



What do people respond to when rating executive function? – a cognitive interviewing investigation of BRIEF-A informant ratings in severe aphasia

Camilla Olsson , Patrik Arvidsson & Monica Blom Johansson

To cite this article: Camilla Olsson , Patrik Arvidsson & Monica Blom Johansson (2020): What do people respond to when rating executive function? – a cognitive interviewing investigation of BRIEF-A informant ratings in severe aphasia, Disability and Rehabilitation, DOI: [10.1080/09638288.2020.1849418](https://doi.org/10.1080/09638288.2020.1849418)

To link to this article: <https://doi.org/10.1080/09638288.2020.1849418>



© 2020 The Author(s). Informa UK Limited, trading as Taylor & Francis Group



Published online: 03 Dec 2020.



Submit your article to this journal [↗](#)



Article views: 251



View related articles [↗](#)



View Crossmark data [↗](#)

What do people respond to when rating executive function? – a cognitive interviewing investigation of BRIEF-A informant ratings in severe aphasia

Camilla Olsson^a , Patrik Arvidsson^b  and Monica Blom Johansson^a 

^aDepartment of Neuroscience, Speech-Language Pathology, Uppsala University, Uppsala, Sweden; ^bCentre for Research and Development, Uppsala University/Region Gävleborg, Gävle, Sweden

ABSTRACT

Purpose: Executive dysfunction is common in persons with severe aphasia. Assessing these functions in this population is challenging. Informant ratings, such as the BRIEF-A, might be a useful alternative to neuropsychological tests. However, research has shown weak relationships between tests and ratings. The aim of this study was to understand how significant others of people with severe aphasia interpret and respond to questions about executive function in the informant report version of BRIEF-A.

Methods: Eleven significant others were interviewed about a subset of the BRIEF-A items, using cognitive interviewing. Interviews were subjected to thematic analysis.

Results: There was variation in the interpretation of the items of BRIEF-A which frequently corrupted the items' relation to what it was intended to measure. Further, informants wavered between considering the person with aphasia's ability or actual performance and many had lowered their expectations. The language problems caused by the aphasia affected the validity of some items.

Conclusions: The quantitative results of BRIEF-A informant ratings should be interpreted with caution, since it is unclear to what extent the responses represent executive function. The use of informant ratings does not solve the problem with aphasia being a confounding factor in assessment of executive function.

ARTICLE HISTORY

Received 31 January 2020
Revised 9 September 2020
Accepted 8 November 2020

KEYWORDS

Aphasia; executive function; informant rating; stroke; BRIEF-A

► IMPLICATIONS FOR REHABILITATION



- Assessing executive function in people with severe aphasia is important but challenging.
- *Quantitative* results of informant ratings of executive function, such as BRIEF-A, in this population should be interpreted with caution, since it is unclear to what extent the ratings represent executive function.
- Using informant ratings does not solve the problem of the aphasia being a confounding factor, since the aphasia impacts on the validity of some of the items.

Introduction

Executive function refers to a group of cognitive control processes. It is usually viewed as an umbrella concept, containing several sub-processes that are required to initiate, plan, execute and monitor goal directed behaviour [1] and that operate mainly on other cognitive processes and behaviours [2]. There are many different definitions of, and theories about, what executive function is and how it should be described [1–6]. It has been proposed that there is relative agreement on what executive functions *do*, but still no consensus as to what they *are* [7]. Executive dysfunction is common after stroke and has negative impact on rehabilitation outcome, for example through disrupting the ability to effectively use remaining areas of functioning as well as the learning and using of compensatory strategies [8–10]. Further, executive dysfunction impacts on daily living, for instance through impaired planning and problem solving skills [11].

A further, common sequela after stroke is aphasia, a language disorder caused by acquired brain damage. Aphasia affects the

production and comprehension of both spoken and written language [12], to varying degrees. In many cases, aphasia is also associated with difficulties in understanding numbers and calculating [13,14]. Persons with severe aphasia (PWSA) have severely limited or absent verbal output, paired with varying levels of language comprehension impairment. In stroke survivors with severe aphasia, an additional executive dysfunction has negative impact on quality of life [15], language rehabilitation outcome [16] and functional communication [17,18]. In severe aphasia, the individuals need to have the strategic competence to bypass their limitations by using compensatory strategies as well as making the most out of the limited linguistic ability they have [19], in order to make themselves understood. Executive functions are crucial to such flexible problem-solving behaviour [10]. In a clinical setting, it is thus important for both prognosis and intervention planning to have reliable information about the executive functions of PWSA. However, obtaining such information is a challenging task.

CONTACT Camilla Olsson  camilla.olsson@neuro.uu.se  Department of Neuroscience, Speech-Language Pathology, Uppsala University, Box 593, Uppsala, SE, 751 24, Sweden

© 2020 The Author(s). Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Table 1. Examples of items from Behavior Rating Inventory of Executive Function – Adult version^a.

Item no	Item text
11.	Has trouble with jobs or tasks that have more than one step
43.	Makes decisions that gets him/her into trouble (legally, financially, socially)
56.	Has trouble remembering things, even for a few minutes (such as directions, phone numbers)

^aReproduced by special permission of the Publisher, Psychological Assessment Resources, Inc. (PAR), 16204 North Florida Avenue, Lutz, Florida 33549, from the Behavior Rating Inventory of Executive Function-Adult Version by Robert M. Roth, PhD, Peter K. Isquith, PhD and Gerard A. Gioia, PhD, Copyright 1996, 1998, 2001, 2003, 2004, 2005 by PAR. Further reproduction is prohibited without permission from PAR.

Executive functions are commonly assessed either with standardised, neuropsychological, performance-based tests and/or by using validated rating instruments for self or informant report [20]. Both kinds of assessments are commonly used by clinical neuropsychologists [21,22] and have been assumed to target more or less the same underlying construct, but with the intention that the latter could provide better ecological validity than the former [23].

The linguistic impairments of PWSA pose a challenge when assessing executive function, since most standardised neuropsychological tests are not adapted for this population [24]. The linguistic demands of understanding instructions, giving verbal responses or dealing with linguistic or numerical content in test materials are serious confounding factors. In addition, most self-rating instruments are also inappropriate due to linguistic demands, unless the instruments are subjected to extensive adaptations, which of course makes it difficult to use existing norms or cut offs. This suggests that informant ratings of executive function, completed by family members or formal caregivers, might seem like a particularly appealing alternative in this population, as a means of bypassing the linguistic challenges.

