Discover

Research

A multi-step linguistic validation for cultural adaptation of the German-language Postpartum Bonding Questionnaire

Pia-Cecilia Steinbrueck[1,2] · Gabriele Meyer[1] · Gertrud Ayerle[1]

Abstract

The four-dimensional Postpartum Bonding Questionnaire (PBQ) is an original English self-report questionnaire aimed for detection of postpartum mother–infant bonding disorders. The aim of this study was to create a German version of the PBQ that corresponds conceptually and linguistically to the original questionnaire. The systematic multistep translation process followed the guidelines of the International Society for Pharmacoeconomics and Outcomes Research, consisting of forward and backward translations, pilot testing, and cognitive interviews of mothers and a panel of experts. Cognitive interviews with users (n = 10) resulted in 15 changes that improved the completeness and accuracy of the translation. The translation was reviewed by a panel of four experts for practicality, resulting in two changes. As a result of the translation process and cognitive interviews, the German PBQ is linguistically equivalent to the original English version and can be evaluated in psychometric validation for further use in primary care.

1 Background

In recent decades, researchers in perinatal psychology have developed objective methods to study mother–infant interaction and assess infant emotion regulation [1]. This has included attempts to operationalize the concept of attachment for use in research and practice, either through observation of the mother's behavior or through the mother's own representation of feelings [2]. Several self-report questionnaires have been developed to measure mother–infant attachment [3] or to screen for attachment disorders [2, 4]. These include the Maternal Postnatal Attachment Scale [5], the Mother–Child Attachment Scale (MIBS) [3], and the Postnatal Bonding Questionnaire (PBQ) [4].

The PBQ and its modified versions provided sufficient evidence of structural validity, internal consistency, and reliability with high quality evidence in several studies [4, 6] and is widely used in both research and clinical practice [7–13]. The PBQ is also the most frequently translated questionnaire, indicating its relevance and popularity [6]. The four-dimensional Postpartum Bonding Questionnaire (PBQ) developed by Brockington and his team [4] is a reliable, validated instrument with 25 items and is used in several countries to detect postpartum mother–infant attachment disorders [13–19].

A previously validated instrument does not necessarily mean that it is valid in a different time, culture, or context [20]. It is important to note that self-report scales are potentially prone to bias due to several (culturally determined) factors,

✉ Pia-Cecilia Steinbrueck, Pia-cecilia.steinbrueck@uniklinik-freiburg.de; Gabriele Meyer, gabriele.meyer@uk-halle.de; Gertrud Ayerle, gertrud.ayerle@uk-halle.de | [1]Medical Faculty, Institute for Health and Nursing Sciences, Martin Luther University Halle-Wittenberg, Halle, Germany. [2]Medical Faculty, Institute for Nursing Science, Albert-Ludwigs-University, Freiburg, Germany.

Check for updates

including social desirability, discrimination of response options, and response style [21]. Foreign language measurement instruments therefore require complex and structured translation processes before they can be translated into other languages [22, 23]. The term "cultural adaptation" describes a process that considers both linguistic and cultural adaptation issues while attempting to establish content equivalence between source and target [22]. Accordingly, cross-cultural adaptation attempts to ensure content consistency and face validity between the source and target versions of a questionnaire [22].

There is no general agreement on how to adapt an instrument for use in a different cultural setting [20, 24]. However, there is agreement that it is inappropriate to simply translate a questionnaire and use it in a different linguistic context [20, 25]. Many instruments are adopted from Anglo-American into German. This often leads to a number of different translations of the same instrument, which, however, prioritize different items in comparison, are no longer accurate, and have partially lost the intent of the original [23]. Therefore, translations need to be standardized to ensure comparability and to prevent the loss of the original intentions, as these are the basis for a valid instrument and reliable data collection [20, 23].

For the assessment of mother–infant attachment, no questionnaire was available in German language until 2006. The German translation of the PBQ was developed using the back-translation procedure of Reck et al. [12]. The questionnaire was translated into German and the accuracy of the translation was verified by back-translation into English. The quality of this translation was confirmed by a back-translation into English performed by a native English speaker [12]. However, no German-language PBQ currently exists that has been linguistically and culturally adapted in a structured manner and to date, the translated version by Reck et al. [12] has not been used in clinical practice by midwives or pediatricians.

The process of cultural adaptation proceeds in two successive steps: linguistic as well as psychometric validation [26]. The aim of this study is the linguistic validation of the PBQ, to perform a structured translation that is conceptually equivalent to the original questionnaire and thus to create the prerequisites for reliably using a German-language validated PBQ in clinical and out-of-hospital research and practice in German-speaking countries.

