Translation and linguistic validation of the German Challenging Behaviour Scale for formal caregivers of people with dementia in nursing homes

Daniela Holle PhD[1,2], Lena Köller, PhD[3], Esme-Moniz-Cook, PhD[4], Margareta Halek PhD*[1,2]

[1]German Center for Neurodegenerative Diseases (DZNE), Stockumer Str. 12, 58453 Witten, Germany

[2]School of Nursing Science, Witten/ Herdecke University, Witten, Germany

[3]Department of hematology and oncology, Knappschaftskrankenhaus, Ruhr-University Bochum, Bochum, Germany

[4]Hull University, Yorkshire, Great Britain

*Corresponding author

Email addresses:

DH: email: daniela.holle@dzne.de, phone: + 49 2302 926-347; fax: + 49 2302 926-239

LK: email: lena.koeller@kk-bochum.de

EMC: email: E.D.Moniz-Cook@hull.ac.uk

MH: email: margareta.halek@dzne.de

Acknowledgements

The study was funded by the German Center for Neurodegenerative Diseases (DZNE), site Witten. The funders had no roles in the study design, data collection, management, analysis, interpretation, manuscript writing or decision to submit the report for publication. Non-financial competing interests: EMC developed the original CBS.

Ethical approval and consent to participate

The Institutional Review Board for Ethics in Research, German Society for Science in Nursing (E-DGP) approved this study. Informed consent was obtained from each participant at the beginning of the study.

Ordering Information and Citation of the German Challenging Behaviour Scale

The German CBS and its user manual is free of charge and is available upon request from the authors of this manuscript.

Copyright

Introduction

Behavioural and psychological symptoms (BPSDs) affect most persons with dementia at some point during the progression of their disease. BPSDs include a broad range of symptoms, such as wandering, physical aggression, screaming, depression, resistance to help with activities of daily living, suspiciousness, accusation, and insomnia (Brodaty & Arasaratnam, 2012). Prevalence studies indicate that more than 80% of nursing home residents with dementia display at least one BPSD (Selbaek, Engedal, & Bergh, 2013), indicating the urgency of addressing such symptoms. Compared with loss of memory, cognitive decline, and functional disabilities, BPSDs are more difficult to manage, and professional caregivers often experience BPSDs as a burden (Edvardsson, Sandman, Nay, & Karlsson, 2008; Schmidt, Dichter, Palm, & Hasselhorn, 2012). BPSDs are also referred to as challenging behaviours (Moniz Cook et al., 2012). This concept considers the consequences of behaviour – distress or suffering – as defining characteristics. Caregivers' emotional response to patient behaviour is often determined by personal attributes, such as staff variables in care homes, rather than by the behaviour itself. Thus, what is considered challenging is often subjective (Bird & Moniz Cook, 2008). Research further indicates that challenging behaviour often emerges from incongruence between a person's needs and the degree to which the environment fulfils those needs (Cohen-Mansfield, Golander, Ben-Israel, & Garfinkel, 2011). However, challenging behaviour can be defined as a manifestation of caregiver distress or of the distress and suffering of the person with dementia (Bird & Moniz Cook, 2008). Consequently, not every BPSD is necessarily a challenging behaviour.

Psychosocial interventions are recommended as the "first line of treatment" for people with dementia and challenging behaviour (National Institute for Health and Care

Excellence [NICE] & Social Care Institute for Excellence [SCIE], 2006; Vasse et al., 2012). To investigate the effect of such psychosocial interventions on nursing home residents' challenging behaviour, a comprehensive scale is needed that incorporates both the problematic resident behaviours and challenges experienced by formal caregivers and residents. The purpose of this study is to translate the English version of the CBS into German and subsequently test the linguistic validity of the newly developed German Challenging Behaviour Scale (CBS-G) in a pilot sample of formal caregivers.

Background and conceptual framework

There are more than 80 scales assessing the BPSDs of people with dementia, but only some are used frequently in research (Gitlin, Marx, Stanley, Hansen, & Van Haitsma, 2014; Jeon et al., 2011; van der Linde, Stephan, Dening, & Brayne, 2014). These measurements cover different aspects of BPSD evaluation: the behaviour's frequency, presence/absence, severity, and received reaction (caregivers or persons with dementia), and the measurements also assess a wide range of behaviours. The neuropsychiatric inventory (NPI) is the most comprehensive measurement available and consists of 12 behavioural domains that assess the frequency and severity of the behaviour and the reaction to the behaviour. The NPI is based on a medical model of behaviour as a symptom regardless of the meaning or cause of the behaviour (Lai, 2014). Thus, a meaningful behaviour could be incorrectly assessed as a neuropsychiatric symptom. The NPI requires careful training and adherence to the manual (Connor, Sabbagh, & Cummings, 2008). In contrast, the theoretical framework of the "Challenging Behaviour Scale (CBS)" is the bio-psychosocial approach to the aetiology of behaviour, implying

that behaviour is not just a symptom of a disease but a responsive reaction to intrinsic and extrinsic factors (Moniz-Cook, Woods, Gardiner, Silver, & Agar, 2001).

Another advantage of the CBS is that the content and language were developed specifically for the nursing setting. This scale reflects the perspective of the nurses regarding challenging behaviours and uses their own words (Moniz-Cook et al., 2001). The scale is applied in both research (Koder, Hunt, & Davison, 2014; Orrell et al., 2007; Wenborn et al., 2013) and nursing practice (Duffy, 2016; Warwick, Higgon, & Edgar, 2011). The CBS incorporates 25 items and four measures. Three of the measures are rated by the caregivers, and the fourth measure is a computed score. First, caregivers assess whether the resident has displayed the behaviour in the past 8 weeks, producing the number of problem behaviours (incidence). Subsequently, the frequency of the behaviour (0 = never present; 1 = occasionally present – less than once a month; 4 = present daily) and its severity (1 = minimal management difficulty; 4 = extreme management difficulty) are rated on a 4-point scale. Scores are calculated for the total number (1-25), frequency (range 0-100) and difficulty (0-100) of challenging behaviours displayed by the person with dementia. The fourth measure indicates the total level of challenge (range 0-400) as the sum of the products of the frequency and difficulty ratings for each behavioural item on the scale (Moniz-Cook et al., 2001).

The psychometric properties of the CBS are promising. The internal consistency of the English version ranges between 0.82 and 0.85 (incidence 0.85; frequency 0.82; difficulty 0.87; challenge 0.85), demonstrating adequate validity and good test-retest reliability. The inter-rater reliability is good when staff receive training (incidence 0.93, frequency 0.88, difficulty 0.82, challenge 0.72) or pairs of staff groups (untrained) of at least three, of mixed qualifications, completed the ratings (incidence 0.96; frequency 0.97; difficulty 0.84; challenge 0.85) (Moniz-Cook et al., 2001).