However, studies across different clinical populations have demonstrated that the correlations between neuropsychological tests and ratings of executive functions are typically weak or at best moderate [23,25], indicating that the overlap in what these assessments capture is in fact rather limited. Several explanations for this phenomenon have been suggested. For example, Toplak et al. [23] propose that tests and ratings assess different levels of cognitive functioning, referring to differences between algorithmic and reflective mind [26]. Other authors assume that tests and ratings capture separate dysexecutive syndromes [27], or that influence of the respondent's personality on ratings can cause discrepancies [7]. It has also been suggested that the structured administration situation during neuropsychological testing simply does not lend itself to assessing executive functions that are primarily needed in novel, non-routine and highly emotional situations [21]. For this reason, it is often suggested that neuropsychological tests and self- or informant ratings should be used as complementary sources of information in clinical settings [25].

Ratings can be made by the person him/herself (self-rating), or by a significant other who have extensive knowledge about the everyday functioning of the person (informant rating). In studies of the agreements between self- and informant ratings in general some discrepancies are usually found, with informants sometimes reporting less problems and sometimes more problems than the person him/herself. Regarding ratings of executive function there are indications that informants have a tendency to report less problems than the persons themselves [25,28]. Simultaneously, it has also been suggested that informant ratings might be more accurate since populations with executive dysfunction can have poor self-awareness of problems [20]. Regarding self versus informant ratings of other areas, such as activities of daily living or quality of life, it is a common finding that informants tend to

overestimate the problems in relation to both stroke patients in general and those with aphasia. Agreement between self and informant rating is usually larger for objective, observable behaviour than for subjective domains [29,30]. It has been suggested that, in spite of the slight disagreements, it is generally better to use informant ratings than to systematically exclude individuals with aphasia from stroke research, as is frequently done [29,31].

There are several instruments for rating of executive function. One of the most commonly used is Behavior Rating Inventory of Executive Function – Adult Version (BRIEF-A) [28], and this is also the instrument used in this study. In BRIEF-A, the respondents are asked to rate how often they experience problems related to executive dysfunction in daily life. This method of accessing information about everyday functioning is similar to other instruments for rating of executive function. The contents of the items are somewhat similar across instruments; thus BRIEF-A is in this study to be regarded as an example of questionnaires about executive function.

BRIEF-A is designed for adults who might have executive dysfunction due to developmental, neurological or psychiatric illnesses, and is thus not specifically developed for the stroke population. It is reported to be reliable across a wide range of demographic contexts [28] and there are parallel self and informant report versions. The questionnaire consists of 75 items belonging to nine clinical scales that target different constructs or aspects of executive function: Inhibit, Shift, Emotional Control, Self-monitor, Initiate, Working Memory, Plan/Organise, Task Monitor, and Organisation of Materials. The informant is asked to rate how often the individual, in this case the PWSA, has had problems with the behaviour described in the item during the last month, see Table 1 for examples. Response is given on a three-point Likert scale; never, sometimes, often. BRIEF-A is not recommended by the developers to be used in isolation as a diagnostic tool, but to be used together with other assessments [28].

In an earlier study of 38 individuals with severe aphasia where BRIEF-A informant ratings were conducted [32], we found even fewer correlations between the neuropsychological tests and the informant ratings of executive function than reported in other populations [23]. We also found a consistent bias towards informant ratings showing significantly less executive dysfunction than neuropsychological tests. In addition, the results of the BRIEF-A ratings did not show any relationship to functional communication. Relationships between functional communication and executive function, when measured with neuropsychological tests, have been previously reported [17,18]. Thus, BRIEF-A does not seem to be sensitive to the aspects of executive function that are important to functional communication. The conclusions of the Olsson et al. [32] study were that informant ratings of executive function cannot be used as a substitute for neuropsychological tests in people with severe aphasia, and that the BRIEF-A ratings do not seem to add information about aspects of executive function relevant to functional communication. The relatively high level of executive functioning indicated by the BRIEF-A results was

surprising, considering many of the individuals with aphasia had suffered severe strokes. The limitations that they, based on clinical experience, could be expected to have did not seem to be reflected in the BRIEF-A scores. The use of informant ratings might be particularly difficult in populations with severe aphasia, since the informants have to base their ratings almost exclusively on their own observations, without being able to consider the person with aphasia's experiences of, or explanations for, behaviours. Some support for this speculation comes from the review by Oczkowski and O'Donnell [29], where the most consistent predictor of disagreement between patient and proxy was increasing general stroke severity (making the assumption that severe aphasia is more common following more severe strokes). Indeed, Hilari et al. [30] found that disagreement between self and informant report increased with aphasia severity for at least some domains of a quality of life instrument.

In the use of informant ratings (as well as self-ratings and other questionnaires), it is essential to have a good understanding of how the respondents go about answering the questions. At present, there is a lack of knowledge about what information is actually gained through informant ratings of executive function, such as the BRIEF-A. It could be argued that what ultimately determines what informant ratings measure is what the respondents think the questions ask about, which in turn decides what knowledge and experience they base their responses on. If a question is not comprehended as intended by the respondent, it is difficult to know what the response means [33]. Further investigation of this issue may provide important information to assist in the interpretation of informant ratings of executive function and the understanding of the discrepancies found between ratings and neuropsychological test results. Regarding PWSA, greater insight into how the respondents interpret the questions is needed to shed light on whether the presence of aphasia has any impact on informant ratings of executive function and, consequently, whether such instruments are appropriate for this population.

Cognitive interviewing is a method specifically developed for exploring the process of responding to surveys and questionnaires [34]. With this method, interviews are used to gain knowledge about how the informants understand the questionnaire items and whether the items are well designed to provide the desired information [35]. Since cognitive interviewing is developed to identify the "content or experiences contained in the respondents' answers" [34], it can be used as a study of validity. The method is often applied at the stage of questionnaire development. The focus of interest is then frequently to detect items that the respondents perceive as difficult or problematic in some way, and to gain information about how to improve those items. There is often a focus on analysing the four main components of the process of responding to a questionnaire item; comprehension, retrieval, judgment, and response process [36]. However, the method can also be applied to existing instruments to investigate the validity and to enhance the understanding of how to interpret the responses.

