

# Psychometric properties of the Dutch revised sense of coherence scale in a firefighter sample

Iris J. L. Egberink <sup>a</sup>, Tim Harms <sup>b</sup> and Miriam J. J. Lommen <sup>b</sup>

<sup>a</sup>Department of Psychometrics and Statistics, University of Groningen, Groningen, The Netherlands; <sup>b</sup>Department of Clinical Psychology and Experimental Psychopathology, University of Groningen, Groningen, The Netherlands

## ABSTRACT

**Background:** Sense of coherence (SOC) has been associated with resilience to posttraumatic stress disorder (PTSD) and seems to be a promising factor in primary prevention of PTSD in high risk populations.

**Objective:** The present study evaluated the psychometric properties of the Dutch revised Sense of Coherence Scale (SOC-R) in a sample of  $N = 527$  firefighters.

**Method:** To investigate the internal structure of this 13-item scale, a Mokken scale analysis and an exploratory factor analysis were conducted (i.e. parallel analysis based on MRFA).

**Results:** The combined results of these analyses suggested that a one-factor solution with 10 out of 13 items was most compelling for our firefighter sample. Reliability estimates for the 10-item version increased compared to the 13-item version (13-item:  $\alpha = .82$ ,  $\lambda_2 = .83$ ; 10-item:  $\alpha = \lambda_2 = .85$ ). As expected, the Dutch version showed positive associations with resilience (convergent validity), and low correlations with neuroticism and extraversion (discriminant validity).

**Conclusions:** The one-factor solution of the Dutch SOC-R with 10 items (excluding item 2, item 3, and item 6) is most convincing. The use of this scale might be specifically interesting regarding its potential to primary prevention of trauma-related psychopathology in high-risk samples.

## Propiedades psicométricas de la escala revisada holandesa de sentido de coherencia en una muestra de bomberos

**Antecedentes:** El sentido de coherencia (SOC) se ha asociado con la resiliencia ante el trastorno de estrés posttraumático (TEPT) y parece ser un factor prometedor en la prevención primaria del TEPT en poblaciones de alto riesgo.

**Objetivo:** El presente estudio evaluó las propiedades psicométricas de la Escala revisada holandesa de Sentido de Coherencia (SOC-R) en una muestra de  $N = 527$  bomberos.

**Método:** Para investigar la estructura interna de esta escala de 13 ítems, se realizó un análisis de escala Mokken y un análisis factorial exploratorio (es decir, análisis paralelo basado en MRFA).

**Resultados:** Los resultados combinados de estos análisis sugirieron que una solución de un factor con 10 de los 13 elementos fue más convincente para nuestra muestra de bomberos. Las estimaciones de confiabilidad para la versión de 10 ítems aumentaron en comparación con la versión de 13 ítems (13 ítems:  $\alpha = .82$ ,  $\lambda^2 = .83$ ; 10 ítems:  $\alpha = \lambda^2 = .85$ ). Como se esperaba, la versión holandesa mostró asociaciones positivas con la resiliencia (validez convergente) y bajas correlaciones con neuroticismo y extraversión (validez discriminante).

**Conclusiones:** La solución de un factor de la SOC-R holandesa con 10 ítems (excluyendo el ítem 2, el ítem 3 y el ítem 6) es muy convincente. El uso de esta escala podría ser específicamente interesante en cuanto a su potencial para la prevención primaria de psicopatología relacionada con el trauma en muestras de alto riesgo.

## 消防员样本中荷兰修订版心理一致感量表的心理测量学性质

背景:心理一致感 (SOC) 与创伤后应激障碍 (PTSD) 的心理韧性相关, 并且似乎是 PTSD 高风险人群的一级预防中有前景的因素。

目标:本研究在一个 527 名消防员组成的样本中, 评估了荷兰修订版心理一致感量表 (SOC-R) 的心理测量学性质。

方法:为考查 13 条目标量表的内部结构, 进行了 Mokken 量表分析和探索性因素分析 (即基于 MRFA 的平行分析)。

结果:这些分析的综合结果表明, 对于我们的消防员样本, 在 13 个条目中有 10 个单因素的解最令人信服。10 条目版本的信度估计值较 13 条目版本相比有所提高 (13 条目:  $\alpha = .82$ ,  $\lambda^2 = 0.83$ ; 10 条目:  $\lambda^2 = .85$ )。同预期一样, 荷兰语版本展现出与心理韧性 (收敛效度) 及与神经质和外向性 (区分效度) 的正相关。

## ARTICLE HISTORY

Received 3 December 2019

Revised 6 April 2020

Accepted 14 April 2020

## KEYWORDS

Sense of coherence; SOC-R; Mokken scaling; exploratory factor analysis; firefighters

## PALABRAS CLAVE

sentido de coherencia; SOC-R; Escala de Mokken; análisis factorial exploratorio; Bomberos

## 关键词

心理一致感; SOC-R; Mokken 量表; 探索性因素分析; 消防员

## HIGHLIGHTS

- Psychometric evaluation increases the valid use of measures.
- A total score based on 10 items provides the best Dutch SOC-R measurement.
- The Dutch SOC-R has the potential to be an interesting measure that can be used to further understand the development of PTSD in high-risk samples.