2 Methods

A clear distinction should be made between translation, adaptation, and cultural validation. Translation is the single process of producing a document from a source version in the target language [24]. Adaptation refers to the process of accounting for differences between the source and target cultures to maintain equivalence of meaning [24]. In this "harmonization" process, phrases based on idiomatic and cultural differences between the source and target languages are to be identified and replaced with commonly understood phrases. This adaptation is called cultural adaptation.

Cultural validation of a questionnaire, on the other hand, is different from cultural adaptation. Cultural validation aims to ensure that the new questionnaire works as intended and has the same characteristics as the original [24]. Thus, the translation, adaptation, and validation of a questionnaire occurred in different steps, but can be part of an iterative process, as the adapted version must be changed if the questionnaire is not valid.

This study performs a structured translation process of the PBQ into German according to the guidelines of the Translation and Cultural Adaptation Group. The Task Force on Translation and Cultural Adaptation, founded by ISPOR (International Society for Pharmacoeconomics and Outcomes Research), developed guidelines and principles for the translation and cultural adaptation of so-called "Patient-Reported Outcome Measures" in 2005 [26].

ISPOR suggests the following procedure: Preparation, forward translation and alignment, back translation, back translation review, harmonization, cognitive debriefing, cognitive debriefing review, proofreading, and finalization [26]. Based on the procedure proposed by ISPOR and other suggestions from the literature on instrument adaptation [22], the translation procedure carried out here is thereby divided into the following seven Steps as shown in Fig. 1:

2.1 Forward translation

The first step involved forward translation of the English original into the German target language. This was done by two non-professional translators working independently of each other, both of whom are native speakers of German and have a command of the source language (English) C1/C2 level (translation versions V1 and V2).
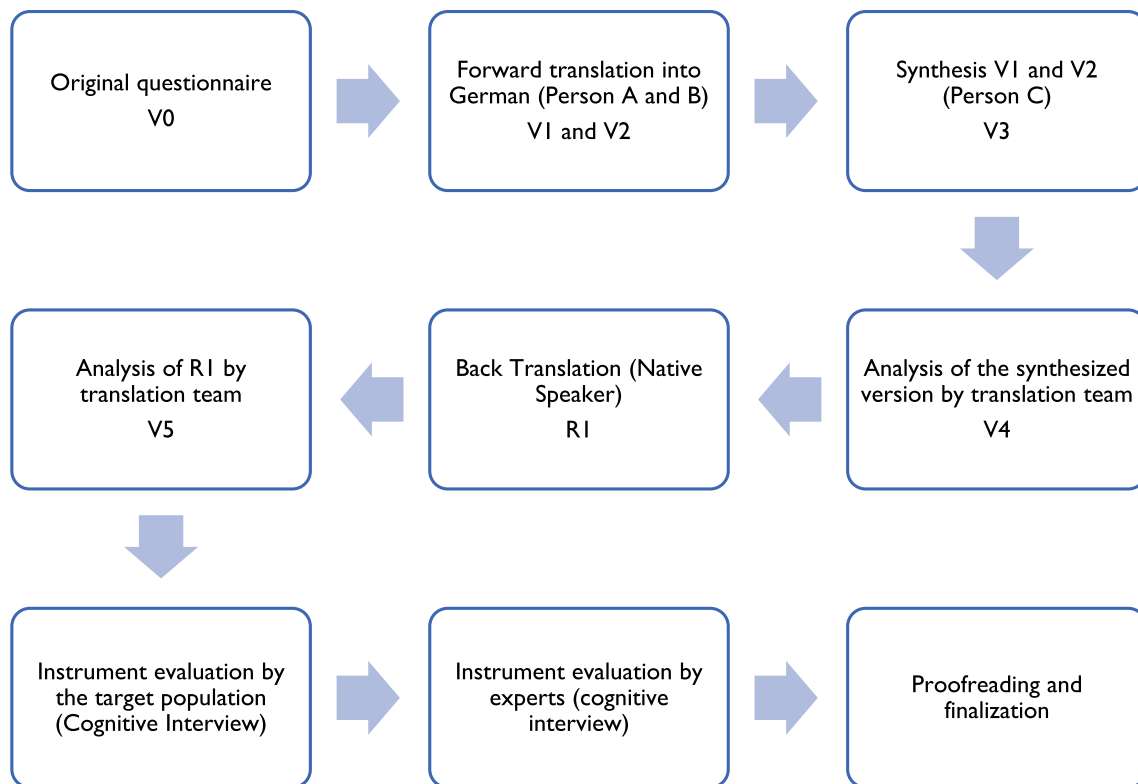
Fig. 1 Translation and linguistic validation sequence

2.2 Synthesis of the translated versions

One author (PCS) of this study served as a third translator who is also a native speaker of the target language and fluent in the source language. Translations V1 and V2 were compared, and a third unified version (V3) was synthesized from the two forward translations. Given different translation versions, the one that appeared to be semantically more accurate was selected.