To our knowledge, BRIEF-A has not been investigated with cognitive interviewing methodology, but there are studies of other cognitive instruments that have presented interesting findings. Perhaps most relevant to the present study, Hill et al. [37] investigated older adults' self-ratings of memory function. They found several instances where the structure and content of the questionnaire items influenced the participants' interpretations and decisional processes in a problematic way. Examples of problems that Hill et al. [37] detected were: vague intent or meaning

in question wording, lack of characterisation of the problems of interest, unspecified point of comparison, and incorrect assumptions regarding consistency of cognitive problems across situations. Considering that their study was concerned with ratings of a cognitive function, with construction of the items similar to those used in ratings of executive function, it might be hypothesised that similar problems could occur with instruments such as BRIEF-A.

To conclude, executive dysfunction impacts many aspects of rehabilitation and life. The use of informant ratings for assessment of executive function is common in clinical settings. However, many questions surround the use of such ratings and what the responses actually reflect. Previously suggested explanations for the discrepancies reported between neuropsychological tests and informant ratings are not entirely satisfactory. In addition, there is a lack of research about what impact communication disorders, such as severe aphasia, might have on informant ratings of executive functions. In this study we investigate the way informants (significant others of PWSA) interpret and respond to the questions in the informant report version of BRIEF-A. Our aim is to deepen the understanding of such instruments' relation to neuropsychological tests, and their validity and applicability in the population of PWSA.

Specific research questions were:

1. How do significant others of PWSA interpret items from BRIEF-A?
2. What experience and knowledge do the significant others of PWSA draw upon when responding to BRIEF-A?
3. Does the aphasia have any specific impact on the relevance or validity of items from the informant ratings version of BRIEF-A?

Method

The study was conducted in Sweden and was approved by the regional ethical review board of Uppsala, Dnr 2017/183 and 2019/00189.

Cognitive interviews were conducted with the focus of gaining insight about what information the responses of BRIEF-A reflect [38,39]. It was not our aim to investigate problems in specific items. We adopted an interpretivist approach, acknowledging that the experiences and life situations of each informant has an influence on their interpretation of the questions, rather than stating that there is one correct interpretation [40].

Participants

Potential participants were strategically selected by the authors from the 38 significant others who had completed the BRIEF-A in a previous study [32]. Selection was made to ensure a variation in age, gender and in the type of relationship with the PWSA (family members as well as formal caregivers). They were contacted by telephone and given information about the study. Twelve significant others were approached, all provided informed consent to participate. One person later cancelled the interview due to health issues.

Demographic information about the participants is presented in Table 2.

Table 2. Interview informants, $n = 11$.

Informant	Gender	Age	Relation to PWSA ^a
1	F	50	Spouse
2	M	35	Child
3	F	72	Sibling
4	F	64	Spouse
5	F	70	Parent
6	F	72	Spouse
7	M	62	Formal caregiver
8	F	55	Child
9	F	57	Spouse
10	F	43	Child
11	F	54	Formal caregiver

^aPerson with severe aphasia.

Material

Behavior rating inventory of executive function – adult version (BRIEF-A)

The main characteristics of BRIEF-A have already been described in the introduction. In the present study, a Swedish translation of BRIEF-A, available from the publisher Hogrefe, was used.

Normative tables are provided for BRIEF-A, the instrument is thus intended not only to detect problems, but also to assist in the interpretations of the level of executive function in the individual. BRIEF-A has been reported to have good internal consistency and test-retest reliability. The instrument is further thought to have good validity, based on the content of the items, comparisons to other similar and dissimilar instruments (other questionnaires about cognitive functions versus depression), factor analysis, and BRIEF-A profiles for different diagnostic groups [28].

Interview guide

It was not considered feasible to interview the informants about all 75 items in BRIEF-A. Thus, 20 items were selected by the authors (items 2, 8, 11, 14–19, 22, 23, 25, 43, 56, 57, 60, 62, 64–66). The selection included items from all the clinical scales. It also aimed to cover both items that seemed clear in their intent, and items that were expected to be more complex to interpret and respond to, based on the authors' judgment. Due to copyright regulations and the protection of test integrity, all the selected items cannot be included here. To exemplify, three items are presented in Table 1.¹ An interview guide was developed, containing the 20 selected items and examples of verbal probes to elicit more detailed information (Table 3). The interview guide also contained questions about following the general instructions of BRIEF-A and how the response alternatives (never, sometimes, often) were understood. To assist in the development of probe questions and to prepare for possible challenges during interviewing, the first author conducted three practice interviews. These were conducted with people outside the targeted population and were thus not included in the sample.

Procedure

The interviews were conducted by the first author (CO) during the period March–May 2019. CO is a female registered speech-language pathologist and at the time of the interviews PhD student, who has long experience of clinical work with PWSA and their families. The interviews were conducted over telephone and audio recorded. The participants were recruited from the sample of a previous study [32] and CO had been in contact with all the interview informants, either by telephone or in person, and also met all the PWSA in person, during the previous data collection. The participants lived over a large geographical area and it was

Table 3. Examples of prepared verbal probes used during interviews.

If you had to choose, which response would you decide on?
What do you think is meant by... (e.g., "activities")?
Can you describe a situation you think of when responding to this question?
How often would you say "sometimes" (or "often") is in this case?
Disregarding X (the person with aphasia) for a moment, could you describe what kinds of problems you think this question asks about?

decided both convenient and feasible to perform the interviews over telephone. Prior to the interviews, all informants received a clean copy of the BRIEF-A questionnaire to be used as a visual support during the interview if needed. The interviewer started by explaining that the purpose of the study was to understand how the questions were interpreted and what the respondents based their answers on, reinforcing that there were no wrong answers. This was considered important since the informants might otherwise interpret probe questions as indications that they have provided the wrong answer [41]. The respondents were instructed to base their responses on the current situation. The interviewer then read the instruction on the questionnaire aloud to the informant and asked if there were any ambiguities. Thereafter, one item at a time was read aloud to the informant, who was encouraged to describe what he/she was thinking about while deciding on a response. Prepared verbal probes (Table 3) were used together with spontaneous questions as considered needed to further explore how the informant understood the item and the response process. Thus a combination of think aloud technique and verbal probing was used [33,42] with the aim of eliciting in-depth descriptions of the current situation, allowing insight into the thoughts and interpretations activated by the items in BRIEF-A. During the interviews, we aimed for an upper time limit of one hour. Since this meant most interviews did not cover all the 20 selected BRIEF-A items, the order of the items was varied to make sure all of them were covered in at least five interviews.

