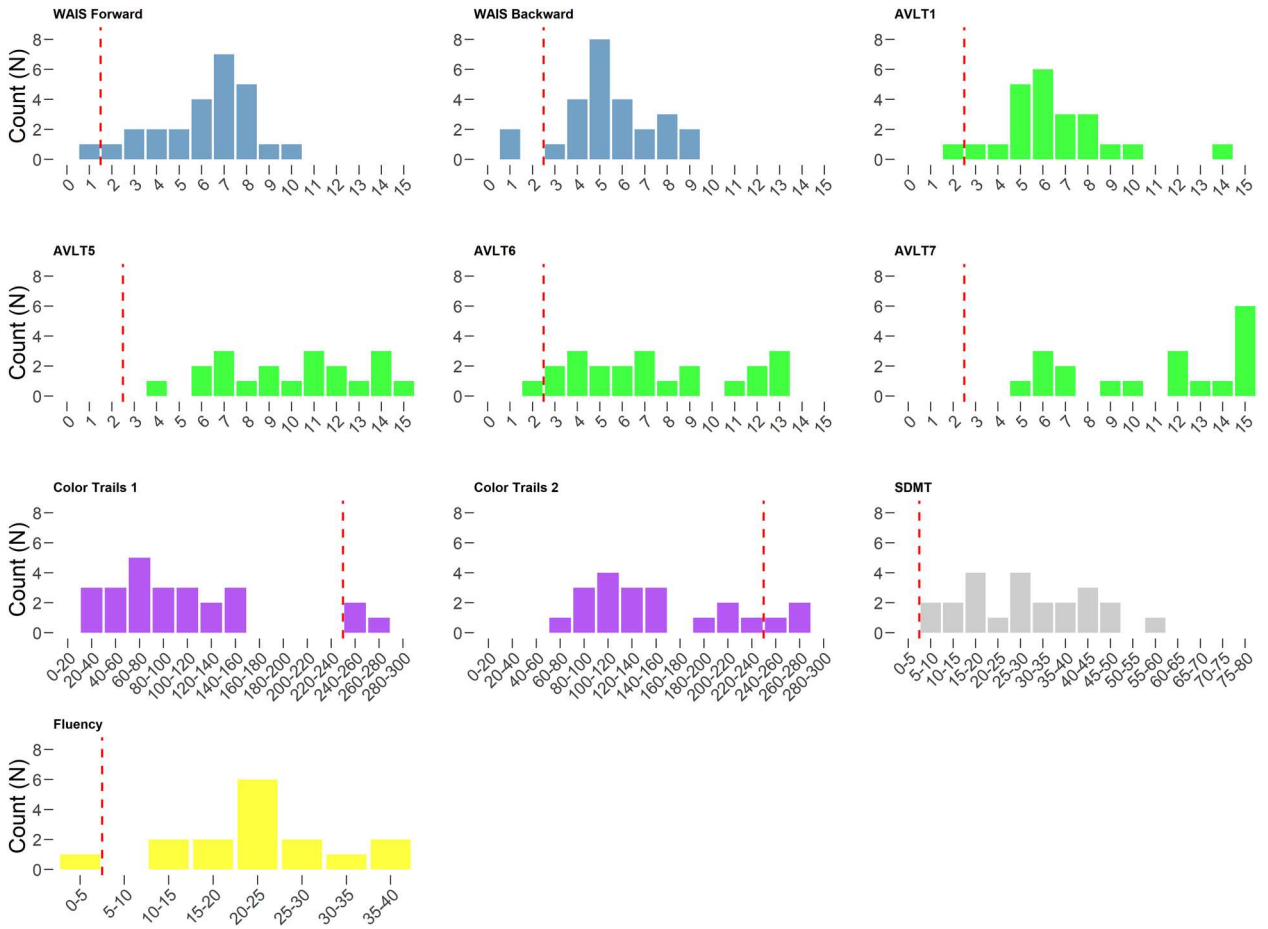

Figure 2. (A–J) Histogram of the distribution of scores.
Note: The red dashed line indicates the defined floor cut-off score.