The internal consistency of the Chinese version ranged between 0.84 and 0.88; the test-retest reliability between 0.96 and 0.98; and the inter-rater reliability between 0.79 and 0.85 when completed by two carers after discussion and consensus. The concurrent validity between the subscales of Chinese CBS and the Cohen-Mansfield Agitation Inventory (CMAI) is good (rho = 0.73–0.86) (Lam, Chan, Mok, Li, & Lam, 2006).

~~The internal consistency of the Chinese version ranged between 0.84 and 0.88; the test-retest reliability between 0.96 and 0.98; and the inter-rater reliability between 0.79 and 0.85 when completed by two carers after discussion and consensus. The concurrent validity between the subscales of CBS and the Cohen-Mansfield Agitation Inventory (CMAI) is good (rho = 0.73–0.86) (Lam, Chan, Mok, Li, & Lam, 2006).~~

~~The CBS incorporates 25 items and four measures. Three of the measures are rated by the caregivers, and the fourth measure is a computed score. First, caregivers assess whether the resident has displayed the behaviour in the past 8 weeks, producing the number of problem behaviours (incidence). Subsequently, the frequency of the behaviour (0 = never present; 1 = occasionally present – less than once a month; 4 = present daily) and its severity (1 = minimal management difficulty; 4 = extreme management difficulty) are rated on a 4-point scale. Scores are calculated for the total number (1-25), frequency (range 0-100) and difficulty (0-100) of challenging behaviours displayed by the person with dementia. The fourth measure indicates the total level of challenge (range 0-400) as the sum of the products of the frequency and difficulty ratings for each behavioural item on the scale (Moniz-Cook et al., 2001).~~

~~To date, the CBS is only available in English and Chinese (Lam et al., 2006; Moniz-Cook et al., 2001), which limits its use to English- and Chinese-speaking countries. Thus, this study aims to translate the English version of the CBS into German and subsequently test its linguistic validity in a pilot sample of formal caregivers.~~

Methods

Translation

For the translation of the CBS, the two-panel approach proposed by Swaine-Verdier, Doward, Hagell, Thorsen, and Mc Kenna (2004) was used. The method involves dual translation panels to produce a high-quality forward translation of a health-related assessment instrument. To ensure a high-quality forward translation, the method recommends recruiting five to seven translators with varied profiles to work as a team in a group meeting. The translation team should be informed about the model underlying the instrument, the way it was developed, its design and content, its target population, as well as the translation requirements (conceptual equivalence, acceptability of wording). The authors of the two-panel approach further suggest to supervise the translation process by an experienced coordinator (Swaine-Verdier et al., 2004). Once the translated version of the instrument has obtained consent in the translation team, the instrument will be assessed by a lay panel working as a focus group in the target language only. The lay panel should only receive the translated version of the instrument because their assessment should not be biased by what they think the translated items should mean but rather what they do mean. The lay panel should be facilitated by the same coordinator involved in the first panel to ensure that the original meaning of the items and instrument structure are retained. Finally, the entire translation process should be reported with specific explanations of changes made following lay panel testing (Swaine-Verdier et al., 2004).

The following section describes the realisation of these recommendations during the translation process of the CBS:

Preparation of forward translation and conduction of translation panel

In preparation of the forward translation of the CBS, English definitions were developed for all 25 items of the CBS to achieve a common understanding of their content, as the original items of CBS either include no definition or were solely described by single examples. The definitions were submitted to the author of the original version of the CBS (EMC) for confirmation or correction.

Additionally, qualification profiles for translators were developed based on the recommendations in the literature (Barandun Schäfer et al., 2009; Eberl & Bartholomeyczik, 2010; Hilton & Skrutkowski, 2002) and accounted for the particular requirements of the CBS (professional experiences in a) dementia and b) with the challenging behaviours of persons with dementia). A former care nurse was recommended to be involved in the translation process as representatives of the target group of formal caregivers in Germany. Persons matching these profiles were contacted and asked to prepare a translation proposal of the CBS.

To translate the instructions, scoring system and background of the CBS, we focused on the comprehensibility of the text passages rather than on absolute equivalence. Therefore, these text passages were forward-translated by only one translator and a second person who performed the role of the project coordinator (LK) during the course of the forward translation process.

All translators and a formal caregiver were invited to participate in a group meeting, which was scheduled for four hours. At the beginning of the group meeting, the participants were informed of the construct of challenging behaviours as well as the development and content of the CBS and its target group. Then, we introduced the

translation method and the goal criteria for translation (equivalence and acceptability of wording) ~~(30 minutes)~~. The translation proposals of the CBS were subsequently discussed ~~(210 minutes)~~. During the discussion, all potential translations were presented simultaneously on a large screen, and either one proposal was selected following a consensus or a new formulation was collectively created.

The formal caregiver was asked to propose particular formulations to adapt them to the language used by German care staff. Questions for the author of the original version of the CBS were collected. The entire process was audiotaped and subsequently protocolled.

Following the first translation panel, the coordinator generated a first interim version of the CBS-G. Issues that required clarification were submitted to the author of the original version, and the comments and explanations that we received in response were considered in the first interim version. This version was thoroughly checked, slightly revised and then forwarded electronically to all translators and the formal caregiver to be corrected or approved conclusively.

Preparation and conduction of lay panel with nursing staff

Six formal caregivers were invited to participate in the lay panel based on the sample size in Hagell, Hedlin, Meads, Nyberg, and Mc Kenna (2010). All formal caregivers were native German speakers and had cared for persons with dementia in nursing homes. The sample was further designed to include caregivers with a 3-year education in nursing/geriatric care as well as persons with no education to ascertain that the CBS-G was comprehensive independent of the education level of the formal caregivers. All caregivers were recruited through telephone calls to managers of three regional nursing homes. Several days prior to the meeting of the lay panel, the formal caregivers

received copies of the first interim CBS-G version and six key questions to discuss either alone or in collaboration with their colleagues in preparation for the meeting:

(1) Are there any items listed in the CBS that are not labelled precisely?
(2) Are there any examples that are not comprehensible?
(3) Are there any response options that are not formulated adequately?
(4) Are there any other terms or formulations that are not comprehensible or that sound incorrect?
(5) Are there any text passages in which the instructions for use of the German CBS are not formulated clearly?
(6) Which pieces of information contained in the instructions do you consider to be redundant?

The lay panel was scheduled for three hours in total. At the beginning, the participants were informed about the CBS, the translation method was introduced, and the goal criteria of the translation were explained using a short presentation, which was similar to the process of the first translation panel. The key questions that had previously been submitted to the members were again presented and explained ~~(30 minutes)~~.

Subsequently, the first interim version of the CBS included the instrument´s instructions, the description of the scoring system and the background information, which were presented successively on a large screen and were reviewed and discussed by the participants with respect to the key questions ~~(150 minutes)~~. The lay panel was facilitated by the same coordinator who had moderated the first translation panel and by one of the translators. All suggestions were collected by the coordinator, and the discussions was audiotaped and subsequently protocolled.