结论:荷兰SOC-R的10条目单因素解(不包括条目2, 条目3和条目6)最具说服力。该量表在高风险样本中对于创伤相关精神疾病的一级预防的应用潜力, 值得深入探讨。

Originally, the concept of sense of coherence (SOC) was introduced as a set of beliefs that the inner and outer environment is predictable, explicable, and comprehensive. These beliefs would serve as a resource to cope with life events in order to maintain good mental health (Antonovsky, 1979). The first measure of SOC concerned the 29-item *Orientation of Life Questionnaire* (SOC-A), scored on a seven-point Likert scale (Antonovsky, 1987). Reliability ( $\alpha$  ranging from .70 to .95) and test-retest reliability ( $r$  ranging from .52 to .80) of this scale were sufficient and with translations in 33 languages it achieved a widespread usage (Eriksson & Lindström, 2005, 2007). Despite this, the SOC-A was repeatedly criticized for its unstable factor structure and validity (e.g. Antonovsky, 1993). Moreover, high correlations with other psychological constructs such as anxiety, depression, neuroticism, and optimism raised the question whether SOC was an independent concept or just a reversed measure of already established constructs (e.g. Feldt, Metsäpelto, Kinnunen, & Pulkkinen, 2007; Gruszczynska, 2006).

To address the aforementioned criticisms, Bachem and Maercker (2016) developed a revised sense of coherence scale (SOC-R) based on a new SOC definition; ‘the general ability to perceive life phenomena as connected to each other and to balance positive and negative appraisals of life experiences’ (Bachem & Maercker, 2016, p. 1). This definition takes the whole range of positive and negative life experiences, as well as the integration process, into account. It is regarded a meta-heuristic that influences reactions towards stressful life events. In this it is similar to resilience, which has been defined as the dynamic process of positive adaptation to previous trauma or adversity (Luthar & Cicchetti, 2000), amongst many other definitions of the concept of resilience. Despite an ongoing debate on the exact definition of resilience (Fletcher & Sarkar, 2013; Southwick, Bonanno, Masten, Panter-Brick, & Yehuda, 2014), resilience and SOC seem overlapping constructs, which would enhance coping with stressful life events to their own degree (Fossion et al., 2014). It has been argued that in contrast to resilience, SOC can be directly changed through therapeutic interventions, which makes it interesting for preventive mental health care (Koutsoukou-Argyriaki et al., 2018). This view has been challenged by consensus about resilience being a trait that changes over time (Fletcher & Sarkar, 2013) and by studies testing the effects of resilience training programmes. However, a review on these

studies show mixed evidence for the effectiveness of resilience training, with only a few, relatively small studies showing positive effects of resilience training programmes using a combination of cognitive behavioural therapy and mindfulness (Joyce et al., 2018). Although promising, given the mixed results, focusing on SOC rather than resilience seems to be an interesting approach to investigate further options for preventive mental health care.

Diving into the psychometric properties of Bachem and Maercker’s SOC-R, this questionnaire includes 13 items divided into three subscales (manageability, reflection, and balance), which are scored on a five-point Likert scale. The original study of Bachem and Maercker (2016) used a bereaved sample ( $N = 334$ ;  $M_{age} = 43.7$ ; 12.6% males) and a control sample ( $N = 157$ ;  $M_{age} = 40.2$ ; 46.9% males) for the validation process. They constructed the scale with an exploratory factor analysis (EFA) and conducted a confirmatory factor analysis (CFA) to affirm their proposed structure. Since the proposed three-factor structure was only confirmed in the bereaved sample, the authors suggested that SOC is more relevant for individuals with recent stressful life experiences. The reliability in the bereaved sample was good for the total scale ( $\alpha = .81$ ) and questionable to acceptable for the subscales (i.e. manageability:  $\alpha = .77$ , reflection:  $\alpha = .76$ , balance:  $\alpha = .63$ ). Results showed a test-retest reliability of  $r = .85$  after four weeks and  $r = .74$  after 15 months. The overlap with other psychological constructs, including optimism ( $r = .39$ ) and neuroticism ( $r = -.23$ ), was lower compared to the SOC-A ( $r = .66$  and  $r = -.85$ , respectively; Feldt et al., 2007).