characteristic in PTSD (Vasterling & Lippa, 2014). Second, the study highlights the importance of accommodating varying degrees of educational levels when planning and interpreting neuropsychological testing for trauma-affected refugees. We found that three participants had very little education raising concerns about their literacy and numeracy. This has implications for feasibility and for the validity relating to the cognitive construct being assessed, as we are interested in investigating the relevant cognitive function, and not the participants' numeracy (e.g. in the CTT) or ability to write numbers (e.g. in the SDMT). For example, one participant was innumerate and therefore could not complete three out of five tests (WAIS-III Digit Span, CTT, and SDMT). Thus, although widely applicable, the current test battery is unfeasible for people who are innumerate. Third, inclusion of shorter assessments may enhance the overall feasibility of the test battery. A pattern emerged, suggesting that lower test completion rates and scores were associated with both the length and order of administration of the neuropsychological tests. For instance, WAIS-III Digit Span, which had the highest completion rate, was administered at the beginning of the test battery and was relatively quick to administer. In contrast, the AVLT7, with the lowest completion rate, was administered at the end of the assessment and was part of the test with the longest administration time. This suggests that tests with longer administration times are difficult for the current sample, and that fatigue may impact performance. A modified test battery could consider investigating the role of fatigue by inclusion of shorter tests and longer tests measuring the same cognitive construct. Fourth, the study supports the feasibility of interpreter-mediated testing with the current test battery. This includes the interpreter taking on a different role, which sometimes requires deviations from the standardized instructions – a phenomena also previously observed in cross-cultural interpreter-mediated testing (Franzen et al., 2023; Nielsen et al., 2024). We find that the inclusion of a practice assignments as in WAIS-III Digit Span, CTT and SDMT enhances the participants understanding of test requirements. Consistent with previous research, we find that interpreter mediated testing is feasible when following these accommodations (Nielsen et al., 2024) However, future research should explore the impact of interpreter-mediation on performance (Gates et al., 2023). This could not be assessed in the current study, as all participants were assessed with support from an Arabic interpreter. Fifth, the use of Performance Validity Tests is a growing and promising field for identifying invalid performance, but base rates for trauma-affected refugees have so far been unexplored. Previous research indicates that lower thresholds are necessary for specific groups including culturally and linguistically diverse populations (Lippa, 2018), non-psychotic psychiatric patients (McWhirter et al., 2020), and when tests are interpreter mediated (Gates et al., 2023), all of which are characteristics of the present sample. Additionally, PTSD-related symptoms such as dissociation and flashbacks can substantially impact performance if present during testing. To enhance the reliability of results, trauma-informed practises should be applied, including creating a safe environment and assessing the occurrence of flashbacks. If flashbacks are a regular part of the patient's daily life, continuing testing even after a flashback occurred may provide a more representative assessment of their cognitive state. Lastly, while the psychologists administering the tests initially expressed concerns regarding the feasibility of testing trauma-affected refugees, these concerns shifted over the study period. The primary concern was a potential negative impact of testing on the therapeutic alliance and the participants' wellbeing. The psychologists were concerned that asking vulnerable patients to complete challenging tests might make them feel misunderstood or alienated, and lead to emotional exhaustion or reduced self-esteem. However, as they gained practical experience with testing, these concerns proved to be unfounded. Based on their observations, patients generally had a positive experience with the testing process. However, these findings have not been systematically documented and warrant further exploration.

This study represents an effort to adapt neuropsychological tests for a trauma-affected, non-Western refugee population. The high prevalence of cognitive impairment among refugees, combined with the lack of culturally appropriate neuropsychological tests, highlights a significant health inequality for trauma-affected individuals. While this study primarily focused on the feasibility of the test battery, the observed neuropsychological scores were so low that their clinical interpretation may be constrained. While PTSD-related cognitive impairments may contribute to these low scores, other factors – such as very low educational levels, linguistic barriers, and cultural differences – could also play a significant role. Furthermore, with generally low scores across all domains, especially when norm-adjusting using Western norms, the test battery's ability to differentiate between low and very low scores is restricted, which limits clinical interpretation. Nevertheless, given the limited research into cognitive impairment in trauma-affected refugees, employing assessments that accommodate a broad range of scores is an important first step. Further research could utilize the current tests to capture a broader range of scores and generate new normative data that is relevant for non-Western refugee populations and explore the clinical and cross-cultural validity.