2.3 Analysis of the synthesized version

All three translators then agreed together on a variant of each item. In addition, regarding the resulting translation versions, the effect of different dialects of the target language was also discussed and the most unique variant was selected. This discussion and review process is referred to as "alignment" and the resulting translation version is referred to as V4.

2.4 Back translation

The translation (V4) was then translated back into English by a female native English speaker who did not work in the field of psychology or medicine and did not know the original English items. This document is referred to as the "back translation" (R1). Back-translation is a quality control step to demonstrate that the target language version does not have different content or conceptual underpinnings that would affect psychometric properties and compromise data quality.

The study leader compared the original English instrument with the back translation (R1) and discussed critical issues in the back translation with the forward translation team. Reconciliation to ensure conceptual equivalence of the translation included identifying discrepancies between the original English version and the back translation and refining the target language versions. The goal was to minimize mistranslations or omissions and analyze causes of the discrepancies. Since the back-translation (R1) could not be unanimously agreed on in all items, a modified version (V5) was jointly

prepared to adapt the German version to the English original and thus achieve harmonization of the two linguistic versions. After the retranslation (R2) of V5 by the native English speaker, the forward translation team agreed to forward V5 for the cognitive interviews.

2.5 Cognitive debriefing

Instrument Evaluation by the Target Population (Cognitive Interview) Conducting cognitive interviews is an evidence-based, qualitative method specifically designed to investigate to assess the degree of comprehensibility and cognitive equivalence of the translations and to highlight elements that are inappropriate at the conceptual level [27]. Cognitive interviewing uses techniques from psychology and traditionally assumes that respondents go through a series of cognitive processes when answering items [28]. These steps include understanding an item and selecting responses, retrieving appropriate information from long-term memory, making judgments based on understanding the item and its memory, and finally selecting a response [29]. An important first step in the cognitive interview process is to establish coding criteria that reflect the survey creator's intended meaning for each item [30]. These can then be used to interpret the responses collected during the cognitive interview.

The two main techniques for conducting a cognitive interview are what is known as "Think aloud" and "General Probing". The "Think aloud" method requires the interviewee to verbalize every thought they have while answering each item. Here, the interviewer simply supports this activity by encouraging the respondent to keep talking and recording what is said for later analysis [27]. This technique can provide valuable information but is unnatural and difficult for most respondents and can result in reams of free-response data that the survey designer must then analyze [28].

A complementary procedure, "Verbal probing," is a more active form of data collection in which the interviewer uses and administers a series of checking questions to elicit specific information [27]. "Probing" is classically divided into concurrent and retrospective probing. In concurrent testing, the interviewer asks respondents specific questions about their thought processes while the respondent answers each question. While concurrent testing is disruptive, it offers the advantage of allowing participants to answer questions while their thoughts are still present [28]. Retrospective testing, on the other hand, occurs after the participant has completed the entire survey (or a portion of the survey) and is generally less disruptive than concurrent testing [28]. The disadvantage of retrospective testing is the risk of recall bias and hindsight effects [31].

For the pretest of the translated questionnaire, concurrent probing was therefore chosen as a method of cognitive interviewing. Unclear meanings or linguistic presentation of items and related difficulties in answering the items were reflected to the translation team. If other difficulties were related to the formatting of the items, they were continuously modified and improved accordingly.

As a qualitative technique, analysis relies on coding and interpretation of transcribed interviews. Therefore, sample sizes for cognitive interviewing are typically small and might include only about ten to 30 participants [27]. The ISPOR guidelines indicate that the tests should involve five to eight respondents who are native speakers of the target language and represent the target population in terms of clinical and sociodemographic characteristics.

Purposive sampling was used to recruit a total of ten female adults from different parts of Germany for the pretest. Participants needed to meet the following criteria: at least 18 years old, have given birth at least 6 weeks ago, describe German as their native language and provide informed consent prior to study entry. The study coordinator contacted the women who were interested in participating to schedule their interview time. The one-on-one cognitive interviews took place online and were conducted by the study coordinator from December 2021 to January 2022.

The items were not previously known to the respondents. During the interview, participants read and answered the items aloud while talking about how they interpreted the item and how they arrived at their answer. If the participant hesitated, the interviewer began asking questions such as, "How would you rephrase this sentence in your own words?" or "How do you interpret [...]?". The interview concluded with a general question, "What do you think of this questionnaire?" to allow respondents to provide additional feedback on specific items or the questionnaire as a whole.