In the aftermath, the coordinator's task was to integrate the suggestions of the lay panel while ensuring that the structure and content of the original CBS version were preserved. The first interim CBS-G version was revised and then sent to the nursing staff of the lay panel for their final consent.

Linguistic validation

Following the translation process, the linguistic validity of the CBS-G was tested. Establishing linguistic validity is important to ensure that a translated instrument is conceptually equivalent to the original and culturally appropriate (Gawlicki, Reilly, Popielnicki, & Reilly, 2006). A pilot sample of formal caregivers reviewed the CBS-G using the approach to testing content validity described by Lynn (1986) and Polit, Beck, and Owen (2007). There are no clear recommendations regarding the number of experts to include. Lynn (1986) suggested including a minimum of 3 experts, but more than 10 was not considered helpful. Other authors recommend between 2 and 20 experts. According to Grant and Davis (1997), the number of experts depends on the required level of expertise and the diversity of knowledge.

We decided to recruit a stratified sample of at least 15 formal caregivers. For that purpose we sent a written inquiry to the managers of 10 different nursing homes in North-Rhine-Westphalia region who subsequently invited formal caregivers from their institution to participate in the linguistic validation of the CBS-G, ~~who were invited via the managers of 10 different nursing homes in the North-Rhine-Westphalia region.~~ The sample was designed to represent the variety of nursing home sponsorships (private, public-welfare, public) as well as the variety of formal caregivers' educational levels (professional training in nursing/geriatric care, social (care) worker, no education) in German nursing homes. The caregivers rated the unambiguity and familiarity of the

items and appropriateness of the examples used to describe the items of the CBS-G on a 4-point Likert scale (very unambiguous/familiar/ appropriate = 4; not at all ambiguous/unfamiliar/inappropriate = 1). Additionally, they evaluated the comprehensibility of the translated instructions and scoring system, the adequateness of the rating options, the practicability and time required to perform the CBS-G and general suitability of the CBS-G for use in nursing practice. For the review, each formal caregiver was asked to apply the CBS-G exemplarily to one resident with dementia.

Data analysis of linguistic evaluation

An item content validity index (I-CVI), which represents the proportion of caregivers (experts) who rated the instrument's content as valid (unambiguous, familiar, appropriate), was calculated for each CBS-G item (Lynn, 1986; Polit et al., 2007). For the total CBS-G, two different measures were applied to calculate the scale content validity index (S-CVI). The first method calculated the average of all the I-CVIs of the individual CBS-G items (S-CVI/AVE). The second method calculated the proportion of CBS-G items that achieved a rating of 3 or 4 by all caregivers (S-CVI/UA). The content validity indexes were rated as good when the I-CVI, S-CVI/AVE, and S-CVI/UA were $\geq$ 0.78, 0.90 and 0.80, respectively (Polit et al., 2007). To counteract the limitations of the I-CVI, a modified kappa, `k`, was calculated to correct the I-CVI for change in agreement (Polit et al., 2007; Wynd, Schmidt, & Schaefer, 2003). The probability of change in agreement was calculated based on the formulas suggested by Polit et al. (2007). Finally, the possibility of chance occurrence (Pc) was calculated to evaluate whether the value of k was excellent (more than 0.74), good (between 0.60 and 0.74) or fair (between 0.4 and 0.59) (Cicchetti & Sparrow, 1981; Fleis, Levin, & Paik, 2003). All

formulas are described in detail in Table 4. Microsoft Excel 2010 was used to perform all calculations.

Results

Translation panel with translators

All contacted translators except one agreed to prepare an independent translation proposal. One translator refused participation due to limited time resources. Two translators had similar educational backgrounds and professional experience, and they developed a proposal jointly (translators 2 and 3). All translators had previous translation experience, and the two-panel approach was new to all of the translators (Table 1).

Table 1: Characteristics of translation participants and lay panel

The three translation proposals diverged in many cases; new formulations were created by the team because often none of the proposed phrases seemed to fit completely. Table 2 provides a representative outline of some items of the CBS for which new formulations were created:

Table 2: Sample items of the CBS with new formulations

Although several clarity issues emerged when defining the items and preparing the translation session, during the translation session, some additional aspects appeared to be ambiguous for German speakers and could not be clarified by the translation team.

All ambiguous issues were summarised and proposed to be discussed with the author of the original version of the CBS.

Integrating all consented translations into one preliminary version was conducted quickly by means of the schedule and was performed by the project coordinator. The responses of the author of the original CBS scale regarding the ambiguities of the CBS were included in the first interim version of the CBS-G.

The translators' feedback on the first interim version of the German CBS revealed that a few consented translation phrases did not appear to be suitable when considering the entire scale. For example, when translated literally into German, the response options "This causes a lot of problems" (German Translation: "Es verursachte viele Probleme") and "This causes quite a lot of problems" (German Translation: "Es verursachte ziemlich viele Probleme") did not seem to differ sufficiently. Thus, the translation of the response option "This causes a lot of problems" was slightly revised to "Es verursachte sehr viele Probleme". The translation of the response option "This causes quite a lot of problems" was changed to "Es verursachte viele Probleme".

Moreover, the translators approved our proposals concerning supplementary changes and put forward additional suggestions for improvement, which were added in the CBS-G, e.g., capitalising the term for the outcomes "HÄUFIGKEIT" [Frequency] and "SCHWIERIGKEIT" [Difficulty] to allow for easier recognition.

Lay panel with nursing staff

Seven formal caregivers from three different nursing homes agreed to participate in the subsequent lay panel. Those seven persons were sent the key questions for the evaluation and discussion of the first interim German CBS version. Further information

was provided about the use of data (i.e., anonymised reporting of the data in specific publications and documentation of the translation process).

The meeting of the lay panel lasted four hours. The actual composition of the panel is outlined in Table 1. One formal caregiver who had agreed to participate did not attend the session. At the time of the meeting of the lay panel, the participating formal caregivers did not have knowledge of the original CBS version. Although we provided participants with materials to prepare for the discussions, some participants did not appear to be familiar with the interim version of the CBS-G or with the key questions. Therefore, both documents were explained in greater detail at the beginning of the meeting.