5. Strengths and limitations

While this study assessed a limited set of feasibility measures in a selected group of refugees from Syria it provides a methodological strength by investigating feasibility in a culturally homogeneous group, thus minimizing the cultural and linguistic influences. However, several limitations should be noted. Recruiting participants and obtaining psychiatric, sociodemographic, and neuropsychological data from this specific population proved challenging. This is characteristic of the refugee population who often lack time, stability or energy to participate in treatment and research activities (Thøgersen et al., 2023). This may have created an unintended selection bias and limits the generalizability of the findings. Although the psychologists were trained to administer the test battery, their supervision was limited to two days of training followed by 10 h of ongoing supervision. This supervision was further constrained by scheduling conflicts and the irregular scheduling of neuropsychological tests, which may have impacted their adherence to the test manual. Furthermore, the study lacks a systematic description of both the perspectives of the clinician, interpreter, or test subject regarding the test battery (face validity) and clinical observations (e.g. flashbacks, missing glasses) for not completing the full test battery. As a result of the lack of representative normative data, we have chosen not to report norm-adjusted scores according to the available norm material (D'Elia et al., 1999; Jørgensen, 2012; Schmidt, 1996; Sheridan et al., 2006; Weschler, 2011), as they could lead to a skewed and potentially unfair interpretation of cognitive performance in the current sample. However, the lack of representative normative data and the non-normal distribution of neuropsychological scores hinders the comparison to relevant normative data and investigation of z-transformed scores. Due to the scope of the study and the small sample size potential correlations between neuropsychological scores and psychiatric scores could not be examined thoroughly, leaving it to future studies. Future research should include qualitative feedback from trauma-affected refugees and interpreters for a more in-depth perspective on feasibility.

6. Conclusion

Despite high levels of cognitive impairment in trauma-affected refugee populations, no neuropsychological test battery has, to our knowledge, been shown to be feasible, leaving cognitive impairment largely unassessed in this population. This study supports the operational and clinical feasibility of an interpreter-mediated neuropsychological test battery for trauma-affected refugees from Syria. However, scores appear on average very low and were skewed and non-normally distributed. Future research should investigate the validity of neuropsychological tests in trauma-affected refugees.

Acknowledgments

The authors would like to thank DIGNITY Danish Institute against Torture and the practitioners for the support and provision of resources in recruitment of participants and data collection. We would also like to thank Center for Hjerneskade (The Danish Center for Brain Injury) for guidance and supervision. All authors declare no conflict of interest. This research was carried out with the financial support of, DIGNITY. All patients gave informed consent to participate in the study. OpenAI's ChatGPT-4 model (https://chat.openai.com/) was utilized for coding assistance with the first author thoroughly reviewing and revising the outputs to ensure accuracy. Due to the nature of this research, the data will not be shared publicly, and supporting data is not available.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Danish Victims Fund, administered by the Department of Civil Affairs under the Ministry of Justice, through Grant [2022/2], case number 22-610-00189.

Data availability statement

The participants in this study did not give written consent for their data to be shared publicly. Due to the sensitive nature of the research, supporting data is not available.

ORCID

Søren Kit Bothe http://orcid.org/0009-0007-8681-9066
T. Rune Nielsen http://orcid.org/0000-0002-8128-2294
Linda Nordin http://orcid.org/0000-0002-8533-520X
Sabina Palic http://orcid.org/0000-0002-9686-9370
Marie Høgh Thøgersen http://orcid.org/0000-0002-2081-3354

References

Ainamani, H. E., Elbert, T., Olema, D. K., & Hecker, T. (2017). PTSD symptom severity relates to cognitive and psycho-social dysfunctioning – A study with Congolese refugees in Uganda. *European Journal of Psychotraumatology*, *8*(1), 1–10. https://doi.org/10.1080/20008198.2017.1283086

Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology Review*, *15*(4), 185–195. https://doi.org/10.1007/s11065-005-9180-y

Ardila, A., Bertolucci, P. H., Braga, L. W., Castro-Caldas, A., Judd, T., Kosmidis, M. H., Matute, E., Nitrini, R., Ostrosky-Solis, F., & Rosselli, M. (2010). Illiteracy: The neuropsychology of cognition without reading. *Archives of Clinical Neuropsychology*, *25*(8), 689–712. https://doi.org/10.1093/arclin/acq079