The ten cognitive interviews were transcribed and anonymized verbatim by the study coordinator, and then the following explanations were created for each item: (a) how did the respondent construct her answers; (b) how did she interpret the questions; and (c) what difficulties did she have in answering the questions [32]. In the next step, the study coordinator coded the marked problems using Willis' [33, 34] coding system. It distinguishes seven main categories with specific subcategories (see Table 1).

The results of the cognitive interviews were discussed in the translation team after every second interview. In the event of significant comprehension problems of an item by the subjects, the items were revisited, and an alternative

Table 1 Coding systems for classifying questionnaire problems of Willis [33, 34]

Clarity	Problems with the intent or meaning of a question
Knowledge	Likely to not know or have trouble remembering information
Assumptions	Problems with assumptions or underlying logic
Response categories	Problems with the response categories
Sensitively	Sensitive nature or wording/bias
Instructions	Problems with introductions or explanations
Formatting	Problems with lay out or question ordering

suggestion was jointly developed. An item was considered not understood if it was misunderstood or the meaning was misinterpreted in two interviews. In this case, the item had to go through the forward and backward translation phases again. In turn, it also had to be examined whether such an item should possibly be replaced by another item or omitted in the event of renewed translation and comprehension uncertainty. After the ten participants were recruited and interviewed, no new information could be generated through the interviews, indicating that saturation was reached.

Instrument evaluation by experts Finally, qualitative comments were obtained from the expert panel of advanced practicing midwives to confirm semantic equivalence and content validity. Inclusion criteria for being called as an expert was a master's degree in midwifery and at least 2 years of professional experience in the context of outpatient care of families with infants. Four experts in the field of midwifery were selected for this study. Each expert was asked to assess the translated questionnaire using the think-aloud method. During the assessment, the experts' verbally expressed thoughts and reflections were recorded and then transcribed. The analysis focused on identifying comprehension problems, semantic differences, and possible cultural nuances.

2.6 Proofreading and finalization

In the final review, corrections regarding typos, grammar and layout errors of the questionnaire were made and resulted in the final German version V6 for use in a test group. Each work step was documented in detail so that the adaptation of the item remains traceable in all steps. This so-called item history of the individual work steps allows recourse to the first versions of the adaptation in the case of modified items.

3 Results

3.1 Translation process

During the translation process of the PBQ into German, some idiomatic and semantic issues emerged. The two independent translations of the PBQ into German were quite similar and, for the most part, easily converted into a synthesized forward translation. As a result of the first backward translation from German into English, nine items were found to be problematic and had to be retranslated, as shown in Table 2. For some items, the words could be easily adapted to the original ones, whereas for other items the correct wording had to be found. Some items that were difficult to translate due to idiomatic and semantic differences, for example, had to be thoroughly discussed among the members of the translation team. In particular, the translation of item pairs 7 and 8, 14 and 15, and 19 and 20 was found challenging at the idiomatic level. For example, the English expressions "My baby winds me up." (Item 7) and "My baby irritates me." (Item 8) produced the same German meaning "My baby irritates me." produced in the forward translation.

The multiple forward and backward translation processes allowed for changes with respect to the English version, where words or terms have a certain meaning in English but a semantically different or secondary meaning in German.

3.2 Semantic equivalence

The demographic data of the interview participants of the ten cognitive interviews are summarized in Table 3.

Of the total 25 items, 14 (56.0%) were commented on in the interviews and were adjusted accordingly on an ongoing basis in terms of their problem category (see Table 4).

Table 2 Translation: concerns and comments of the translation team explaining the changes

Item number (original)	Concerns	Changes (back translation)
Alle Items with "child" („baby")	Translation "child" also refers to children over 1 year, not just infants	Baby
Item 7: My child annoys me ("My baby winds me up.")	According to the first backward translation, "My child annoys me."; "annoys" emphasizes the child (as in the original), "nerves" emphasizes the mother	My baby upsets me
Item 8: "My child unsettles me. („My baby irritates me.")	According to first backward translation, "My child unsettles me." which is not a synonym for "irritates me"	My baby bothers me
Item 11: I like to play with my child. ("I enjoy playing with my baby.")	According to the first backward translation "I like to play with my child."; "like" is less emotional than "make fun"	I enjoy playing with my baby
Item 18: I have harmed my child before. („I have done harmful things to my baby.")	According to first backward translation "I have harmed my child before."; means "I have hurt my baby."; emphasis here is on pain, not harm	I did harmful things to my baby
Item 19: My child makes me nervous. („My baby makes me anxious.")	After first backward translation "My child makes me nervous."; absence of the anxious-nervous aspect	My baby makes me feel anxious
Item 21: My child upsets me. („My baby annoys me.")	After first backward translation "My child upsets me.", too much similarity with item 7	My baby is annoying me
Item 23: I think the only way out is for someone else to take care of my child. („I feel the only solution is for someone else to look after my baby.")	According to first backward translation "I think the only way out is for someone else to take care of my child."; "way out" more dramatic than "solution"	I think the only solution is to have someone else take care of my baby
Item 24: I want to hurt my child. („I feel like hurting my baby.")	According to first backward translation "I want to hurt my child."; "would like" more active verb (intention) than "feel"	I feel like hurting my baby
Item 25: My child calms down easily. („My baby is easily comforted.")	According to first backward translation "My child calms down easily." child as an active part of the statement, passive in the original	My baby is easily comforted