Data analysis

Interviews were transcribed verbatim. A thematic analysis, based on the method for analysing cognitive interviews described by Miller et al. [43], was conducted by the first author in cooperation with the co-authors. All interview segments regarding each BRIEF-A item were read through several times. Data were condensed by writing summaries of each participants' response, keeping content that was relevant to the research questions. Again, the summaries were read through, and a general summary was written for each item. Since the purpose was to reach a general understanding of how the participants went about responding to the BRIEF-A, the interview analysis was continued one step further than in the description provided by Miller et al. [43], by looking for common themes across the items. The emerging themes were regularly checked back against the original interview texts during the analysis process in order to ensure the themes' close connection to the interview data. The number of participants was not decided *a priori* but was based on the analysis that partly ran parallel with data collection. The analysis started after eight interviews, and a preliminary set of themes was developed. Another two interviews were conducted and added to the analysis, which further strengthened the preliminary themes and did not lead to the emergence of new themes. Yet another interview was added with the same result. It was thus concluded that a sufficient level of inductive thematic saturation [44] had been reached and data collection was terminated. Thus, 11 interviews were conducted and

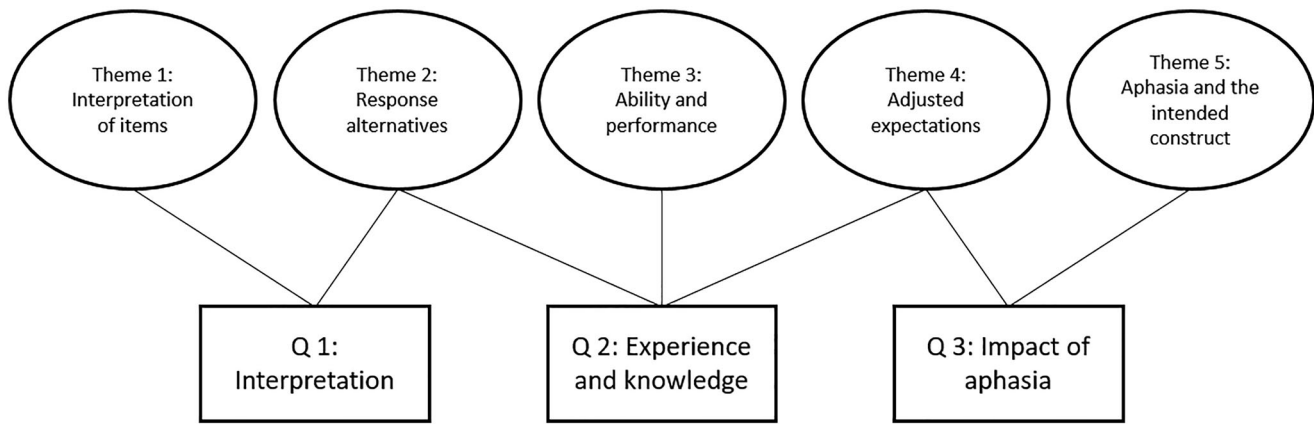


Figure 1. Themes and their relations to the research questions.

analysed. The themes were further refined through continued analysis until the final theme structure and labelling were reached.

Results

The length of the interviews varied between 17 and 76 min, (median 57 min). The second shortest interview was 49 min, thus the 17-min interview (with participant 11) was by far the shortest. The reason for this was that the participant, who was a formal caregiver, wished the interview to be conducted during her working hours, which meant there was a limited time at our disposal. It was nonetheless considered important to include this participant since recruitment of formal caregivers was somewhat difficult.

The data analysis resulted in five themes related to the three research questions (Figure 1). The first theme illustrates issues in response to research question 1. If the content of a questionnaire item is interpreted very differently by different informants, this compromises the validity of the item. There were two main ways in which this could happen; through variation in the interpretation of specific concepts used in the items or through the entire wording of the item leading the respondents' thoughts in an unintended direction.

Theme 2 concerns the response alternatives and is connected to the first two research questions, as shown in Figure 1. The interviews revealed a variation in interpretations of the response alternatives, related to research question 1. We also found evidence that the informants, using their experience and knowledge about the PWSA, were reluctant to consign them into categories perceived to be somewhat crude, thus relating this theme to research question 2. Theme 3 is firmly connected to research question 2, dealing with the discrepancy between the ability and the actual performance of the PWSA. Theme 4, revealing the informants' adjusted expectations on the PWSA, is mainly connected to the question about experience and knowledge used when responding to the questionnaire. But there is also a relation to the third research question, since the adjustment of expectations is likely to come about partly as an effect of the aphasia. The final theme illustrates the challenge in distinguishing between executive function and language in this population, thus responding to research question 3. Below follows a closer description of the themes, illustrated by quotes which have been translated from Swedish. BRIEF-A items will be referred to by their numbers in the questionnaire (with the exception of the three items presented in Table 1). Due to copyright regulations and protection of test integrity no specific description of item contents is allowed. Thus, very general indications of

the items' target behaviours will be provided in the text. The interested reader is advised to refer to a copy of the BRIEF-A.

Theme 1. Varying interpretation of items: "what can be an activity or a task?"

This theme captures the large variation in the informants' interpretation of the items. Some words occurring in BRIEF-A items are quite general (e.g., tasks, things, activities) and for the most part no examples of what they may refer to are given (with a few exceptions, such as item 56 below). This allows informants the important freedom to interpret the item in relation to the everyday life of the particular person he/she is rating. However, it also potentially makes comparisons across individuals or to norms precarious.

Sometimes, informants clearly expressed insecurity about how to interpret certain words.

Oh, I'll have to think, what can be an activity or a task? (1)

But more commonly, the informants did not express any particular insecurity about the interpretation, but the variation became clear when comparing between informants. For item 11 "Has trouble with jobs or tasks that have more than one step" from the scale Working Memory, three different interpretations of the central concept (tasks that have more than one step) emerged:

1. Performing an activity that consists of a sequence of steps.

The dishwasher, that's also different steps; first, he has to empty the dishwasher, then he puts everything up on the bench and then he closes the door of the dishwasher, and then he takes care of that, and when he has taken care of that he opens the door again and puts in the dirty dishes (1)

2. Performing different, unrelated activities, one after another.

If, for example, she cooks and then tidies up and then perhaps she puts on the computer and sits down to look at that a bit and then she moves on to the sitting room. (10)

3. Having the stamina to take part in several activities during one day.

And also he doesn't cope, he's old you know plus now he has this so, it has come to that he can only do one thing, if he's going to the hairdresser then that's it, if he's going to the aphasia group then that's it, so it's like he can only do one thing (8)

The variations in interpretation meant different informants in reality responded to different questions. The item above belongs to the clinical scale Working Memory, but it is only with the first of the interpretations that working memory might be reflected.