Axelrod, B. N. (2001). Administration duration for the Wechsler Adult Intelligence Scale-III and Wechsler Memory Scale-III. *Archives of Clinical Neuropsychology*, *16*(3), 293–301. https://doi.org/10.1016/S0887-6177(00)00048-2

Badawi, A., Steel, Z., Mahoney, C., & Berle, D. (2021). Feasibility of an adjunctive cognitive task in the treatment of posttraumatic stress disorder. *Discover Psychology*, *1*(1), 11. https://doi.org/10.1007/s44202-021-00011-5

Bendixen, A. M. C. (2023, June 30). *Flygtninge i Danmark historisk*. refugees.dk.https://refugees.dk/da/fakta/tal-og-statistik/hvor-mange-flygtninge-har-danmark-taget-imod-gennem-aarene-og-hvor-kom-de-fra/

Carlsson, J., Sonne, C., Vindbjerg, E., & Mortensen, E. L. (2018). Stress management versus cognitive restructuring in trauma-affected refugees—A pragmatic randomised study. *Psychiatry Research*, *266*, 116–123. https://doi.org/10.1016/j.psychres.2018.05.015

Daugherty, J. C., Querido, L., Quiroz, N., Wang, D., Hidalgo-Ruzzante, N., Fernandes, S., Pérez-García, M., De Los Reyes-Aragon, C. J., Pires, R., & Valera, E. (2021). The coin in hand–extended version: Development and validation of a multicultural performance validity test. *Assessment*, *28*(1), 186–198. https://doi.org/10.1177/1073191119864652

D'Elia, L., Satz, P., Uchiyama, C., & White, T. (1999). *Color Trails Test (CTT) Professional Manual*. PAR. https://www.parinc.com/Products/Pkey/77

Díaz, M., & Wallin Lundell, I. (2023). *Tortyrens påverkan på exekutiva funktioner: Konsekvenser av PTSD och traumatisk hjärnskada hos tortyröverlevare vid traumabehandling, asylprocess och integration*. Röda Korsets Högskola; Svenska Röda Korset. https://urn.kb.se/resolve?urn=urn:nbn:se:rkh:diva-4527

Esterman, M., Fortenbaugh, F. C., Pierce, M. E., Fonda, J. R., DeGutis, J., Milberg, W., & McGlinchey, R. (2019). Trauma-related psychiatric and behavioral conditions are uniquely associated with sustained attention dysfunction. *Neuropsychology*, *33*(5), 711–724. https://doi.org/10.1037/neu0000525

Franzen, S., European Consortium on Cross-Cultural Neuropsychology (ECCroN), Watermeyer, T. J., Pomati, S., Papma, J. M., Nielsen, T. R., Narme, P., Mukadam, N., Lozano-Ruiz, Á, Ibanez-Casas, I., Goudsmit, M., Fasfous, A., Daugherty, J. C., Canevelli, M., Calia, C., van den Berg, E., & Bekkhus-Wetterberg, P. (2022). Cross-cultural neuropsychological assessment in Europe: Position statement of the European Consortium on Cross-Cultural Neuropsychology (ECCroN). *The Clinical Neuropsychologist*, *36*(3), 546–557. https://doi.org/10.1080/13854046.2021.1981456

Franzen, S., Papma, J. M., van den Berg, E., & Nielsen, T. R. (2020). Cross-cultural neuropsychological assessment in the European Union: A Delphi expert study. *Archives of Clinical Neuropsychology*, *36*(5), 815–830. https://doi.org/10.1093/arclin/acaa083

Franzen, S., Van Den Berg, E., Bossenbroek, W., Kranenburg, J., Scheffers, E. A., Van Hout, M., Van De Wiel, L., Goudsmit, M., Van Bruchem-Visser, R. L., Van Hemmen, J., Jiskoot, L. C., & Papma, J. M. (2023). Neuropsychological assessment in the multicultural memory clinic: Development and feasibility of the TULIPA battery. *The Clinical Neuropsychologist*, *37*(1), 60–80. https://doi.org/10.1080/13854046.2022.2043447

Gates, T. M., Metcalf, K., Sabaz, M., Young, K., Anandakumar, T., & Simpson, G. K. (2023). What are the predictors of TOMM failure in clinical TBI populations? A retrospective analysis. *Journal of the International Neuropsychological Society*, *29*(4), 336–345. https://doi.org/10.1017/S1355617722000339

Jørgensen, K. (2012). *Danske normer til neuropsykologiske tests* (1., pp. 15–33, 79–90). Dansk Psykologisk Forlag.