Subsequently, the lay panel discussed the different units of meaning, the instrument´s instructions, the description of the scoring system and the background information of the first interim CBS-G in relation to the six key questions. Some examples used to describe the items of the CBS-G were controversial, resulting in a few reformulations. In item 8. Unruhe [Restlessness] the example "auf und abgehen" [pacing] was deleted to ensure a clear distinction from item 7 Wandern [wandering]. In item 9 Motivationslosigkeit [Lack of motivation] the example "Teilnahmslosigkeit" [apathy] was removed to ensure a clear distinction from item 25 Teilnahmslosigkeit [lack of occupation]. In item 21 "Schlafstörungen" [sleep problems) the example "Tag-Nacht-Umkehr" [reversal of circadian rhythm] was added to clarify the item. ~~(Table 3).~~

~~Table 3: Revision of CBS-G items after lay panel~~

Some participants were not familiar with the terms "Item" or "Coping" used in the descriptions of the scoring system and the background information. Additional

information was added to explain the term “item”. The foreign word “coping” was replaced by a term “Bewältigung durch das Personal” that is closer to every language of nursing staff. Moreover, caregivers pointed to difficulties in understanding the scoring system and in interpreting the resulting scores. Consequently, a sample calculation of the CBS-G score was provided in the description of the scoring system (Figure 1). Suggestions related to changing the structure or content of the scale (e.g., an additional field for indicating whether the resident lives in a single or double room) were recorded but not adopted at that time.

Finally, the meeting of the lay panel resulted in the second interim German CBS version, which was conclusively checked once again by two translators from the translation panel, resulting in the final CBS-G, which was subsequently reviewed and evaluated regarding its linguistic validity (Figure 1).

Figure 1: CBS-German

Linguistic validation

Twenty-three formal caregivers from 7 different nursing homes (3 private, 3 public welfare, 2 public) reviewed the final CBS-G. Eighteen persons had professional education in nursing or geriatric care, 2 had an education in social care and 3 had no education (Table 3).

Table 3: Characteristics of participants in the linguistic validation

Regarding unambiguity, 22 of the 25 items showed excellent content validity (I-CVI ≥ 0.78 and k > 0.74), two items (11 and 14) showed good content validity (I-CVI 0.7.-0.73, k ≥ 0.69) and item 6 showed low content validity (I-CVI 0.33, k = 0.29) (Table 4). Regarding familiarity, 23 items showed excellent content validity (I-CVI ≥ 0.78 and k > 0.74), item 14 showed fair content validity (I-CVI = 0.61, k = 0.57) and item 6 showed low content validity (I-CVI 0.26, k = 0.25) (Table 4).

Table 4: Content validity of the CBS-G

Eighteen items were reviewed to determine the adequateness of the examples. All items except one (item 13) showed excellent validity indexes (I- CVI ≥ 0.78 and k > 0.74) (Table 4).

The average scale content validity (S-CVI/AVE) was 0.85 for the unambiguity, 0.83 for the familiarity and 0.89 for the adequateness of the examples, all of which are below the cut-off of 0.90. The scale content validity universal agreement (S-CVI/UA) was low for all indexes (≤ 0.16), indicating that the unanimity among caregivers was low (Table 4).

The majority of the formal caregivers (95%) reported that the instructions of the CBS-G were (very) understandable ~~(n = 21)~~. Fifteen of the 22 persons also rated the scoring system as (very) understandable. The adequateness of the rating options was confirmed by 19 formal caregivers, and the average time to complete the CBS-G was 21.09 (± 13.45) minutes. Seventeen persons also reported that the CBS-G was practicable and suitable (Table 4).

Table 4: Review of the CBS-G

Discussion

The purpose of this study was to translate the English CBS into German according to the two-panel approach and test its linguistic validity in a pilot sample of formal caregivers. ~~The results of the linguistic validity indicated that twenty-two of the 25 items of the CBS-G showed excellent CVI and k regarding the unambiguity of the items. Twenty-three items showed excellent CVI and k values in relation to the familiarity of the items. All examples used to describe the items were viewed as adequate.~~

Although the two-panel approach is seldom used in research to translate health-related outcome measures and guidelines primarily propose a forward/backwards-translation (Guillemin, Bombardier, & Beaton, 1993; Wild et al., 2005), in this study, the two-panel approach proved to be a useful alternative for the translation of the CBS, which is not solely due to the convincing results of the linguistic validation of the CBS-G. A major strength of this approach is the possibility to involve and amalgamate various qualifications and experiences by means of different expert panels and, thus, not rely solely on the abilities of one or two persons (Swaine-Verdier et al., 2004). The strength of the translation panel was its linguistic and clinical expertise, whereas the strength of the lay panel was that it represents the target population of the CBS-G, both of which important for the validation of the translation process (Danielsen, Pommergaard, Burcharth, Angenete, & Rosenberg, 2015). The diversity of persons within the translation panel often led to intensive and substantial discussions in our study, and new translations were generated through teamwork. Thus, our study does not share the problems arising from members´common views or perceived pressure to reach a consensus that were noted by Hilton and Skrutkowski (2002).

~~The lay panel with the formal caregivers proved to be a useful addition to the translation panel, because it provided further cues for reformulation and for the difficulties in understanding the scoring system, which was later confirmed by some participants of the linguistic validation of the CBS-G. We believe that difficulties in understanding the scoring systems mainly resulted due to the complexity of the scoring system and the interpretation of its results rather than due to low translation quality. Therefore, we recommend adding information to the instructions of the CBS-G that facilitates the use and interpretation of the scoring system.~~

The two-panel approach is a time-consuming method because it takes time to recruit the two different expert panels and to merge the different translation proposals that result during the different phases of the two-panel approach. Thus, the whole translation process lasted nine months.

Working together with the author of the original version of the instrument proved to be a clear advantage as it contributed to the clarification of the exact meaning of certain terms and is recommended for further translations.

The findings of the validation study mostly support the linguistic validity of the CBS-G. Experts judged all items except two (item 6: perseveration; item 14: manipulative) as unambiguous and familiar and all examples used to characterise the items as adequate. The results of the two measurements, I-CVI and k, were consistent. Items meeting the I-CVI criterion of 0.78 also showed excellent k* values and vice versa, indicating that both methods resulted in the same conclusion and strengthening the current evidence. Item 6 was viewed as ambiguous and unfamiliar, and item 14 was primarily unfamiliar. Therefore, both items should be replaced by terms that are more common in the everyday language of German professional caregivers.

The S-CVI/AVE and the S-CVI/UA found in this study did not reach the limits of agreement recommended by Polit et al. (2007). The low S-CVIUA values of 0.04, 0.12 and 0.16 might be explained by the large number of experts included, as the likelihood of achieving total agreement decreases as the number of experts increases (Polit & Beck, 2006). The time required to complete the CBS-G was, on average, 21 minutes, which is more than twice as long as that specified in the original article on the CBS (Moniz-Cook et al., 2001). This result may be due to the fact that the CBS-G was being used for the first time.

Future directions

A German version of the CBS is now available for further evaluation of its psychometric properties, as the translation and linguistic validation of an instrument represents only the first step in the process of adapting an instrument from one country to another (Acquadro, Conway, Hareendran, & Aaronson, 2008). Further research is needed to account for the reliability and validity of the CBS-G, as these metrics are important for the evaluation of the effect of psychosocial interventions in the management of challenging behaviours of individuals with dementia.