Table 3 Demographic characteristics of cognitive interview participants (n = 10)

	Mean	SD[a]	V[b]
Age, years	36,3	6,3	36,0
Gender			
Female	10		
Male	0		
Divers	0		
Number of children	1,7	0,95	0,9
Age of child(ren), years	6,7	5,5	28,9
School education			
High school diploma	10		
University degree	6		
Region of Germany			
North	4		
West	1		
East	2		
South	3		

[a]Standard Deviation; [b]Variance

Table 4 Outcome of cognitive debriefing of the German-translated PBQ and number of modified items

PBQ-dimension	Number of items	Number of items with comment	Questionnaire problems*				Modified items according to con-sensus
			C	A	RC	S	
1	12	5	2	4	2	3	5
2	7	5	4	2	1	1	5
3	4	3	2	2			3
4	2	1	1			1	1

* Coding systems for classifying questionnaire problems of Willis (1999, 2009)

C: Clarity; A: Assumption; RC: Response categories; S: Sensitively

During the cognitive debriefing, the four experts identified one item that might need to be corrected. Based on their comments, a wording adjustment was made to a total of one item without changing the original meaning of the item, resulting in the final German version of the PBQ, which is linguistically valid and corresponds to the original scale (see Table 5).

4 Discussion

In the present study, the PBQ was successfully translated into German, culturally adapted, and linguistically validated. The produced German version corresponds to the original English version in content and scope and is now ready for psychometric validation. The culturally adapted and linguistically validated questionnaire differs from the translation of Reck et al. [12] in 20 out of 25 items.

The cultural adaptation of psychological instruments is a complex task that requires careful planning in terms of content preservation, psychometric properties, and generality to the target population [35]. This involves providing evidence of both semantic equivalence of items and adequate psychometric properties of the new version of the instrument [35].

4.1 Strengths and limitations

A major strength of this study was the adherence to established guidelines in the translation and language validation process, which included several steps. The English items of the PBQ, following the guidelines for high quality translation of questionnaires with a multi-step translation process, proved to translate well. The guidelines and

Table 5 Semantic equivalence: comments of the test persons and experts as well as final changes made by the translation team

Item number	Comments	Changes
Item 1: I feel close to my baby	Often interpreted spatially or "feel close" as inaccurately evaluated (physically or emotionally)	I feel closeness to my baby
Item 3: I feel distant from my baby	Exact meaning of "feeling distant" not clear; naming causes would be helpful; "being distant/ indifferent" is clearer than "feeling distant"	I am indifferent to my baby
Item 4: I love to snuggle with my baby	Physical expression better with "cuddling"; "I love it" sounds too absolute	I like to cuddle with my baby
Item 6: The baby doesn't seem to be mine	Unclear category: Physical or emotional; if emotional level is addressed, more of a sense of belonging, other- wise, great similarity to Item 1	The baby feels strange
Item 11: I enjoy playing with my baby	"Playing" is not a given at the beginning, so " spend time" is more suitable	I like to spend time with my baby
Item 12: My baby cries too much	Problem term "too much", since no objective measure is given	My baby cries too much for me
Item 16: My baby is the most beautiful baby in the world	Problem term "beautiful"; Too exaggerated statement	My baby is gorgeous
Item 19: My baby makes me anxious	"Anxious" means "to be afraid" and weighs more heavily than "to be restless/worried"; similar- ity to Item 20	My baby makes me restless
Item 20: I am afraid of my baby	Vague meaning: General anxiety in everyday situations	My baby scares me
Item 21: My baby is irritating me	Identical meaning with items "my baby makes me angry/my baby upsets me"; "nerves" occurs at the physical rather than emotional level	My baby is a nuisance to me
Item 23: I think the only solution is to have someone else take care of my baby	"Solution" does not emphasize the emotional aspect, rather the result of a rational thought process	I wish someone else would take care of my baby
Item 24: I feel like hurting my baby	Unclear meaning: thinking or executing	I have the urge to hurt my baby

recommendations of an experienced working group regarding questionnaire translation that are already widely used and established internationally [26] were applied in the present study. It helped to ensure that a German-language instrument was formulated in a clearly understandable and conceptually equivalent way and could therefore be used in Germany.