A similar issue arose with item 8 from the Shift scale. Again, within our sample there were at least three distinctly different interpretations of what ability the item asked about:

1. Stopping an ongoing activity/task
2. Completing an activity/task and then moving on to another
3. Changing back and forth between different tasks (e.g., being interrupted and then being able to pick up the activity again)

In our material, behaviours according to the first and third interpretation above were more frequently reported as being difficult than the second interpretation. Thus, the informants' different interpretations had a marked impact on the responses. Some informants clearly expressed this, as illustrated in the citation below:

Let's take writing his signature again; it takes a lot of concentration for him to be able to write his name now, so if someone starts talking and perhaps asks him something in the middle of it, then he can't continue writing ... but that's one interpretation, ... I guess one could also interpret it as if he has done one thing and is starting on something else without being interrupted, in that case it's not a problem. (2)

As described previously, the items in BRIEF-A belong to 9 different clinical scales, each intended to capture a different construct, or aspect of executive function. The clinical scales are not indicated on the questionnaire and are thus unknown to the informant. Whether a response actually reflects the intended construct is a matter of what the informant thinks the item enquires about. Item 56 "Has trouble remembering things, even for a few minutes (such as directions, phone numbers)", from the scale Working Memory, frequently caused the informants to associate to memory in general. For someone with a bit of knowledge about cognition and executive function, the phrase "even for a few minutes" in combination with the provided examples indicate that the target for this item is working memory. However, without such knowledge, one might just as well think about long term memory instead:

Those screws ... it was at least six months since we bought them and he remembered that (7)

There were several instances where the entire wording of an item seemed to direct the informants' thoughts in unintended directions. For example, some items in BRIEF-A (e.g., 57 and 64) start with the phrase "People say that...", and then describes a behaviour considered being a sign of executive dysfunction. This allows for the informants' experience of other peoples' perception of the person to have some impact on the ratings. In these items, the focus is still supposed to be the behaviour described in the item, targeting for example Emotional Control or Self-monitoring. The informants expressed no difficulty in understanding these items. However, due to the initiating phrase, the targeted behaviours were frequently not considered by the informants. Instead they focused on whether people say, or would say, such things about the PWSA:

I don't think anyone would say to me that he had done something strange; I don't think so, out of consideration maybe you just don't do that (3)

Friends that come to visit him often, they would never say that, I don't think so (5)

Such responses obviously do not contain any information about the executive functioning of the PWSA.

Theme 2. Varying use of response alternatives: "you kind of end up in the middle"

The response alternatives in BRIEF-A are *never*, *sometimes* and *often* and they were interpreted quite differently between informants. To some, *never* seemed to be an absolute:

I wouldn't say often, but never, I don't know, I don't dare to say ... I don't think it is a problem for her but, well, I will have to say sometimes on that one because I don't really know for sure (10)

Others rather thought of it as "very seldom":

Well never or sometimes, oh that's difficult, something in between there ... no, I would think of the most common everyday things and he manages those, so I would say never (3)

In the general instructions of the questionnaire, the informant is specifically asked to respond to all items, and there is no "don't know" or "not applicable" alternative. Clearly, some informants used the *sometimes* response as an equivalent of "don't know".

When I kind of didn't know, I responded sometimes. (5)

There was also evidence that the informants, with their knowledge about and relation to the PWSA, were reluctant to use the alternative *often*, because they wanted to avoid being too harsh in their judgments. Several informants mentioned that they would have preferred a broader scale with more steps. The three-step scale seems so have been perceived as too crude. There was a tendency to "round down", especially in the choice between *sometimes* and *often*.

I thought that I would rather have a scale from one to ten, and that sometimes would have been five, then I could have put like six or seven, you see (4)

You kind of end up in the middle ... one should be able to choose a response that isn't as clear, but still leans more to one side (2)

Theme 3. A mixture of ability and actual performance: "when it is a routine he is very pedantic"

From the instructions of the BRIEF-A and the wordings of the items, it was not always clear to the informants whether they were being asked to think about the ability or the actual performance of the PWSA. That is, what it would be like if the PWSA had to do things on his/her own, versus how it actually worked in everyday life with existing support. Some informants were relatively consistent during the interview, while some seemed to waver back and forth between these two viewpoints. Perhaps the most obvious example of this discrepancy concerned item 43 "Makes decisions that gets him/her into trouble (legally, financially, socially)", from the scale Inhibition. Two informants responded at opposite ends of the scale though describing similar situations; they both took care of the finances of the PWSA, who could no longer manage to pay bills or make decisions about financial issues on his/her own. In one case, the informant decided on the response *often* because of earlier experience of problems when the PWSA did take care of his finances himself, thus basing the response on the PWSA's ability:

Informant: Often, I've had to take charge many times, he can't have electronic ID on his cell phone and I have complete power of attorney at the bank.

Interviewer: ... if I understand you correctly, there would often be problems if he did have access to his finances.

Informant: Yes.

Interviewer: But you have found a solution so that doesn't happen.

Informant: Yes, exactly. (1)

In the other case, the informant's response was *never*, based on the fact that with existing support no problems occurred.

He has handed over all those things to me, so he doesn't make those decisions (8)

In this theme, the importance of routines also became apparent, and caused some confusion for the informants, since an activity that is routine caused no problems, while new activities were difficult. The solution for the informants was often to respond *sometimes*, which did not reveal correct information about either the routine or new activities. The example below concerns item 2.

I would choose sometimes ... when it is a routine then he is very pedantic. But, if it is something new then he can do it [make mistakes] often, so I would choose sometimes (1)

Theme 4. Expectations adjusted to the perceived ability: "if you consider what is possible with his, you know, disease"

We found several indications that the informants had adjusted their expectations on the PWSA, based on their disabilities. This means they judged the functioning of the PWSA against what they thought was reasonable to expect, given the circumstances, which of course varied between participants. One example is item 65, which targets the construct Organisation of Material. Regarding this item, the response *never* could, according to one informant's description, mean that the PWSA was able to find an everyday, frequently used object (his shaver) when explicitly asked to fetch it:

Because sometimes I ask him, well, to fetch something, for example if I'm going to shave him, he goes straight to fetch it (5)

Another informant also responded *never* to this same question, but based on experience of a more complex task; the PWSA could find specific documents about legal and financial issues among his many papers.