In a study among older Swiss participants ( $N = 268$ ;  $M_{age} = 67$ ; 28.6% males), the three-factor structure was replicated and an acceptable total scale reliability of  $\alpha = .78$  was found. Reliabilities of the subscales varied from poor to acceptable (manageability:  $\alpha = .69$ , reflection:  $\alpha = .74$ , balance:  $\alpha = .54$ ), and acceptable convergent (i.e. assessed with optimism, satisfaction with life, and self-compassion) and discriminant validity (i.e. assessed with negative affect, pessimism, and chronic stress) were found (Mc Gee, Höltge, Maercker, & Thoma, 2018a).

A recent study evaluated the SOC-R in a representative German sample ( $N = 2510$ ;  $M_{age} = 48.2$ ; 47.7% males; Thoma et al., 2018). The overall reliability was good ( $\alpha = .87$ ) and for the subscales scores acceptable to good (manageability:  $\alpha = .72$ , reflection:  $\alpha = .86$ , balance:  $\alpha = .71$ ). The results indicated a moderation effect of SOC-R on the

relation between both emotional neglect and traumatic events on the one hand and depression on the other hand, suggesting a protective effect of high SOC.

All previous studies of the SOC-R have used the proposed three-factor structure in their factor analysis, although the total score rather than the three subscales are used in clinical practice. In line with this, it seems necessary to explore the factorial structure without the restriction of the proposed three-factor structure. This way, we can test whether these three subscales can be derived from our data using exploratory methods. To the best of the authors' knowledge, this study will be the first to apply non-parametric item response theory (NIRT), as well as EFA, to explore the internal structure of the questionnaire.

Furthermore, to allow a more widespread usage of the SOC-R, the goal of the current study is to evaluate the psychometric properties (i.e. internal structure, reliability, and construct validity) of the Dutch SOC-R. Since a validated Dutch SOC-R, to the best of the authors' knowledge, does not exist yet and might have specific clinical value in high-risk populations, the psychometric properties of the Dutch SOC-R were assessed in a sample of 527 firefighters. This is of special interest because firefighters, as other rescue workers, encounter highly stressful events on a daily basis and are regarded as a high-risk group for the development of trauma-related psychopathology (Berger et al., 2012). Once the Dutch SOC-R has good psychometric properties and is validated in this high-risk sample, its predictive value for the development of PTSD can be tested and will further our insight in the potential of SOC trainings to be used in the primary prevention of PTSD.

## 1. Method

### 1.1. Participants and procedure

Data were collected in the first assessment of a larger longitudinal study in collaboration with Dutch firefighters. After providing information about the study, 11 out of 25 safety regions in the Netherlands showed an interest in participation. At selected firefighter stations across these 11 safety regions, firefighters were orally and in writing informed about the study. Selection of fire stations was done by the individual safety regions and was based on a combination of willingness to participate, convenience and practical considerations. Those individuals who were present at the time of testing, who were willing to participate and who provided informed consent, filled out online questionnaires and completed computer tasks (a selection within the scope of the current article will be described below). Participation was voluntary and could be ended at any moment without negative consequences. The study was

approved by the Ethical Committee Psychology of the University of Groningen.

The sample consisted of 527 Dutch firefighters (94.1% males) with a mean age of 39.91 years ( $SD = 10.01$ ). On average, participants had 15.07 years of work experience ( $SD = 9.47$ ), of which 25.8% was volunteer, 42.3% full-time firefighter as profession, and 31.9% both full-time and volunteer.

### 1.2. Measurement instruments

The Dutch SOC-R, based on the 13-item SOC-R (Bachem & Maercker, 2016), was used to measure sense of coherence. For the Dutch translation, native Dutch clinical psychologists translated the SOC-R independently into Dutch. After a back translation into English the group judged which translations were most appropriate to be used for the final Dutch SOC-R. The items are scored on a five-point Likert-scale ranging from 'not at all true' (score '1') to 'very true (score '5') and yield one total score. Item 6 is reversed scored. A higher total score indicates a stronger sense of coherence.

The Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975) was used to assess neuroticism (EPQ-N; 22 items) and extraversion (EPQ-E; 19 items). All items were answered with 'no' (score '0') or 'yes' (score '1'), with items 11, 15 and 22 reversed coded. In the present study, Cronbach's alpha was .87 for the EPQ-N and .81 for the EPQ-E. This questionnaire was used to assess discriminant validity.

The Connor-Davidson Resilience Scale (CD-RISC; Connor & Davidson, 2003) is a 25-item measurement tool to assess resilience. Items were answered on a 5-point Likert-scale ranging from 'not true at all' (score '0') to 'true nearly all the time' (score '4'), based on the participant's experience in the past month. The total score ranges from 0 to 100, with higher scores indicating higher resilience. In the present study, Cronbach's alpha was .85. This questionnaire was used to assess convergent validity.

The 9-item version of the Resilience Evaluation Scale (RES; Van der Meer et al., 2018) was used to assess resilience. Items were rated on a five-point Likert scale from 'completely disagree' (score '0') to 'completely agree' (score '4'). In the current study, Cronbach's alpha was .87. This questionnaire was used to assess convergent validity.