Translation The chosen translation method enabled two independent translations, which were characterized by a simple and target group-oriented formulation and were largely free of colloquial expressions [26]. During the translation process, the necessity of a detailed translation methodology became apparent: Due to the cultural language differences, margins of interpretation arose for individual items, which could only be captured by a multiprofessional team in individual steps of pre- and back-translation. In this context, the back-translation step, which has been judged in all studies to be crucial regarding the comparison and requirements of high-quality translation processes [26], proved to be crucial to detect comprehension errors or excessively wide margins of interpretation in the outward translation.

Equivalence testing Backward translation was recommended to detect weaknesses in forward translation [26]. The PBQ items were first pre-translated and then back translated into English in an elaborate process. The procedure chosen considered the measures essential for backward translation and subsequent review of backward translation [36]: The back translator was a native English speaker with fluency in German who had neither been involved in the previous translation steps nor had access to the original items and their definitions. Various items had to be corrected in the German target language after their back-translation into the original English language. In this process, the back-translation process, which is normally ensured by only one native speaker, is more prone to errors than the forward translation, which is usually produced by at least two translators and finally decided by consensus [36]. However, the ISPOR working group [26] identifies equivalence testing as a very important measure for the cultural adaptation of quality-of-life measurement instruments. During the backward translation review of this work, difficult concepts were discussed, and inappropriately translated items were identified.

Cognitive testing Cognitive interviews are conducted in the development phase of a questionnaire to clarify what problems may arise when answering each individual question. They minimize the risk of missing or incorrect data in surveys caused by respondents' unconscious misunderstanding of questions [26].

The involvement of users of the questionnaire and experts was instrumental in the linguistic validation of the German version. The assessment of the users through cognitive interviews was particularly important in verifying the semantic equivalence of the translated PBQ. These cognitive interviews were conducted using the "Think Aloud", "General Probing" and "Paraphrasing" as evidence-based qualitative methods, and the sample size used (n = 10) was comparable to that used in published research of a similar nature and exceeded the ISPOR guideline recommendations of 5–8 patients.

"General probing" and "Think aloud" gave each study participant the opportunity to give their own thoughts and comments on each answered question and to assess the difficulty of the question themselves. It turned out that most of the items were perceived as clearly understandable by the participants. Some questions immediately emerged as difficult to understand. As these were discussed with the participants during the interview, this step often resulted in helpful suggestions for improvement.

The cognitive technique of "General probing" has some weaknesses, so that it should not be used alone to adequately assess the understanding of questions. "General probing" seems only productive in cases in which the respondent is aware that they do not understand the question. The respondent may not be aware of possible discrepancies between their own understanding of the question and that of the developer of the question because they do not know the developer's intention.

The use of a single cognitive method did not provide sufficient information about the quality of the translated items. The additional method of "Paraphrasing" revealed that some translations had considerable comprehension difficulties. "General probing" was essential to identify subjective problems with the participants' understanding of the questions and to obtain possible suggestions for improvement. The combination of the cognitive methods of "General probing" and "Paraphrasing" enabled a more comprehensive understanding of the items from the participants' perspective. As a result, conceptual equivalence, comprehensibility, and target group suitability of all items were achieved.

The major limitation of the study is the recruitment of a highly qualified sample in which all participating women had a high school degree and 60% had a college degree. A more representative approach would be to survey women with different levels of education.

The use of the "think-aloud" method allowed the experts to share their thoughts and reflections in real time, resulting in a detailed recording of their cognitive processes. The identified comprehension problems and semantic differences highlight the complexity of questionnaire translation and the need for careful validation procedures. Also, the results emphasize the importance of involving experts in the evaluation of translated instruments to ensure quality and accuracy.

Country-specific approach There are three main approaches to translating a quality-of-life measurement instrument into a language that is spoken in different countries [37]. While the country-specific approach and the adaptation approach produce a separate language version in the same language for each country, the universal approach produces a common language version for all countries in a linguistic and cultural area. In this work, we have opted for the country-specific approach, as the universal approach (joint translation for Germany, Austria, and German-speaking Switzerland) does not take into account the cultural and linguistic characteristics and peculiarities of the individual subcultures and the pursuit of a common language version is comparatively time-consuming.

In our study, this assumption of avoiding idiomatic as well as dialectical expressions was met. The translators did not belong to any linguistic subgroup and the subject group consisted of a random composition of women from all regions of Germany. It can be assumed that the translation result would have varied with the participation of translators or subjects from a dialectical subgroup. A new cognitive debriefing by colleagues from German language communities outside Germany, e.g., Switzerland or Austria, before introducing the German PBQ in one of the two countries therefore seems necessary.

