BOX 1: ***NEOSi development according to the first 3 steps of Rattray & Jones***[1]

Step 1: Specifying assessment goals

For the clinical practice, the NEOSI should have the potential to:

1. improve clinical practice by facilitating discussion and striving to mutual understanding between the patient and her physicians concerning the nature and extent of unbearable suffering as experienced by psychiatric patients.
2. support physicians in their evaluation of the key criterion 'unbearable suffering' in psychiatric patients with and without euthanasia request by providing a comprehensive overview of the patients' suffering experiences within a variety of life domains, which can then be employed for preventive and curative perspectives.
3. enable new ways of gaining deeper insight into the potential of (avoiding) suffering becoming unbearable.

For further research endeavours, the NEOSI must meet the following goals:

1) to exhaustively cover and map the nature and extent of suffering experiences;
2) to enable examining differences between suffering experiences from psychiatric patients with or without a euthanasia request; and
3) to enable examining differences between suffering experiences from psychiatric patients with granted, denied, persistent or withdrawn euthanasia requests.

Step 2: Item generation

The initial pool of items was derived from the results of the qualitative study on unbearable suffering inciting psychiatric patients to request for euthanasia. In this study, 'testimonials' from psychiatric patients with a euthanasia request were analysed. In total, 71 descriptive themes on suffering experiences were distinguished: 62 aspects (see list A in OSF) and nine indicators of unbearable suffering experiences (see list B in OSF). The 62 aspects were grouped into a hierarchical tree structure with the following three layers: 35 inductively sorted codes capturing each suffering aspect in its essence; sorted in 12 more abstract descriptive subcategories and then generally classified in 5 overarching main categories (suffering domains), see list A in OSF.

Deductive methods were used in the development of the new questionnaire. The 71 reported terms that were used to describe the nature (62 aspects) and extent (nine indicators) of the suffering experience were rephrased into two sets of questions: one set to map the nature of the suffering experience (the sub-questionnaire Nature Of Suffering index, NOSi), and one set to map the extent of the suffering experience (the sub-questionnaire Extent of Suffering index, EOSi).

In this study, these lists were thus used as a starting point for the development of the NEOSI as the codes from list A were converted/transferred into questions for the sub-questionnaire NOSi and from list B into questions and anchors (by means of scale construction) for the sub-questionnaire EOSi. These questions were formulated in accordance with six formal and stylistic criteria as listed in Drenth & Sijtsma[2]: 1) not more than 20 words per question; 2) formulated in accessible language; 3) double negations avoided; 4) content should be formulated unambiguously; 5) questions should cover but one theme, and 6) items to be confirmed by everyone or no one should be eliminated. The questions were listed in WORD (see 'NEOSI_ItemDevelopment' in OSF, in Dutch and English) and presented to experts with experience in the euthanasia practice (as a researcher or psychiatrist) and adjusted to their feedback.

Step 3: Scale construction

As for the sub-questionnaire NOSi, response registration was based on the SOS-V measurement instrument, specifically addressing frequency and intensity of each aspect in the NOSi. In formulating the answer options, the research team attempted to balance precision and objectivity (e.g. sharply delineating time for the frequency measurements) with concreteness and emotional safety (intensity). Each item was measured via LimeSurvey's 'Double Array Scale' question type, which presents two side-by-side Likert scales for each question. Each of the five clusters (medical, intrapersonal, interpersonal, societal and existential) was followed by an invitation to answer the open question whether or not items were missing addressing additional contributors related to patients' suffering experiences within the specific cluster. The answer scales were chosen in accordance to literature study. A frequently used and user-friendly Likert scale for the NOSi was used. A 4-point or 5-point Likert scale is more sensitive to nuanced answer options than a 3-point scale or less, though less confusing and complicated than a 7-point scale.[2] Although a 5-point Likert is the most popular scale, the decision to use a 4-point Likert scale was used to measure the most subjective concept 'intensity' to reduce central tendency bias.

As for the EOSI, each of the 9 indicators from list B (see OSF) and its opposite were used in a Visual Analogue Scale (VAS) as 'opposite anchor answer options' (e.g. "hopeless" versus "hopeful"). These anchors were preluded with 9 questions that assessed the extent of suffering experiences with respect to intensity, duration, chronicity and perspectives. The Visual Analogue Scale (VAS) used a slider with minimum value of 0 and maximum value of 100. A VAS-scale for EOSi was chosen as the exploration of the extent of suffering implied more gradual answers on a continuum of many stages between one extreme versus another. As a short VAS-scale has often been used in research on pain, perceived quality of life, and changes in response on medical treatments, it also generates high face validity when straightforwardly examining patients' experiences.[3]

1. Rattray J, Jones MC. Essential elements of questionnaire design and development. *J Clin Nurs*. 2007;16(2):234-243. doi:10.1111/j.1365-2702.2006.01573.x.
2. Drenth PJD, Sijtsma K. *Test Theorie (Test Theory)*. 4th ed. Houten, the Netherlands: Bohn Stafleu Van Loghum; 2006.
3. Miller SD, Duncan BL, Brown J, Sparks JA, Claud DA. A Preliminary Study of the Reliability , Brief Visual Analog Measure. *J Br Ther*. 2003;2(2). https://www.researchgate.net/profile/Scott_Miller12/publication/242159752_The_Outcome_Rating_Scale_A_Preliminary_Study_of_the_Reliability_Validity_and_Feasibility_of_a_Brief_Visual_Analog_Measure/links/02e7e52e7b2daab9a1000000/The-Outcome-Rating-Scale-A-P.