Kalmar, K., Novack, T. A., Nakase-Richardson, R., Sherer, M., Frol, A. B., Gordon, W. A., Hanks, R. A., Giacino, J. T., & Ricker, J. H. (2008). Feasibility of a brief neuropsychologic test battery during acute inpatient rehabilitation after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *89*(5), 942–949. https://doi.org/10.1016/j.apmr.2008.01.008

Lippa, S. M. (2018). Performance validity testing in neuropsychology: A clinical guide, critical review, and update on a rapidly evolving literature. *The Clinical Neuropsychologist*, *32*(3), 391–421. https://doi.org/10.1080/13854046.2017.1406146

McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical populations—A systematic review. *Journal of Neurology, Neurosurgery & Psychiatry*, *91*(9), 945–952. https://doi.org/10.1136/jnnp-2020-323776

Mirabolfathi, V., Schweizer, S., Moradi, A., & Jobson, L. (2022). Affective working memory capacity in refugee adolescents. *Psychological Trauma: Theory, Research, Practice, and Policy*, *14*(6), 983–988. https://doi.org/10.1037/tra0000552

Mollica, R. F., Caspi-Yavin, Y., Bollini, P., Truong, T., Tor, S., & Lavelle, J. (1992). The Harvard trauma questionnaire: Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. *The Journal of Nervous and Mental Disease*, *180*(2), 111–116. https://doi.org/10.1097/00005053-199202000-00008

Nielsen, T. R., Franzen, S., Watermeyer, T., Jiang, J., Calia, C., Kjærgaard, D., Bothe, S., & Mukadam, N. (2024). Interpreter-mediated neuropsychological assessment: Clinical considerations and recommendations from the European Consortium on Cross-Cultural Neuropsychology (ECCroN). *The Clinical Neuropsychologist*, *38*(8), 1775–1805. https://doi.org/10.1080/13854046.2024.2335113

Nielsen, T. R., Segers, K., Vanderaspoilden, V., Bekkhus-Wetterberg, P., Minthon, L., Pissiota, A., Bjørkløf, G. H., Beinhoff, U., Tsolaki, M., Gkioka, M., & Waldemar, G. (2018). Performance of middle-aged and elderly European minority and majority populations on a Cross-Cultural Neuropsychological Test Battery (CNTB). *The Clinical Neuropsychologist*, *32*(8), 1411–1430. https://doi.org/10.1080/13854046.2018.1430256

Nordbrandt, M. S., Sonne, C., Mortensen, E. L., & Carlsson, J. (2020). Trauma-affected refugees treated with basic body awareness therapy or mixed physical activity as augmentation to treatment as usual—A pragmatic randomised controlled trial. *PLoS One*, *15*(3), e0230300. https://doi.org/10.1371/journal.pone.0230300

Nordin, L. (2020). *Posttraumatic stress reactions in tortured refugees: Relationship to pain, cognitive impairments, and negative appraisals*. Lund Universitet.

Nordin, L., Bothe, S. K., Perrin, S., & Rorsman, I. (2024). Severe cognitive impairment in trauma-affected refugees—Exploring the impact of traumatic brain injury. *Journal of Clinical Medicine*, *13*(17), Article 17. https://doi.org/10.3390/jcm13175096

Peconga, E. k., & Høgh Thøgersen, M. (2020). Post-traumatic stress disorder, depression, and anxiety in adult Syrian refugees: What do we know? *Scandinavian*

Journal of Public Health, *48*(7), 677–687. https://doi.org/10.1177/1403494819882137

Rabelo, I., Pacanaro, S., Rossetti, M., Leme, I., Castro, N., Güntert, C., Miotto, E., & Lucia, M. (2010). Color trails test: A Brazilian normative sample. *Psychology & Neuroscience*, *3*(1), 93–99. https://doi.org/10.3922/j.psns.2010.1.012