### 1.3. Statistical analysis

The statistical analysis of this study consisted of three parts. A Mokken scale analysis and an EFA were conducted to explore the internal structure of the Dutch SOC-R. Convergent and discriminant validity were also investigated.

### 1.3.1. Item response theory

The core idea of item response theory (IRT) is that psychological constructs are only indirectly observable (i.e. latent). As a consequence, information about a construct can only be gathered by investigating the item response patterns of a group of persons. In order to explain the resulting manifest responses, IRT models assume a latent trait ( $\theta$ ) and each person and item to be located on the same latent trait continuum (Embretson & Reise, 2000; Sijtsma & Molenaar, 2002).

IRT models can be distinguished in parametric and nonparametric models. Compared to parametric models, nonparametric item response theory (NIRT) models have less restrictive assumptions with regard to empirical data (i.e. no logistic model is assumed, and requires smaller sample sizes). Therefore, NIRT models are a suitable method to investigate the internal structure of questionnaires. For a more detailed discussion, see Junker and Sijtsma (2001), Meijer and Baneke (2004), or Sijtsma and Molenaar (2002).

### 1.3.2. Mokken scale analysis

In the current study, the nonparametric Mokken's monotone homogeneity (MMH) model is applied to explore the internal structure of the Dutch SOC-R. Coefficient  $H$  is used as an index of scalability of a set of items and reflects the discriminative power of an item set. Therefore, higher  $H$  values lead to a more precise ordering of respondents with regard to their total scores. Furthermore, the  $H$  value is an indication of how well the items are related to each other and form a coherent scale. Coefficient  $H_i$  is used as an index of scalability of single items and reflects how well an item  $i$  fits to the other  $k - 1$  items in a scale (Egberink & Meijer, 2011; Meijer & Baneke, 2004). Sijtsma and Molenaar (2002) recommended the following rules of thumb for practical test construction; scalability is regarded weak if  $.3 \leq H < .4$ , medium if  $.4 \leq H < .5$ , and strong if  $.5 \leq H < 1$ . The default lower bound was used for both test scalability ( $H = .30$ ) and the items ( $H_i = .30$ ).

Mokken Scale Analysis for Polytomous Items 5.0 (MSP5) was used to conduct the NIRT analyses. The 'Test'-function was used for a preliminary investigation of the entire 13-item scale. In order to investigate the internal structure, the 'Search normal'-function was used. This automatic stepwise bottom-up procedure starts with the item pair with the highest  $H$ -coefficients. In the next step, items with the highest  $H_i$ -values are added one-by-one to the scale, as long as they are above the cut-off score of  $H_i = .30$ . For the analysis the following settings were used: lower bound  $H_i = .30$ ,  $\alpha = .05$  and maximum number of scales set to 6.

### 1.3.3. Exploratory factor analysis

The programme FACTOR (Lorenzo-Seva & Ferrando, 2006) was used to conduct EFA. More specifically, a parallel analysis (PA) with a 95% threshold based on

a minimum rank analysis (PA\_MRFA) was conducted.<sup>1</sup> Although both PA\_MRFA as PA based on principal component analysis (PA\_PCA) perform well in the analysis of polytomous items, PA\_MRFA is slightly superior, because PA\_PCA shows more irregular results compared to PA\_MRFA (Timmerman & Lorenzo-Seva, 2011). Therefore, PA\_MRFA was used with polychoric correlations and the Promin criterion for oblique rotation.

### 1.3.4. Convergent and discriminant validity

Convergent validity was assessed by calculating bivariate Pearson correlations between the scores of the Dutch SOC-R and two resilience measures (i.e. CD-RISC and RES). Since construct validity is primarily concerned with the accumulation of research evidence and construct validation research is never completed, we decided to use two instead of one resilience instrument. Based on the correlation between the scores on the short version, 13 item French SOC and the Resilience Scale for Adults (i.e.  $r = .68$ ; Fossion et al., 2014), it was hypothesized that the Dutch SOC-R is positively associated with the CD-RISC and the RES (i.e.  $r$  around  $.60$ ). To assess discriminant validity, total scores from the Dutch SOC-R, the EPQ-N (neuroticism), and the EPQ-E (extraversion), were correlated. Research on the SOC-R reported low correlations with neuroticism ( $r = -.23$  in Bachem & Maercker, 2016 and  $r = -.32$  in; Mc Gee et al., 2018a) and extraversion ( $r = .22$  in Bachem & Maercker, 2016 and  $r = .20$  in; Mc Gee et al., 2018a). Therefore, it is hypothesized that both EPQ-N and EPQ-E total scores will have low correlations with Dutch SOC-R total scores (i.e.  $r$  around  $-.25$  for neuroticism and  $r$  around  $.20$  for extraversion).