One wonders how he can find anything in there, with all those papers, but if I ask him about something, he goes in there and then he comes out and shows it, so there is order in his mess (3)

Hence, both informants could think of situations when the PWSA managed to locate specific items and thus both responded that there were *never* any problems, but these responses seem to reflect rather different levels of functioning. The PWSA described in the first citation would probably not be able to perform the task described in the second citation.

Similarly, regarding item 25 targeting Initiate, the PWSA seemed to be judged against widely varying standards.

Well, for example, he doesn't exercise, he has trouble starting to do that (1)

He gets up and sits on the balcony and ... sometimes, he can bring the pitcher ... he puts it on the table outside and goes back in to fetch a glass, so in that way he can get going (5)

During the interviews, the informants were sometimes asked what kind of activity they would think about if they were asked about themselves, a technique that revealed that they applied different standards to themselves and the PWSA. A particularly striking example concerns item 66 from the scale Planning/Organisation. One informant, when asked what kind of activity she thought about when she responded to the question about the PWSA, responded:

Well, it depends. If it is playing music, for example, he does that very often actually ... that works very well ... the TV too, that's also something that he manages himself (5)

When asked to give an example related to herself, the response was quite different:

Well, for example, let's say I'm going to clean the house on Thursday, well then I have to make sure I have detergent at home and, you

know, things like that I guess, kind of making sure I have the stuff I need for the things I'm going to do (5)

Some informants quite explicitly stated that they did not judge their PWSA against the standards they would normally use.

And you can't think of too difficult things (9)

Well instructions about things, it has to be small things if it's going to work; it takes some reminding too probably (6)

No, he has good memory, if you consider what is possible with his, you know, disease (7)

Theme 5. The complexity of distinguishing executive dysfunction from language and communication problems: "practical tasks and language tasks, it's a very big difference"

It became evident that it is challenging to distinguish language from executive function in some of the investigated items. Also, in some cases it became evident from the informants' descriptions that they could and did make this distinction, but it could not be captured by the items of BRIEF-A. One example is item 56 "Has trouble remembering things, even for a few minutes (such as directions, phone numbers)", from the scale Working Memory. All informants explained that remembering numbers was out of the question. Several chose to respond that the PWSA never had problems remembering things, having clearly decided to base their responses on other examples than phone numbers or the like.

With the same item, it also occurred that the response was based primarily on the ability to talk about what was remembered, and the informant expressed insecurity about distinguishing between that and the working memory itself.

And then, he tries to tell me how he did it and all that, but "shucks" he says finally, because he doesn't remember the names of things ... but perhaps that's mostly the speech, maybe he knows but maybe he can't explain (3)

For some items (e.g., item 8 and 18 from the scales Shift and Task Monitoring, respectively) informants described that the PWSA experienced problems if the tasks were somehow language dependent, whereas there was no problem with more practical tasks, resulting in the response *sometimes*.

Well, you kind of mix together a response so you get sometimes, because I think that, well sometimes with these tasks and almost always with those tasks and so you choose sometimes ... because, you know, practical tasks and language tasks, it's a very big difference (4)

For several items, for example regarding activities that require several steps, remembering things for a short while or not managing certain tasks, the informants were quite clear that it depended on whether the PWSA had understood the instruction for the task. However, their insights about this systematic variation between linguistic and other cognitive functioning was not possible to convey through the questionnaire responses.

I think it's mainly the instruction, if he can think for himself and kind of see the logic in what he is expected to do, well, then it works (2)

I notice that he makes mistakes because he didn't understand the instruction (4)

Discussion

The results of this study indicate that there is considerable variation regarding what information BRIEF-A informant reports are actually based upon. The analysis shows that central concepts were interpreted in various ways and the way the informants

understood an item sometimes corrupted its relation to the intended construct. There was some variation in how the response alternatives were used and a tendency to “round down”, especially when the informant hesitated between *sometimes* and *often*. Further, the informants seemed to waver between basing their responses on the PWSAs’ ability or their actual performance and seemed to have adjusted their expectations or demands and thus responded in relation to a separate standard for the PWSA. The language and communication problems of PWSA affected the validity of some of the items.

The items of BRIEF-A contain many concepts that are not clearly specified, such as “tasks”, “activity” and “have trouble”. This is a challenge when designing questionnaires. If the concepts are too narrowly exemplified, some respondents will not think the item applies to them. If they are too unspecified there is a risk that respondents form very different interpretations. For instance, what does it really mean to be able to plan and organise? Putting on a cd, planning to clean the house, or make arrangements for a big birthday party? The interpretation of the item can determine whether the informant reports that there are never any problems or often. This is important to bear in mind when interpreting data from questionnaires like BRIEF-A.

One of the proposed assets of rating instruments of executive function is ecological validity; that is the measurement’s relationship to actual behavioural outcomes in the real world [45]. Since the items of BRIEF-A explicitly ask about behaviour in everyday situations during the last month, it certainly can be said to have high face validity. This is not, however, necessarily equivalent to ecological validity. As evident from our interview data, the behaviours rated by the informants in BRIEF-A are indeed behaviours from the everyday lives of the PWSA. In that sense, the responses are ecologically valid. Whether they are an ecologically valid measure of *executive function* is a somewhat different matter. The interview data provide numerous examples of responses that are based on behaviours and facts that are not related to the intended target construct. That is, the responses simply do not contain information about executive function. Also, since the responses are mainly based on recurring everyday situations, it is questionable how well the results can be assumed to predict functioning in new situations (which is also an aspect of ecological validity). In fact, it has been suggested that behaviours that are automatic or routine are, by definition, not a reflection of executive function [46]. This is reinforced by our interview data, where we repeatedly encountered descriptions of there being no problems in routine situations but more so in new situations.

Although using a different analysis method than Hill et al. [37], we found indications of some of the same problems as they described in their study of self-ratings of memory function. For instance, Hill et al. [37] found that many of the items they investigated required a generalisation of performance across times and situations, whereas their participants experienced variations in performance. This is very much in line with our findings of the aforementioned variation between routine versus new situations, or linguistic versus practical tasks. It could be argued that responding to questionnaires always requires estimation of an average. But in these cases, the average fails to capture important information about systematic differences in performance between different kinds of situations.