Conceptual equivalence The central object of linguistic validation is to ensure conceptual equivalence. This means that the translation and the original version must represent the same underlying construct. This enables international data collection and comparisons between countries. To achieve this goal, special attention was paid to content equivalence in numerous steps of the translation process. Even at the translation stage, the translators preferred a meaning-based wording to a literal one.

According to the definition of conceptual equivalence, an equivalent translation is linked to the condition that the constructs to be translated are equally relevant or acceptable in the target culture [35]. As also described in the literature, a translated instrument may lack content equivalence if there are interculturally different conceptions of social desirability in the construct to be measured [38]. Due to the response categories used in the original instrument (Response Options: Always–Very often–Quite often–Sometimes–Rarely–Never), it can be assumed that these absolute answers are usually given in English. Based on the results of the cognitive interviews, however, it must be assumed that a German-speaking participant would choose a less extreme answer for fear of appearing unloving or unrealistic. As a result, the response categories for the German-speaking cultural area were adapted to minimize social desirability bias (German Response Options: Almost always–Very often–Often–Sometimes–Rarely–Almost never).

Despite a guideline-compliant translation and linguistic validation, an item in the target language might be stronger or weaker in capturing a trait expression than the original item, resulting in a shift in item parameters. Therefore, it is imperative that the translation process be followed by quantitative methods for test-statistical validation and calibration of the items to produce the final German PBQ. The translation and linguistic validation in itself is not sufficient to generate a measurement instrument equivalent to the English items. The goal of future work must therefore be to verify the psychometric properties of the German PBQ to be able to compare it with the existing results of the English PBQ.

5 Conclusion

Adequate methods must be used for the linguistic and cultural validation of questionnaires. Translation and adaptation procedures, cultural adaptations, and cognitive interviews with target groups are proven methods to ensure the accuracy and comprehensibility of the instruments. The involvement of experts from the relevant care contexts is crucial to ensure validity.

The PBQ has been successfully translated into German, culturally adapted, and linguistically validated. The final German translation will be validated in the next step by the authors for psychometric properties such as validity and reliability before it can be recommended for use in future research.

Data availability The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

References

1. Klier C. Mother–infant bonding disorders in patients with postnatal depression: the postpartum bonding questionnaire in clinical practice. Arch Women's Ment Health. 2006;9:289–91.
2. Bicking Kinsey C, Hupcey JE. The state of the science of maternal-infant bonding: a principle-based concept analysis. Midwifery. 2013;29:1314–20.
3. Taylor A, Atkins R, Kumar R, Adams D, Glover V. A new Mother-to-Infant Bonding Scale: links with early maternal mood. Arch Women's Ment Health. 2005;8:45–51.
4. Brockington IF, Oates J, George S, Turner D, Vostanis P, Sullivan M, et al. A screening questionnaire for mother-infant bonding disorders. Arch Women's Ment Health. 2001;3:133–40.
5. Condon JT, Corkindale C. The assessment of parent-to-infant attachment: development of a self-report questionnaire measurement. J Reprod Infant Psychol. 1998;16:57–77.
6. Wittkowski A, Vatter S, Muhinyi A, Garrett C, Henderson M. Measuring bonding or attachment in the parent-infant-relationship: a systematic review of parent-report assessment measures, their psychometric properties and clinical utility. Clin Psychol Rev. 2020;82: 101906.
7. Dearman L, Musonda P, Roberts F, Bowles K, Morris E. Bonding in women with postnatal anaemia: a pilot case control study looking at postnatal bonding in women who have been diagnosed with anaemia at a University Hospital in the East of England. Arch Gynecol Obstet. 2012;285:1243–8.
8. Edhborg M, Matthiesen AS, Lundh W, Widstrom AM. Some Early indicators for depressive symptoms and bonding 2 months postpartum—a study of new mothers and fathers. Arch Women's Ment Health. 2005;8:221–31.
9. Edhborg M, Nasreen HE, Kabir ZN. Impact of postpartum depressive and anxiety symptoms on mothers' emotional tie to their infants 2–3 months postpartum: a population-based study from rural Bangladesh. Arch Women's Ment Health. 2011;14:307–16.
10. Hornstein C, Trautmann-Villalba P, Hohm E, Rave E, Wortmann-Fleischer S, Schwarz M. Maternal bond and mother-child interaction in severe postpartum psychiatric disorders: is there a link? Arch Women's Ment Health. 2006;9:279–84.
11. Pearson RM, Heron J, Melotti R, Joinson C, Stein A, Ramchandani PG, Evans J. The Association between observed non-verbal maternal responses at 12 months and later infant development at 18 months and IQ at 4 years: a longitudinal study. Infant Behav Dev. 2011;34:525–33.
12. Reck C, Klier CM, Pabst K, Stehle E, Steffenelli U, Struben K, et al. The German version of the postpartum bonding instrument: psychometric properties and association with postpartum depression. Arch Women's Ment Health. 2006;9:265–71.
13. Wittkowski A, Wieck A, Mann S. An evaluation of two bonding questionnaires: a comparison of the mother-to-infant bonding scale with the postpartum bonding questionnaire in a sample of primiparous mothers. Arch Women's Ment Health. 2007;10(4):171–5.
14. Brockington IF, Fraser C, Wilson D. The postpartum bonding questionnaire: a validation. Arch Women's Ment Health. 2006;9:233–42.
15. Kerstis B, Aarts C, Tillman C, Persson H, Engström G, Edlund B, et al. Association between parental depressive symptoms and impaired bonding with the infant. Arch Women's Ment Health. 2016;19:87–94.
16. Mitchell EA, Lynne Hutchison B, Thompson JMD, Wouldes TA. Exploratory study of bedsharing and maternal-infant bonding. J Paediatr Child Health. 2015;8:820–5.
17. Moehler E, Brunner R, Wiebel A, Reck C, Resch F. Maternal depressive symptoms in the postnatal period are associated with long-term impairment of mother-child bonding. Arch Women's Ment Health. 2006;9(5):273–8.
18. van Bussel JC, Spitz B, Demyttenaere K. Three self-report questionnaires of the early mother-to-infant bond: reliability and validity of the Dutch version of the mpas, pbq and mibs. Arch Women's Ment Health. 2010;13(5):373–84.
19. Wittkowski A, Williams J, Wieck A. An examination of the psychometric properties and factor structure of the post-partum bonding questionnaire in a clinical inpatient sample. Br J Clin Psychol. 2010;49(Pt 2):163–72.
20. Gjersing L, Caplehorn J, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. BMC Med Res Methodol. 2010;10:13.