## 2. Results

### 2.1. Descriptive statistics

Table A1 depicts the descriptive statistics for all 13 items and the reliability of the total scale (both Guttman's lambda-2 and Cronbach's alpha). In general, Guttman's lambda-2 is a more accurate reliability estimate compared to Cronbach's alpha. According to the rules of thumb formulated by the Dutch Committee on Tests and Testing (COTAN) in their 'COTAN review system for evaluating test quality' (Evers, Lucassen, Meijer, & Sijtsma, 2010), reliability coefficients equal to or larger than  $.80$  are considered 'good' for tests that are used for relatively less important decisions at the individual level, for example general descriptive use of test scores for admission to therapy. The participants' mean total score of 48.99 ( $SD = 6.31$ ) indicated medium to high SOC. The item-rest correlations were between  $.34$  and  $.69$ , except for item 6 ( $r_{ir} = -.03$ ). According to the rules of thumb formulated by the COTAN

(Evers et al., 2010), item-total correlations equal to or larger than .30 are considered 'good'. This was a first indication that all items, except for item 6, relate relatively well to the total score and discriminate satisfactorily.

## 2.2. Mokken scale analysis

The  $H$  coefficient,  $H_i$  coefficients and reliability (i.e. rho) based on the 'Test'-function for all items of the Dutch SOC-R are displayed in column 2 of Table A2. It was apparent that the Dutch SOC-R with all 13 items did not perform well. Although, the reliability of the scale was sufficient ( $\rho = .82$ ), the overall scalability was weak ( $H = .30$ ). Especially, item 6 was not related to the other scale items ( $H_6 = -.02$ ), which indicated that item 6 should be removed. In addition, item 2 ( $H_2 = .22$ ), item 3 ( $H_3 = .24$ ), item 4 ( $H_4 = .27$ ), and item 10 ( $H_{10} = .27$ ) might also be excluded from the scale, due to  $H_i < .30$ .

Column 3 of Table A2 depicts the results of the exploratory 'Search normal'-function; the scale was reduced to 10 items. As expected, item 6 was excluded due to a negative  $H_i$  coefficient. Also, Item 2 ( $H_2 = .26$ ) and item 3 ( $H_3 = .27$ ) were excluded because  $H_i < .30$ . As a result of excluding these items, the test scalability and reliability increased ( $H = .41$ ;  $\rho = .85$ ) compared to the 13-item scale (i.e.  $H = .31$ ;  $\rho = .83$ ).

## 2.3. Exploratory factor analysis

PA\_MRFA suggested one factor. Table A3 displays the factor loadings and the variance accounted for (VAF). Also in this analysis, item 6 did not seem to fit with the other items. Given the Mokken scale results the PA\_MRFA results, PA\_MRFA was repeated for the 10-item scale (i.e. without items 2, 3, and 6). PA\_MRFA again suggested one factor; factor loadings and VAF can be found in column 3 of Table A3. The reliability estimates for the 10-item version increased compared to the 13-item version (13-item:  $\alpha = .82$ ,  $\lambda_2 = .83$ ; 10-item:  $.85$  for both  $\alpha$  and  $\lambda_2$ ). Although the factor loadings remained similar for the 10-item version, VAF increased from 62.9% to 73.8%.

## 2.4. Combining results of the structural analyses

These results combined with the content of the items suggested that the one-factor solution was most convincing and that the total score on 10 out of 13 items of the Dutch SOC-R might be a direct measure of the individuals' SOC level in our firefighter sample. All analyses suggested that item 6 'Difficult situations overstrain me' did not function properly and

should therefore be excluded. One explanation might be that 'overstrained' is a strong word, which induces strong emotional reactions, causing an automatic rejection of this statement; 75.5% of the participants chose 'not at all true' or 'slightly true'. The item might oppose the subcultural role of firefighters who need to be strong in face of challenging situations. Another explanation could be that firefighters are resilient and only decide to become a firefighter if they are not easily overstrained by stressful situations. These potential explanations might be interesting for further empirical investigation.

Although items 2 and 3 were only excluded in the Mokken scale analysis, the item-rest correlations (Table A1, column 6) and PA\_MRFA results (Table A3, column 2) also indicated that these two items performed worse compared to the other scale items. The reason for exclusion of item 3, 'I could suddenly experience something really horrible or shocking' was similar to item 6 in the Mokken scale analysis. Possibly, firefighters develop a general vigilance state, because their profession includes exposure to horrible and shocking situations. This is supported by recent evidence that firefighters showed higher physiological arousal baselines compared to policemen and the general population (Salters-Pedneault, Ruef, & Orr, 2010). Furthermore, this explanation fits with the result that item 3 had the highest mean of 4.11 (81.2% chose either 'quite true' or 'very true').