Another problem discovered by Hill et al. [37] was the use of an unspecified reference point. In our analysis, theme 3 and 4 both touch on this issue. The instruction in BRIEF-A is to “indicate whether he or she has had any problems with these behaviours over the past month” [28]. It was not clear to all informants

whether this meant just thinking about situations and behaviours as they actually play out with the existing adaptations and support, or if it meant thinking about how the PWSA would hypothetically behave if left to his/her own devices. These are of course in many cases two very different things. The adjusted expectations of theme 4 can be hypothesised to be related to the fact that the concept “problems” is not well defined. If you have adjusted your expectations you are likely to respond based on what actually causes problems under the current conditions, while others might think about what would have been considered problems before the stroke. Again, those are two very different things. The phenomenon of significant others adjusting their expectations is probably a natural and, in many cases, functional adaptation to the impact of the stroke on everyday life. However, there is evidence that aphasia and other communication disorders have specific, negative impact on the perception of an individual [47,48], which might indicate that aphasia can be a particularly strong factor for lowering of expectations. In this there is an inherent risk of underestimation of the competence of the PWSA [49]. But low expectations might also lead to under-reporting of executive dysfunctions that are important for rehabilitation staff to know about, and which might be mitigated.

The final theme that emerged from our analysis made clear that the aphasia, and the resulting communication disorders, indeed do interfere with the constructs targeted in BRIEF-A. As previously mentioned, participants described discrepancies between linguistically demanding and more practical situations. Thus, they frequently did have a notion about a distinction between problems caused by executive dysfunction (i.e., making mistakes due to insufficient task monitoring) or by the aphasia (making mistakes due to not understanding instructions). However, this distinction could not be unambiguously captured by the BRIEF-A, since there might very well be problems with the behaviours described in the items, only not for the reason assumed in the BRIEF-A. For example, the aphasia seemed to be a confounding factor in several items from the scale Working Memory, since the PWSA due to the aphasia could not be expected to remember, for instance, phone numbers or instructions in several steps.

There are both theoretical and empirical reasons to believe that executive function, as measured with standardised neuropsychological tests, is important to functional communication in severe aphasia [17–19]. However, the aspects of executive function that are important to functional communication seem not to be captured by BRIEF-A informant ratings [32]. Still, one would have hoped that informant ratings could be a way of assessing executive function in this population without the aphasia acting as a confounding factor. Alas, according to our results, this is not the case. The aphasia does interfere, and for some of the BRIEF-A items it is impossible to know whether the response reflects the aphasia or some aspect of executive function, or indeed some other cognitive ability.

One of the aims of this study was to deepen the understanding of the relationship between neuropsychological tests and informant ratings of executive function, since these have repeatedly been shown to have weak relations. Our results indicate that, at least considering the population of PWSA, there is a large variation in the extent to which the BRIEF-A responses actually reflect the intended executive functions. It seems logical that this could be part of the explanation for the weak correlations between informant ratings and the more strictly controlled neuropsychological test results.

As can be glimpsed from the citations presented in this paper, interviews based on the BRIEF-A questionnaire can provide detailed, rich and clinically relevant information about the everyday lives and functioning of the PWSA and the informants. Had these interviews been conducted within a clinical setting, with a clinical purpose, they would have offered opportunity for valuable discussions, explanations and understanding, and could have generated ideas for feasible interventions. For example, learning that a person struggles with new, unfamiliar activities but has the capacity to manage if the activities are worked into being routines, or learning that a person tends to get stuck in linguistically demanding tasks but not in practical ones, can be highly relevant to rehabilitation staff. Such information is important when planning intervention aiming to help the patient and family to find ways to manage everyday challenges. Our results indicate that pure quantitative results from BRIEF-A should be interpreted with great caution, since it is uncertain what they represent. But if completion of the questionnaire is followed by a clinical interview, the instrument may contribute a lot to the gathering of relevant information about the individual in a clinical context.

PWSA are, apart from the aphasia, no different than other stroke patients and some of our results mirror the findings of Hill et al. [37], who were investigating healthy older adults' self-ratings of memory. Based on these facts, it is likely that similar issues could be expected in other self- and informant reports of cognitive functioning in similar populations.

Limitations and further research

This is a small study, with eleven informants and covering only 20 of the 75 BRIEF-A items. Data collection was continued until no new themes emerged, but of course it is impossible to know if further interviews could have generated new insights. To our knowledge, BRIEF-A has not previously been investigated with cognitive interviewing methodology. It does not seem to be a common way to validate other cognitive self or informant reports either, with the exception of the study by Hill et al. [37]. The present study can be considered a first attempt to use cognitive interviewing in this context, and our results as an indication of issues that need further investigation.

The use of verbal probes in cognitive interviewing has been criticised by some authors, with the argument that sometimes the process of responding to a question is not conscious, you "just know" the answer. Forcing the informants to make it conscious by asking probe questions might then not give an accurate description of the response process [41]. However, it could also be argued that to be able to respond to, for instance, the question "Has trouble with jobs or tasks that have more than one step.", the informants have to decide what they think "jobs" and "tasks" mean. Their spontaneous response might be based on many different situations, but asking them to give examples appears to give an indication of the width of the variation.

We do not present suggestions to solve the problems we found regarding BRIEF-A. Such suggestions would demand a larger study, covering all items and applying a different analytic methodology. Our aim was not to describe problems in individual items, but rather to enhance the understanding of more general processes involved in informant ratings of executive function in this population.

Further research is needed to learn more about how to best gain information about the executive (and other cognitive) functions of PWSA, and how to make it possible for this population to be directly involved in their own assessments. Since both

traditional testing and self-report is challenging, a path worth exploring might be structured observation of activity. Several such assessment instruments exist, they are however usually unsuited to PWSA due to linguistic demands and thus need to be adapted. Structured observation could be combined with questions to the PWSA during or directly after the activity. With a concrete, standardised activity to talk about, supporting material (such as pictures, visual scales) could more easily be developed and used to allow the PWSA to give his/her own view of the issue at hand.

Conclusions

BRIEF-A informant ratings of executive function in PWSA should be interpreted with great caution. Varying interpretations of the items compromise their relation to the intended constructs and frequently the responses may not reflect executive function. Informant ratings cannot be considered to be free from the problem of aphasia interfering with the assessment of executive function in PWSA, since the linguistic and communication problems interfere with the constructs of the questionnaire. Preferably, the completion of BRIEF-A should be followed by interviews about the responses. This could provide detailed, rich and clinically relevant information about the everyday life and functioning of the PWSA.