References

Acquadro, C., Conway, K., Hareendran, A., & Aaronson, N. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health, 11*, 509-521. doi:10.1111/j.1524-4733.2007.00292.x

Barandun Schäfer, U., Massarotto, P., Lehmann, A., Wehrmuller, C., Spirig, R., & Marsch, S. (2009). [Translation of the Richmond Agitation-Sedation Scale (RASS) for use in German Swiss intensive care units]. *Pflege, 22*, 7-17. doi:10.1024/1012-5302.22.1.7

Bird, M., & Moniz Cook, E. D. (2008). Challenging behavior in dementia: a psychosocial approach to intervention. In B. Woods & L. Clare (Eds.), *Handbook of the clinical psychology of ageing* (Vol. 2, pp. 571-594). West Sussex: John Wiley & Sons, Ltd.

Brodaty, H., & Arasaratnam, C. (2012). Meta-analysis of nonpharmacological interventions for neuropsychiatric symptoms of dementia. *American Journal of Psychiatry, 169*, 946-953. doi:10.1176/appi.ajp.2012.11101529

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*, 127-137.

Cohen-Mansfield, J., Golander, H., Ben-Israel, J., & Garfinkel, D. (2011). The meanings of delusions in dementia: a preliminary study. *Psychiatry Research, 189*, 97-104. doi:10.1016/j.psychres.2011.05.022

Connor, D. J., Sabbagh, M. N., & Cummings, J. L. (2008). Comment on administration and scoring of the Neuropsychiatric Inventory in clinical trials. *Alzheimers & Dementia, 4*, 390-394. doi:10.1016/j.jalz.2008.09.002

Danielsen, A. K., Pommergaard, H. C., Burcharth, J., Angenete, E., & Rosenberg, J. (2015). Translation of questionnaires measuring health related quality of life is not standardized: a literature based research study. *PLoS One, 10*, e0127050. doi:10.1371/journal.pone.0127050

Duffy, F. (2016). `Look at all of me`-a CLEAR model for dementia care. *Journal of Dementia Care, 22*, 27-30.

Eberl, I., & Bartholomeyczik, S. (2010). [The Belgian Nursing Minimum Data Set II (B-NMDS II) and its transfer to German hospitals: results of the first investigation phase, the translation and adaption process of the instrument]. *Pflege, 23*, 309-319. doi:10.1024/1012-5302/a000064

Edvardsson, D., Sandman, P. O., Nay, R., & Karlsson, S. (2008). Associations between the working characteristics of nursing staff and the prevalence of behavioral symptoms in people with dementia in residential care. *International Psychogeriatrics, 20*, 764-776. doi:10.1017/S1041610208006716

Fleis, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (Vol. 3). New York, NY: John Wiley.

Gawlicki, M. C., Reilly, M. C., Popielnicki, A., & Reilly, K. (2006). Linguistic validation of the US Spanish work productivity and activity impairment questionnaire, general health version. *Value in Health, 9*, 199-204. doi:10.1111/j.1524-4733.2006.00101.x

Gitlin, L. N., Marx, K. A., Stanley, I. H., Hansen, B. R., & Van Haitsma, K. S. (2014). Assessing neuropsychiatric symptoms in people with dementia: a systematic review of measures. *International Psychogeriatrics, 26*, 1805-1848. doi:10.1017/S1041610214001537

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health, 20,* 269-274. doi:10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46,* 1417-1432. doi:10.1016/0895-4356(93)90142-N

Hagell, P., Hedlin, P.-J., Meads, D. M., Nyberg, L., & Mc Kenna, S. P. (2010). Effects of method of translation fo patient-reported health outcome questionnaires: a randomized study of the translation of the Rheumatoid Arthritis Quality of Life (RAQoL) instrument for Sweden. *Value in Health, 13,* 424-430. doi:10.1111/j.1524-4733.2009.00677.x

Hilton, A., & Skrutkowski, M. (2002). Translating instruments into other languages: development and testing processes. *Cancer Nursing, 25,* 1-7.

Jeon, Y. H., Sansoni, J., Low, L. F., Chenoweth, L., Zapart, S., Sansoni, E., & Marosszeky, N. (2011). Recommended measures for the assessment of behavioral disturbances associated with dementia. *American Journal of Geriatric Psychiatry, 19,* 403-415. doi:10.1097/JGP.0b013e3181ef7a0d

Koder, D., Hunt, G. E., & Davison, T. (2014). Staff's views on managing symptoms of dementia in nursing home residents. *Nursing Older People, 26,* 31-36. doi:10.7748/nop.26.10.31.e638

Lai, C. K. (2014). The merits and problems of Neuropsychiatric Inventory as an assessment tool in people with dementia and other neurological disorders. *Clinical Interventions in Aging, 9,* 1051-1061. doi:10.2147/CIA.S63504

Lam, C. L., Chan, W. C., Mok, C. C., Li, S. W., & Lam, L. C. (2006). Validation of the Chinese Challenging Behaviour Scale: clinical correlates of challenging behaviours in nursing home residents with dementia. *International Journal of Geriatric Psychiatry, 21*, 792-799. doi:10.1002/gps.1564

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*, 382-385.

Moniz-Cook, E., Woods, R., Gardiner, E., Silver, M., & Agar, S. (2001). The Challenging Behaviour Scale (CBS): development of a scale for staff caring for older people in residential and nursing home. *British Journal of Clinical Psychology, 40*, 309-322. doi:10.1348/014466501163715

Moniz Cook, E. D., Swift, K., James, I., Malouf, R., De Vugt, M., & Verhey, F. (2012). Functional analysis-based interventions for challenging behaviour in dementia. *Cochrane Database of Systematic Reviews, 2*, CD006929. doi:10.1002/14651858.CD006929.pub2

National Institute for Health and Care Excellence [NICE], & Social Care Institute for Excellence [SCIE]. (2006). NICE clinical guideline 42: dementia supporting people with dementia and their carers in health and social care. Retrieved from https://www.nice.org.uk/guidance/cg42

Orrell, M., Hancock, G., Hoe, J., Woods, B., Livingston, G., & Challis, D. (2007). A cluster randomised controlled trial to reduce the unmet needs of people with dementia living in residential care. *International Journal of Geriatric Psychiatry, 22*, 1127-1134. doi:10.1002/gps.1801

Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*, 489-497. doi:10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*, 459-467. doi:10.1002/nur.20199

Schmidt, S. G., Dichter, M. N., Palm, R., & Hasselhorn, H. M. (2012). Distress experienced by nurses in response to the challenging behaviour of residents-evidence from German nursing homes. *Journal of Clinical Nursing, 21*, 3134-3142. doi:10.1111/jocn.12066