The exclusion of item 2, 'Evil has also its place in the world', based on the Mokken scale analysis had a different explanation. It could be argued that the question is somewhat open for interpretation and requires an elaborated reflection process to rate the item accurately. Therefore, many participants may have been unsure how to rate this statement and rather chose the middle option ( $M = 3.39$ ).

## 2.5. Convergent and discriminant validity

The convergent and discriminant validity of the 10-item version of the Dutch SOC-R were assessed by correlating its total score with those of the CD-RISC and RES, and with the EPQ-N and EPQ-E, respectively. As expected, the Dutch SOC-R was positively related to the resilience measures ( $r_{CD-RISC} = .46$ ,  $p < .01$ ;  $r_{RES} = .37$ ,  $p < .01$ ) and showed low correlations with neuroticism ( $r_{EPQ-N} = -.19$ ,  $p < .01$ ) and extraversion ( $r_{EPQ-E} = .11$ ,  $p < .01$ ). These results support the construct validity of the Dutch SOC-R and are also in line with the formulated hypotheses that the Dutch SOC-R is positively associated with the CD-RISC and the RES. The discriminant validity coefficients were low (i.e. negative and/or between .00 and .20), as one would expect. The convergent validity coefficients were positive as hypothesized and with values of  $r = .37$  and  $r = .46$  could be interpreted as moderate.

### 3. Discussion

The aim of the current study was to evaluate the psychometric properties of the Dutch SOC-R in a firefighter sample. This is the first step to further investigate the potential value of SOC-R in the primary prevention of trauma-related psychopathology, since SOC might be a protective factor in developing trauma-related psychopathology and seems to be adaptable through training. The strengths of the current study include the use of a high-risk sample, the application of a non-parametric IRT approach to evaluate the psychometric properties of the Dutch SOC-R, and the combination of assessing internal structure and validity of the Dutch SOC-R. Our results suggested that a one-factor solution with 10 out of 13 items was most compelling, based on the PA\_MRFA analysis, additional evidence from the Mokken scaling, and the content of the items. Reliability estimates also increased for the 10-item version. Therefore, this is the first study to provide psychometric support for the use of the total score in clinical practice. With regard to the validity, as hypothesized, the Dutch SOC-R related positively to two resilience measures (convergent validity) and showed low correlations with neuroticism and extraversion (discriminant validity).

The one-factor solution of the Dutch SOC-R is different from the three-factor structure of balance, reflection, and manageability from previous studies (Bachem & Maercker, 2016; Mc Gee et al., 2018a; Thoma et al., 2018). Interestingly, on an item level there were similarities with these previous studies, as items 2, 3, and 6 did not function well in those studies either. Thoma et al. (2018) also found insufficient factor loadings of item 6 ('manageability'-subscale) and co-variances with item 8. Their conclusion was that item 6 does not reflect the essence of manageability and therefore should be rephrased to 'I perceive difficult situations as challenges to be/I can overcome'. However, so far this item was not rephrased and tested. Therefore, the exclusion of item 6 in high-risk profession samples is advised. Item 2 ('Evil has also its place in the world') and item 3 ('I could suddenly experience something really horrible or shocking') are both part of the 'balance'-subscale. Previous research already recognized that the reliability of the 'balance'-subscale is suboptimal and that the two items covary to some degree (Bachem & Maercker, 2016; Mc Gee et al., 2018a). It was argued, that it resulted from small sample sizes and non-representative samples. Therefore, Mc Gee et al. (2018a) suggested a rephrasing or replacement of item 3, but without a concrete example. Our results supported the exclusion of items 2 and 3.

The positive relationship between the Dutch SOC-R and resilience is in line with previous studies (Mc Gee, Hölte, Maercker, & Thoma, 2018b). It suggests that these two concepts overlap to a certain degree, but can also be considered as distinct constructs. That is, the constructs have a positive correlation suggesting

a certain overlap, but the correlation is moderate which suggests a unique part as well. Furthermore, the associations between the Dutch SOC-R and the personality measures were weaker than in other studies (Bachem & Maercker, 2016; Mc Gee et al., 2018a; Thoma et al., 2018). This supports the SOC-R conceptualization of Bachem and Maercker (2016) that SOC should be viewed as a coping ability rather than a personality trait. This is further supported by the fact that SOC is changeable through direct interventions, which is possible for coping abilities, but rather difficult for personality traits (Koutsoukou-Argyriaki et al., 2018).

The limitations of this study might help to guide future research. First, the current firefighter sample is a highly specific (mostly male) sample. This might compromise the generalizability of the results to other (non high-risk) samples. Second, no information about intra-individual changes is available, which might give valuable information about the stability of the Dutch SOC-R. Future research should evaluate the Dutch SOC-R in representative samples, as well as in other high-risk professions to replicate the results. Replication might strengthen the current conclusion that the SOC-R needs adjustments for high-risk samples. To investigate the clinical validity of SOC, it is important to test whether SOC-R has predictive value in the development of trauma-related psychopathology. Lastly, it might be worthwhile to investigate different methods to actively increase SOC-R scores through interventions in high risk samples and test its effect on the development of psychopathology.