Reyes, A., LaBode-Richman, V., Salinas, L., & Barr, W. B. (2019). WHO-AVLT recognition trial: Initial validation for a new malingering index for Spanish-speaking patients. *Applied Neuropsychology: Adult*, *26*(6), 564–572. https://doi.org/10.1080/23279095.2018.1470974

Riley, E., Mitko, A., Stumps, A., Robinson, M., Milberg, W., McGlinchey, R., Esterman, M., & DeGutis, J. (2019). Clinically significant cognitive dysfunction in OEF/OIF/OND veterans: Prevalence and clinical associations. *Neuropsychology*, *33*(4), 534–546. https://doi.org/10.1037/neu0000529

Sandahl, H., Jennum, P., Baandrup, L., Poschmann, I. S., & Carlsson, J. (2017). Treatment of sleep disturbances in trauma-affected refugees: Study protocol for a randomised controlled trial. *Trials*, *18*(1), 520. https://doi.org/10.1186/s13063-017-2260-5

Schmidt, M. (1996). *Rey auditory verbal learning test: A handbook*. Western Psychological Services.

Sheridan, L. K., Fitzgerald, H. E., Adams, K. M., Nigg, J. T., Martel, M. M., Puttler, L. I., Wong, M. M., & Zucker, R. A. (2006). Normative symbol digit modalities test performance in a community-based sample. *Archives of Clinical Neuropsychology*, *21*(1), 23–28. https://doi.org/10.1016/j.acn.2005.07.003

Smith, A. (1982). *Symbol digit modalities test*. Western Psychological Services.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed., pp. 326–340). Oxford University Press.

Steel, Z., Chey, T., Silove, D., Marnane, C., Bryant, R. A., & van Ommeren, M. (2009). Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: A systematic review and meta-analysis. *JAMA*, *302*(5), 537. https://doi.org/10.1001/jama.2009.1132

Stergiopoulos, V., Cusi, A., Bekele, T., Skosireva, A., Latimer, E., Schütz, C., Fernando, I., & Rourke, S. B. (2015). Neurocognitive impairment in a large sample of homeless adults with mental illness. *Acta Psychiatrica Scandinavica*, *131*(4), 256–268. https://doi.org/10.1111/acps.12391

Thøgersen, M. H., Bager, L., Bangsgaard, S. G., Palic, S., Auning-Hansen, M., Møller, S. B., Larsen, K. B., Tækker, L., Jensen, B. S., Bothe, S., & Nordin, L. (2023). The Danish Trauma Database for Refugees (DTD): A multicenter database collaboration—overcoming the challenges and enhancing mental health treatment and research for refugees. *International Journal of Environmental Research and Public Health*, *20*(16), 6611. https://doi.org/10.3390/ijerph20166611

UNHCR, T. U. R. A. (2023). *Mid-year trends* 2023. UNHCR. https://www.unhcr.org/mid-year-trends-report-2023

Üstün, T. B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korff, M., & Pull, C. (2010). Developing the World Health Organization disability assessment schedule 2.0. *Bulletin of the World Health Organization*, *88*(11), 815–823. https://doi.org/10.2471/BLT.09.067231

Vasterling, & Lippa, S. (2014). Neurocognitive Alterations Associated with PTSD. In M. J. Friedman, T. M. Keane, & P. A. Resick, *Handbook of PTSD, Second Edition: Science and Practice* (pp. 185–199). Guilford Publications.

Veliu, B., & Leathem, J. (2017). Neuropsychological assessment of refugees: Methodological and cross-cultural barriers. *Applied Neuropsychology: Adult*, *24*(6), 481–492. https://doi.org/10.1080/23279095.2016.1201483

Weschler, D. (2011). *WAIS-III DK. Wechsler Adult Intelligence Scale-III. Vejledning* (pp. 1–40). Pearson Sweden.

Wind, T. R., van der Aa, N., de la Rie, S., & Knipscheer, J. (2017). The assessment of psychopathology among traumatized refugees: Measurement invariance of the Harvard Trauma Questionnaire and the Hopkins Symptom Checklist-25 across five linguistic groups. *European Journal of Psychotraumatology*, *8*(sup2), 12. https://doi.org/10.1080/20008198.2017.1321357