Selbaek, G., Engedal, K., & Bergh, S. (2013). The prevalence and course of neuropsychiatric symptoms in nursing home patients with dementia: a systematic review. *Journal of the American Medical Directors Association, 14*, 161-169. doi:10.1016/j.jamda.2012.09.027

Swaine-Verdier, A., Doward, L. C., Hagell, P., Thorsen, H., & Mc Kenna, S. (2004). Adapting quality of life instruments. *Value in Health, 7*, S27-S30. doi:10.1111/j.1524-4733.2004.7s107.x

van der Linde, R. M., Stephan, B. C., Dening, T., & Brayne, C. (2014). Instruments to measure behavioural and psychological symptoms of dementia. *International Journal of Methods in Psychiatric Research, 23*, 69-98. doi:10.1002/mpr.1414

Vasse, E., Vernooij-Dassen, M., Cantegreil, I., Franco, M., Dorenlot, P., Woods, B., & Moniz-Cook, E. (2012). Guidelines for psychosocial interventions in dementia care: a European survey and comparison. *International Journal of Geriatric Psychiatry*, *27*, 40-48. doi:10.1002/gps.2687

Warwick, R., Higgon, J., & Edgar, F. (2011). IAPT and people with dementia in residential care-a training programme for care home staff. *PSIGE Newsletter, 113*, 63-67.

Wenborn, J., Challis, D., Head, J., Miranda-Castillo, C., Popham, C., Thakur, R., . . . Orrell, M. (2013). Providing activity for people with dementia in care homes: a cluster randomised controlled trial. *International Journal of Geriatric Psychiatry, 28*, 1296-1304. doi:10.1002/gps.3960

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value in Health, 8*, 94-104. doi:10.1111/j.1524-4733.2005.04054.x

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research, 25*, 508-518.

Tables:

Figure:

Figure 1

SKALA ZUR ERFASSUNG VON HERAUSFORDERNDEM VERHALTEN ÄLTERER MENSCHEN IN PFLEGEEINRICHTUNGEN (CBS-GERMAN)

Name:

Alter: **Geschlecht:** ☐**Mann** ☐Frau **Demenzdiagnose vorhanden?** ☐ **ja** ☐ **nein** ☐**unbekannt**

Pflegeeinrichtung: **Wohnbereich:**

Skala ausgefüllt von: **Datum:**

Körperliche Fähigkeiten (Unzutreffendes durchstreichen)

1. **kann ohne Hilfe gehen/ kann mit einem Hilfsmittel gehen/ sitzt im Rollstuhl**
2. **kontinent/ harninkontinent/ stuhlinkontinent/ harn- und stuhlinkontinent**
3. **kann ohne Hilfe aufstehen/ sich hinsetzen/ hinlegen (Bett/Stuhl)/ benötigt Hilfe beim Aufstehen/ Hinsetzen/ Hinlegen (Bett/Stuhl)**
4. **kann sich ohne Hilfe waschen und anziehen/ benötigt Hilfe beim Waschen und Anziehen**
5. **kann ohne Hilfe essen und trinken/ benötigt Hilfe beim Essen und Trinken**

Auf der nächsten Seite finden Sie eine Liste von herausfordernden Verhaltensweisen, die ältere Menschen in Pflegeeinrichtungen zeigen können. Denken Sie an das Verhalten der Person in den letzten 8 Wochen zurück und bewerten Sie es wie folgt:

KOMMT VOR: **ja/nein. Falls ja, fahren Sie mit HÄUFIGKEIT fort.**

HÄUFIGKEIT:

4: Diese Person zeigt dieses Verhalten **täglich**

3: Diese Person zeigt dieses Verhalten **mehrmals pro Woche**

2: Diese Person zeigt dieses Verhalten **mehrmals im Monat**

1: Diese Person zeigt dieses Verhalten **gelegentlich**

SCHWIERIGKEIT:

Dann bewerten Sie anhand der Skala, wie schwierig es ist, mit *jeder zutreffenden Verhaltensweise* umzugehen.

4: Es verursacht sehr viele Probleme

3: Es verursacht viele Probleme

2: Es verursacht einige Probleme

1. Es ist kein Problem

Anm.: Falls ein Verhalten nicht zutrifft, entfallen Angaben zu HÄUFIGKEIT/ SCHWIERIGKEIT. Falls ein Verhalten für Sie in den letzten (acht) Wochen unterschiedlich schwierig war, kreuzen Sie den für Sie höchsten Schwierigkeitsgrad an.

	HERAUSFORDERNDES VERHALTEN	KOMMT VOR		HÄUFIGKEIT		SCHWIERIGKEIT		HERAUSFORDERUNG	
		ja	nein	1 = gelegentlich 2 = mehrmals/Monat 3 = mehrmals/Woche 4 = täglich		1 = kein Problem 2 = einige Probleme 3 = viele Probleme 4 = sehr viele Probleme		HÄUFIGKEIT x SCHWIERIGKEIT	
1.	**Körperliche Aggression** (z.B. schlagen, treten, kratzen, zugreifen)								
2.	**Verbale Aggression** (z.B. beleidigen, fluchen, bedrohen)								
3.	**Selbstverletzung** (z.B. sich selbst schneiden/schlagen, Nahrung verweigern)								
4.	**Rufen**								
5.	**Schreien/ Lautstarkes Jammern**								
6.	**Perseveration** (z.B. ständiges Wiederholen von Wörtern/ Sätzen/ Fragen/ Handlungen oder Singen)								
7.	**Wandern** (scheinbar zielloses Umhergehen)								
8.	**Unruhe** (z.B. herumzappeln, nicht zur Ruhe kommen können, auf- und abgehen, immer ‚auf dem Sprung sein')								
9.	**Motivationslosigkeit** (z.B. schwer zu motivieren, fehlendes Interesse an Aktivitäten, Apathie/ Teilnahmslosigkeit)								
10.	**Hinterherlaufen/ Folgen** (z.B. anderen Bewohnern/Personal nachlaufen, ‚klammern')								
11.	**Andere Menschen stören/ sich einmischen**								
12.	**Sammeln/ Horten** (z.B. persönliche Dinge, Abfall, Papier, Essen)								
13.	**Misstrauen** (z.B. andere beschuldigen)								
14.	**Manipulativ** (z.B. andere Bewohner/Personal zum eigenen Vorteil nutzen)								
15.	**Selbstpflegedefizit** (z.B. mangelnde/ fehlende Körperpflege)								
16.	**Spucken**								
17.	**Mit Kot schmieren**								
18.	**Urinieren an unpassenden Orten/ in unpassenden Situationen** (z.B. nicht in der Toilette, in der Öffentlichkeit)								
19.	**Unpassendes Entkleiden** (z.B. unangemessenes Ausziehen, Entblößen nackter Haut)								
20.	**Unangemessenes sexuelles Verhalten** (z.B. in der Öffentlichkeit masturbieren, unangemessene ‚Annäherungsversuche' machen)								
21.	**Schlafstörungen** (z.B. nachts aufwachen, Schlaflosigkeit)								
22.	**Fehlende Kooperationsbereitschaft** (z.B. ausdrücklich Anleitungen/ Aufforderungen des Personals ignorieren, Nahrung verweigern, Hilfe bei der Selbstpflege ablehnen)								
23.	**Gefährliches Verhalten** (z.B. Brände oder Überschwemmungen verursachen)								
24.	**Aufmerksamkeit einfordern**								
25.	**Teilnahmslosigkeit** (z.B. herumsitzen und nichts tun)								
	GESAMTPUNKTZAHL Addieren Sie die Punkte (1-25) aus jeder Spalte.		**25**		**100**		**100**		**400**