In sum, this study evaluated the psychometric properties of the Dutch SOC-R and was the first study to apply a non-parametric IRT approach for this aim. The one-factor solution of the Dutch SOC-R with 10 items (item 2, item 3, and item 6 were excluded) is most convincing. This contradicts the three-factor solution of previous research on the SOC-R. This might be attributable to the specific sample used in this study, that is firefighters, as this sample is different from the samples used in previous studies. This evaluation in such a high-risk sample is important because of its potential clinical value: a Dutch SOC-R with good psychometric properties is a requirement to assess the predictive value of SOC-R in the development of posttraumatic stress disorder in high-risk samples, which can consequently inform us about the potential of SOC training programmes as primary prevention of trauma-related psychopathology.

### Note

1. PA\_PCA was also performed, but due to superiority of PA\_MRFA, only results of PA\_MRFA will be reported here.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Iris J. L. Egberink  <http://orcid.org/0000-0003-0191-2733>  
Miriam J. J. Lommen  <http://orcid.org/0000-0001-8845-4338>

## References

- Antonovsky, A. (1979). *Health, stress, and coping*. San Francisco, CA: Jossey-Bass.
- Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well*. San Francisco, CA: Jossey-Bass.
- Antonovsky, A. (1993). The structure and properties of the sense of coherence scale. *Social Science & Medicine*, 36(6), 725–733.
- Bachem, R., & Maercker, A. (2016). Development and psychometric evaluation of a revised sense of coherence scale. *European Journal of Psychological Assessment*, 34(3), 206–215.
- Berger, W., Coutinho, E. S. F., Figueira, I., Marques-Portella, C., Luz, M. P., Neylan, T. C., ... Mendlowicz, M. V. (2012). Rescuers at risk: A systematic review and meta-regression analysis of the worldwide current prevalence and correlates of PTSD in rescue workers. *Social Psychiatry and Psychiatric Epidemiology*, 47(6), 1001–1011.
- Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and Anxiety*, 18(2), 76–82.
- Egberink, I. J. L., & Meijer, R. R. (2011). An item response theory analysis of Harter's self-perception profile for children or why strong clinical scales should be distrusted. *Assessment*, 18(2), 201–212.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Eriksson, M., & Lindström, B. (2005). Validity of Antonovsky's sense of coherence scale: A systematic review. *Journal of Epidemiology and Community Health*, 59(6), 460–466.
- Eriksson, M., & Lindström, B. (2007). Antonovsky's sense of coherence scale and its relation with quality of life: A systematic review. *Journal of Epidemiology and Community Health*, 61(11), 938–944.
- Evers, A. V. A. M., Lucassen, W., Meijer, R. R., & Sijsma, K. (2010). *COTAN Beoordelingssysteem voor de Kwaliteit van Tests* (geheel herziene versie, mei 2019; gewijzigde herdruk mei 2010) [COTAN Rating system for test quality (completely revised version, May 2009, reprinted 2010)]. Amsterdam: NIP. Retrieved from <https://www.psynip.nl/en/dutch-association-psychologists/about-nip/psychological-testing-cotan/>
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck personality questionnaire*. San Diego, CA: Educational and Industrial Testing Service.
- Feldt, T., Metsäpelto, R., Kinnunen, U., & Pulkkinen, L. (2007). Sense of coherence and five-factor approach to personality: Conceptual relationships. *European Psychologist*, 12(3), 165–172.
- Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist*, 18(1), 12–23.
- Fossion, P., Leys, C., Kempnaers, C., Braun, S., Verbanck, P., & Linkowski, P. (2014). Disentangling sense of coherence and resilience in case of multiple traumas. *Journal of Affective Disorders*, 160, 21–26.
- Gruszczynska, E. (2006). What is measured by the orientation to life questionnaire? Construct validity of the instrument for the sense of coherence measurement. *Polish Psychological Bulletin*, 37(2), 74–83.
- Joyce, S., Shand, F., Tighe, J., Laurent, S. J., Bryant, R. A., & Harvey, S. B. (2018). Road to resilience: A systematic review and meta-analysis of resilience training programmes and interventions. *BMJ Open*, 8, e017858.
- Junker, B. W., & Sijsma, K. (2001). Nonparametric item response theory in action: An overview of the special issue. *Applied Psychological Measurement*, 25(3), 211–220.
- Koutsoukou-Argyryaki, A., Bertolino, A., Krefß, V., Hofmann, F., Sperth, M., & Holm-Hadulla, R. (2018). Veränderungen von Kohärenzgefühl, Symptombelastung und Lebens- sowie Studienzufriedenheit im Verlauf psychosozialer Beratungen. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 68(2), 66–74.
- Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, 38(1), 88–91.
- Luthar, S. S., & Cicchetti, D. (2000). The construct of resilience: Implications for interventions and social policies. *Development and Psychopathology*, 12(4), 857–885.
- Mc Gee, S. L., Hölzge, J., Maercker, A., & Thoma, M. V. (2018a). Evaluation of the revised sense of coherence scale in a sample of older adults: A means to assess resilience aspects. *Ageing & Mental Health*, 22(11), 1438–1447.
- Mc Gee, S. L., Hölzge, J., Maercker, A., & Thoma, M. V. (2018b). Sense of coherence and stress-related resilience: Investigating the mediating and moderating mechanisms in the development of resilience following stress or adversity. *Frontiers in Psychiatry*, 9, 378.
- Meijer, R. R., & Baneke, J. J. (2004). Analyzing psychopathology items: A case for nonparametric item response theory modeling. *Psychological Methods*, 9(3), 354–368.
- Salters-Pedneault, K., Ruef, A. M., & Orr, S. P. (2010). Personality and psychophysiological profiles of police officer and firefighter recruits. *Personality and Individual Differences*, 49(3), 210–215.
- Sijsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. Thousand Oaks, CA: Sage.
- Southwick, S. M., Bonanno, G. A., Masten, A. S., Panter-Brick, C., & Yehuda, R. (2014). Resilience definitions, theory, and challenges: Interdisciplinary perspectives. *European Journal of Psychotraumatology*, 5(1), 25338.
- Thoma, M. V., Mc Gee, S. L., Fegert, J. M., Glaesmer, H., Brähler, E., & Maercker, A. (2018). Evaluation of the revised sense of coherence scale in a representative German sample. *PLoS One*, 13(12), e0209550.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209–220.
- Van der Meer, C. A. I., Te Brake, H., van der Aa, N., Dashtgard, P., Bakker, A., & Olf, M. (2018). Assessing psychological resilience: Development and psychometric properties of the English and Dutch version of the resilience evaluation scale (RES). *Frontiers in Psychiatry*, 9, 169.