21. Yasir A. Cross cultural adaptation & psychometric validation of instruments: stepwise description. Int J Psychiatry. 2016;1(1):1–4.
22. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25:3186–91.
23. Radtke F, Franck M, Drews T, Weiß-Gerlach E, Twigg E, Harbeck-Seu A, et al. Die posttraumatische Stressskala-14 (PTSS-14) - Validierung der richtlinienkonformen Übersetzung für den deutschsprachigen Raum. Anästhesiologie, Intensivmedizin, Notfallmedizin und Schmerztherapie. 2010;45:688–95.
24. Epstein J, Santo R, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. J Clin Epidemiol. 2015;68:435–41.
25. Hambleton RK, Li S. Translation and adaptation issues and methods for educational and psychological tests. In: Frisby CL, Reynolds CR, editors. Comprehensive handbook of multicultural school psychology. Hoboken: Wiley; 2005.
26. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. 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 Health. 2005;8(2):94–104.
27. Willis G, Artino A. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. J Grad Med Educ. 2013;5(3):353–6.
28. Artino A, La Rochelle J, Dezee K, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach. 2014;36:463–74.
29. Tourangeau R, Rips L, Rasinski K. The psychology of survey response. New York: Cambridge University Press; 2000.
30. Karabenick S, Woolley M, Friedel J, Ammon B, Blazevski J, Bonney C, et al. Cognitive processing of self-report items in educational research: do they think what we mean? Educ Psychol. 2007;42(3):139–51.
31. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. J Adv Nurs. 2003;42(1):57–63.
32. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. Public Opin Q. 2007;71:287–311.
33. Willis G. Cognitive interviewing. A tool for improving questionnaire design. Thousand Oaks: Sage Publications; 2005.
34. Willis G. Analysis ode the cognitive interview in questionnaire design. Oxford: University Press; 2015.
35. Borsa JC, Damásio BF, Bandeira DR. Cross-cultural adaptation and validation of psychological instruments: some considerations. Paidéia. 2012;22(53):423–32.
36. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. Eval Health Prof. 2005;28:212–32.
37. Wild D, Eremenco SL, Mear I, Martin M, Houchin C, Gawlicki M, Hareendran A, Wiklund I, Chong LY, Von Maltzahn R, Cohen L, Molsen E. Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: The ISPOR patient-reported outcomes translation and linguistic validation good research practices task force report. Value Health. 2009;12:30–440.
38. Marquis P, Keininger D, Acquadro C, De La Loge C. Translating and evaluating questionnaires - cultural issues for international research. In: Peter M, Fayers RDH, editors. Assessing quality of life in clinical trials: methods and practice. 2nd ed. Oxford: Oxford University Press; 2005.