Table 1: Characteristics of translation participants and lay panel

Participants of translation panel							Participants of lay panel						
Translator	Translator 1 (German) native)	Translator 2 (Polish native)	Translator 3 (German native)	Translator 4 (English native)	Practitioner (German) native)	Coordinator (German) native)	Participants	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Professional education/ studies	Communication Sciences, (B.A.) Language and communication studies (M.A.) Health Sciences (Ph.D.)	Professional training in geriatric care (3 years) Nursing Sciences (B.A. and M.A.) Nursing Sciences (Ph.D.)	Professional 3-years training in nursing, Nursing Sciences (Ph.D.)	Not specified	Professional training in geriatric care (3 years) Education in geronto-psychiatric care	Diploma in psychology + gerontology	Professional education	Professional training in geriatric care (3 years)	Professional training in geriatric care (3 years)	(Ongoing) Professional training in geriatric care (3 years)	Professional training in geriatric care (1 years)	No education	No education
Working position	Scientist	Scientist	Scientist	Assistant in research	Elderly care nurse	Scientist	Working position	Elderly care nurse	Elderly care nurse	Trainee	Nursing aid	Nursing aid	Nursing aid
English Language skills	Lived 4 years in New Zealand; Experienced in writing English papers, giving English presentation	Experienced in writing English papers, giving English presentation	Experienced in writing English papers, giving English presentation		/	Experienced in writing English papers, giving English presentation	Working area	Nursing unit for people with and without dementia	Nursing unit for people with and without dementia	Nursing unit for people with and without dementia	Nursing unit for people with and without dementia	Nursing unit for people with and without dementia	Nursing unit for people with and without dementia
Professional experiences with challenging behaviour of	Yes (research)	Yes (nursing practice and research)	Yes (nursing practice and research)	Yes (research)	Yes (nursing practice)	Yes (nursing practice and research)	Working hours (per week)	39	39	39	38,5	39,5	38,5
Professional experiences in dementia	Yes (research)	Yes (nursing practice and research)	Yes (nursing practice and research)	Yes (research)	Yes (nursing practice)	Yes (nursing practice, research, clinical diagnostics)	Work experience (years)	5	6	-	25	20	20

Table 2: Sample items of the CBS with new formulations

Item	Original CBS	1st proposal	2nd proposal	3rd proposal	New formulation
10	Clinging (follows/ holds on to other residents/ staff, etc.)	Anklammern (folgt anderen Bewohnern oder Personal oder hält sich an ihnen fest usw.)	Klammern (er/ sie folgt/ haftet sich an andere Bewohner/ Personal, etc.)	Anhänglichkeit (er/ sie folgt anderen Bewohnern/ Mitarbeitern bzw. klammert sich an sie etc.)	Folgen/Hinterherlaufen: z.B. anderen Bewohnern/Personal nachlaufen, "klammern"
15	Lack of Self Care (hygiene problems, dishevelled, etc.)	Mangel/ Fehlen von Selbstfürsorge (Probleme mit der persönlichen Hygiene, ist unordentlich/ zerzaust)	Fehlende Selbstfürsorge (Hygieneprobleme, ungepflegt, etc.)	Fehlende Körperpflege (hygienische Probleme, ungepflegt, etc.)	Selbstpflegedefizit (z.B. mangelnde/fehlende Körperpflege)
18	Inappropriate Urinating (in public, not in toilet, etc.)	Unangemessenes/ unpassendes Harnlassen/Urinieren (in der Öffentlichkeit, nicht auf der Toilette usw.)	Unangemessenes Urinieren (in der Öffentlichkeit, nicht in die Toilette, etc.)	Unangemessenes Urinieren (in der Öffentlichkeit, nicht in einer Toilette, etc.)	Urinieren an unpassenden Orten/ in unpassenden Situationen (z.B. in der Öffentlichkeit, nicht in der Toilette)
19	Stripping	Entkleiden (zieht	Entkleiden (er/ sie	Entkleiden (er/ sie	Unpassendes

	(removes clothes inappropriately, flashes, etc.)	sich unpassender-weise aus, ist ein Flitzer usw.)	entkleidet sich unangemessen, entblößt sich, etc.)	zieht unangebracht Kleidungsstücke aus, entblößt nackte Haut, etc.)	Entkleiden: z.B. unangemessenes Ausziehen, Entblößen nackter Haut
22	Non-compliance (deliberately ignores staff requests, refuses food, resists self-care help, etc.)	Verweigerung (Bitten vom Personal vorsätzlich/ absichtlich ignorieren, Essen verweigern, Hilfe bei der Selbstfürsorge widerstehen/widersetzen usw.)	Fehlende Compliance (absichtlich die Wünsche des Personals ignorieren, Nahrung verweigern, Hilfe zur Selbstpflege abwehren, etc.)	Fehlende Compliance (er/ sie ignoriert bewusst Mitarbeiteraufforderungen, verweigert die Nahrungsaufnahme, lehnt Hilfe bei der Körperpflege ab, etc.)	Fehlende Kooperationsbereitschaft (z.B. ausdrücklich Anleitungen/Aufforderungen des Personals ignorieren, Nahrung verweigern, Hilfe bei der Selbstpflege ablehnen)
25	Lack of Occupation (sits around doing nothing, etc.)	Mangel an Beschäftigung/ Beschäftigungslosigkeit (sitzt herum ohne etwas zu tun/machen usw.)	Mangel an Beschäftigung (sitzt herum, ohne etwas zu tun, etc.)	Fehlen von Beschäftigung (er/ sie sitzt herum und tut nichts, etc.)	Teilnahmslosigkeit (z.B. herumsitzen und nichts tun)