## Appendix A

**Table A1.** Dutch SOC-R item descriptive statistics and reliability for the total sample.

Item	Mean	SD	% not at all true	% slightly true	% some-what true	% quite true	% very true	$r_{ir}$
1	4.08	0.77	0.8	2.3	14.8	52.8	29.4	.51
2	3.39	1.02	5.1	11.8	34.2	36.4	12.5	.34
3	4.11	0.98	2.7	5.3	10.8	40.8	40.4	.36
4	3.10	1.00	6.6	18.4	40.2	27.5	7.2	.41
5	3.55	0.93	3.2	9.9	27.3	48.2	11.4	.49
6	4.07	0.90	0.6	4.9	19.0	38.5	37.0	-.03
7	3.78	0.87	0.8	8.9	19.7	52.9	17.6	.58
8	3.94	0.79	1.7	2.8	15.6	59.2	20.7	.62
9	3.65	0.87	0.9	9.3	27.9	47.8	14.0	.50
10	3.84	0.73	0.6	3.4	22.0	59.4	14.6	.43
11	3.87	0.75	1.1	2.8	19.7	60.2	16.1	.64
12	3.77	0.83	1.1	5.7	24.5	52.6	16.1	.61
13	3.85	0.75	0.6	4.0	21.3	58.4	15.7	.69
$\lambda_2/\alpha$								.83/.82

$\lambda_2$  = Guttman's lambda-2;  $\alpha$  = Cronbach's alpha; SD = standard deviation;  $r_{ir}$  = item-rest correlation.

**Table A2.** Results of Mokken scale analysis for 'Test'-function and 'Search'-function.

Item	Test	Search
	$H_i$	$H_i$
1	.34	.39
2	.22	
3	.24	
4	.27	.30
5	.32	.38
6	-.02	
7	.36	.42
8	.40	.47
9	.32	.38
10	.27	.33
11	.40	.48
12	.38	.46
13	.43	.50
H	.30	.41
$\alpha/\rho$	.82/.82	.85/.85

H = scalability coefficient for the total scale;  $H_i$  = scalability coefficient for the item;  $\alpha$  = Cronbach's alpha;  $\rho$  = rho.

**Table A3.** Factor loadings for PA-MRFA for the Dutch SOC-R items based on 13 items and 10 items.

Item	13 items	10 items
	Factor 1	Factor 1
1	.61	.59
2	.42	n.a.
3	.43	n.a.
4	.47	.46
5	.63	.61
6	.01	n.a.
7	.69	.68
8	.76	.77
9	.60	.60
10	.55	.53
11	.80	.82
12	.76	.78
13	.85	.87
VAF	62.9%	73.8%

PA\_MRFA = a parallel analysis with a 95% threshold based on a minimum rank analysis; n.a. = not applicable; VAF = variance accounted for; bold factor loadings are highest on the corresponding factor.