Table 3

Table 3: Characteristics of participants in the linguistic validation

Characteristics	Number (n = 23)
Sex (female)	17 (73.91%)
Professional education	
Professional training in nursing (3 years)	7
Professional training in geriatric care (3 years)	11
Social care worker (non-academic)	1
Social worker (academic)	1
No education	3
Working position	
Nursing management	3
Head nurse	4
Elderly care nurse	11
Nursing aide	4
Social worker	1
Working area	
Nursing unit for people with/without dementia	9
Nursing unit for people with dementia	4
Small-scale living unit for people with dementia	3
Day care for people with dementia	3
Gerontopsychiatry	1
Management	3
	Mean (min-max)
Working hours (per week)	32.11 (5-42)
Work experience (years)	14.25(1-35)
Age (years)	42.96 (25-64)

Table 4: Content validity of the CBS-G

		Unambiguity of items						Familiarity of items						Adequateness of examples					
No	Items	number of experts	Number giving a rating of 3 or 4	I-CVI[a]	p_c[b]	*k**[c]	evaluation[d]	number of experts	Number giving a rating of 3 or 4	I-CVI a	pc[b]	k*[c]	evaluation[d]	number of experts	Number giving a rating of 3 or 4	I-CVI[a]	p_c[b]	*k**[c]	evaluation[d]
1	Physical Aggression (hits, kicks, scratches, grabs, etc.)	23	22	0.97	0.000	0.96	****	23	21	0.91	0.000	0.91	****	22	19	0.86	0.000	0.86	****
2	Verbal Aggression (insults, swearing, threats, etc.)	23	19	0.83	0.001	0.83	****	23	20	0.87	0.000	0.87	****	23	20	0.87	0.000	0.87	****
3	Self-Harm (cuts/hits self, refuses food/starves self, etc.)	23	22	0.96	0.000	0.96	****	23	20	0.87	0.000	0.87	****	23	20	0.87	0.000	0.87	****
4	Shouting	23	19	0.83	0.001	0.83	****	23	19	0.83	0.001	0.83	****						
5	Screaming/Crying out	22	20	0.91	0.000	0.91	****	22	20	0.91	0.000	0.91	****						
6	Perseveration (constantly repeating speech or actions, repetitive questioning or singing)	21	7	0.33	0.055	0.29	*	23	6	0.26	0.012	0.25	*	18	16	0.89	0.001	0.89	****
7	Wandering (walks aimlessly around home)	23	20	0.87	0.000	0.87	****	23	22	0.97	0.000	0.96	****	23	20	0.87	0.000	0.87	****
8	Restlessness (fidgets, unable to settled down, pacing, ´on the go´, etc.)	23	18	0.78	0.004	0.78	****	23	21	0.91	0.000	0.91	****	23	20	0.87	0.000	0.87	****
9	Lack of motivation (difficult to engage, shows no interest in activities, apathy, etc.)	23	20	0.87	0.000	0.87	****	23	21	0.91	0.000	0.91	****	23	22	0.96	0.000	0.96	****

10	Clinging (follows/holds on to other residents/staff, etc.)	23	19	0.83	0.001	0.83	****	23	21	0.91	0.000	0.91	****	23	22	0.96	0.000	0.96	****
11	Interfering with other people	22	16	0.73	0.018	0.72	***	22	18	0.82	0.002	0.82	****						
12	Pilfering or Hoarding (possessions, rubbish, paper, food, etc.)	22	20	0.91	0.000	0.91	****	22	20	0.91	0.000	0.91	****	22	21	0.95	0.000	0.95	****
13	Suspiciousness (accusing others, etc.)	22	20	0.91	0.000	0.91	****	23	19	0.83	0.001	0.83	****	23	15	0.65	0.058	0.63	***
14	Manipulative (takes advantage of others, staff, etc.)	23	16	0.70	0.029	0.69	***	23	14	0.61	0.097	0.57	**	21	17	0.81	0.003	0.81	****
15	Lack of self care (hygiene problems, dishevelled, etc.)	23	19	0.83	0.001	0.83	****	23	19	0.83	0.001	0.83	****	23	18	0.78	0.004	0.78	****
16	Spitting	23	22	0.96	0.000	0.96	****	23	20	0.87	0.000	0.87	****						
17	Faecal Smearing	23	22	0.96	0.000	0.96	****	23	22	0.96	0.000	0.96	****						
18	Inappropriate Urinating (in public, not in toilet, etc.)	23	22	0.96	0.000	0.96	****	23	23	1	0.000	1	****	22	22	1	0.000	1	****
19	Stripping (removes clothes inappropriately, flashes, etc.)	23	19	0.83	0.001	0.83	****	23	19	0.83	0.001	0.83	****						
20	Inappropriate Sexual Behaviour (masturbates in public, makes inappropriate ´advances´ to others, etc.)	23	20	0.87	0.000	0.87	****	23	20	0.87	0.000	0.87	****						
21	Sleep Problems (waking in night, insomnia, etc.)	23	22	0.96	0.000	0.96	****	23	23	1	0.000	1	****						
22	Non-compliance (deliberately ignores staff requests, refuses food, resists self-care help, etc.)	23	21	0.91	0.000	0.91	****	23	21	0.91	0.000	0.91	****						
23	Dangerous Behaviour (causes fires or floods, etc.)	23	18	0.78	0.004	0.78	****	23	18	0.78	0.004	0.78	****						
24	Demands Attention	22	19	0.86	0.000	0.86	****	22	21	0.95	0.000	0.95	****						
25	Lack of Occupation (sits around doing nothing, etc.)	23	23	1	0.000	**1**	****	23	23	1	0.000	1	****						
		S-CVI/AVE 0.85						S-CVI/AVE 0.83						S-CVI/AVE 0.89					

		S-CVI/UA 0.04				S-CVI/UA 0.12				S-CVI/UA 0.16			

[a] I-CVI (content validity index) = number of experts giving rating of 3 or 4/number of experts

[b] p_c (probability of chance occurrence) = [N!/A!(N-A)!] x 0.5^N , N = number of experts; A = number of experts agreeing on rating 3 or 4

[c] k^* (modified kappa) = (I-CVI-p_c)(1-p_c)

[d] Evaluation criteria for level of content validity: relationship between I-CVI and k^*; excellent validity = CVI-I $\geq$ 0.78 and k* >0.74 (****); good validity CVI-I $<$ 0.78 and 0.60 $\leq k^* \leq$0.74 (***); fair validity CVI-I $<$ 0.78 and 0.40 $\leq k^* \leq$0.59 (**); poor validity CVI-I $<$ 0.78 and $k^* <$0.40 (*)

Table 5: Review of the CBS-G

	Total number of experts	Number of experts rating 1[a]	Number of experts rating 2	Number of experts rating 3	Number of experts rating 4[b]
Understandability of instructions	22 (23)	-	1	6	15
Understandability of scoring system	22 (23)	4	3	7	8
Adequateness of rating options	23 (23)	1	3	9	10
Practicability of CBS-G	22 (23)	1	3	11	7
Suitability of CBS-G	23 (23)	1	5	9	8

Key:

[a] 1 = not at all understandable; not at all adequate; very poor

[b] 4 = very understandable; very adequate; very good