Notes

1. The items are reproduced by special permission of the Publisher, Psychological Assessment Resources, Inc. (PAR), 16204 North Florida Avenue, Lutz, Florida 33549, from the Behavior Rating Inventory of Executive Function-Adult Version by Robert M. Roth, PhD, Peter K. Isquith, PhD and Gerard A. Gioia, PhD, Copyright 1996, 1998, 2001, 2003, 2004, 2005 by PAR. Further reproduction is prohibited without permission from PAR.

Acknowledgments

The authors thank the participants for their time, effort and patience as well as the speech-language pathologists, The Aphasia Association and Forsa folk high school, who assisted in recruitment.

Disclosure statement

The authors report no conflict of interest.

Funding

This work was made possible by generous donation from Solveig and Arne Morin and a grant from The Swedish Stroke Association.

ORCID

Camilla Olsson  <http://orcid.org/0000-0002-7703-2188>
 Patrik Arvidsson  <http://orcid.org/0000-0002-3067-2794>
 Monica Blom Johansson  <http://orcid.org/0000-0002-6095-6130>

References

- [1] Lezak MD, Howieson DB, Bigler ED, et al. Executive functions. In: Lezak MD, Howieson DB, Bigler ED, et al. editors. *Neuropsychological assessment*. 5th ed. New York (NY): Oxford university press; 2012. p. 666–711.
- [2] Friedman NP, Miyake A. Unity and diversity of executive functions: individual differences as a window on cognitive structure. *Cortex*. 2017;86:186–204.
- [3] Suchy Y. *Executive functioning: a comprehensive guide for clinical practice*. New York (NY): Oxford University Press; 2015.
- [4] Schumacher R, Halai AD, Lambon Ralph MA. Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. *Brain*. 2019;142:3202–3215.
- [5] Constantinidou F, Wertheimer JC, Tsanadis J, et al. Assessment of executive functioning in brain injury: collaboration between speech-language pathology and neuropsychology for an integrative neuropsychological perspective. *Brain Inj*. 2012;26:1549–1563.
- [6] Burgess PW, Stuss DT. Fifty years of prefrontal cortex research: impact on assessment. *J Int Neuropsychol Soc*. 2017;23:755–767.
- [7] Suchy Y, Ziemnik RE, Niermeyer MA. Assessment of executive functions in clinical settings. In: Goldberg E, editor. *Executive functions in health and disease*. San Diego(CA): Elsevier Science & Technology; 2017.
- [8] Leśniak M, Bak T, Czepiel W, et al. Frequency and prognostic value of cognitive disorders in stroke patients. *Dement Geriatr Cogn Disord*. 2008;26:356–363.
- [9] Shea-Shumsky NB, Schoeneberger S, Grigsby J. Executive functioning as a predictor of stroke rehabilitation outcomes. *Clinical Neuropsychol*. 2019;33:854–872.
- [10] Lewis MW, Babbage DR, Leathem JM. Assessing executive performance during cognitive rehabilitation. *Neuropsychol Rehabil*. 2011;21:145–163.
- [11] Lipskaya-Velikovsky L, Zeilig G, Weingarden H, et al. Executive functioning and daily living of individuals with chronic stroke: measurement and implications. *Int J Rehabil Res*. 2018;41:122–127.
- [12] Papathanasiou I, Coppens P. *Aphasia and related neurogenic communication disorders*. 2nd ed. Burlington (MA): Jones & Bartlett Learning; 2017.
- [13] Ardila A, Rosselli M. Cognitive rehabilitation of acquired calculation disturbances. *Behav Neurol*. 2019;2019:3151092.
- [14] Basso A, Caporali A, Faglioni P. Spontaneous recovery from acalculia. *J Int Neuropsychol Soc*. 2005;11:99–107.
- [15] Nicholas M, Hunsaker E, Guarino AJ. The relation between language, non-verbal cognition and quality of life in people with aphasia. *Aphasiology*. 2017;31:688–702.
- [16] Simic T, Rochon E, Greco E, et al. Baseline executive control ability and its relationship to language therapy improvements in post-stroke aphasia: a systematic review. *Neuropsychol Rehabil*. 2019;29:395–439.
- [17] Olsson C, Arvidsson P, Blom Johansson M. Relations between executive function, language, and functional communication in severe aphasia. *Aphasiology*. 2019;33:821–845.
- [18] Fridriksson J, Nettles C, Davis M, et al. Functional communication and executive function in aphasia. *Clin Linguist Phon*. 2006;20:401–410.
- [19] Light J, McNaughton D. Communicative competence for individuals who require augmentative and alternative communication: a new definition for a new era of communication? *Augment Altern Commun*. 2014;30:1–18.
- [20] Suchy Y, Niermeyer MA, Ziemnik RE. Assessment of executive functions in research. In: Goldberg E, editor. *Executive functions in health and disease*. San Diego (CA): Elsevier Science & Technology; 2017. p. 197–216.
- [21] Egeland J, Løvstad M, Norup A, et al. Questionnaire use among Nordic neuropsychologists: shift from assessing personality to checking ecological validity of neuropsychological assessments? *Prof Psychol Res Pr*. 2017;48:227–235.
- [22] Egeland J, Lovstad M, Norup A, et al. Following international trends while subject to past traditions: neuropsychological test use in the Nordic countries. *Clin Neuropsychol*. 2016;30:1479–1500.
- [23] Toplak ME, West RF, Stanovich KE. Practitioner review: do performance-based measures and ratings of executive function assess the same construct?. *J Child Psychol Psychiatry*. 2013;54:131–143.
- [24] Keil K, Kaszniak AW. Examining executive function in individuals with brain injury: a review. *Aphasiology*. 2002;16:305–335.
- [25] Lovstad M, Funderud I, Endestad T, et al. Executive functions after orbital or lateral prefrontal lesions: neuropsychological profiles and self-reported executive functions in everyday living. *Brain Inj*. 2012;26:1586–1598.
- [26] Stanovich KE. *What intelligence tests miss: the psychology of rational thought*. New Haven (CT): Yale University Press; 2009.
- [27] Roussel M, Martinaud O, Henon H, et al.; GREFEX Study Group. The behavioral and cognitive executive disorders of stroke: the GREFEX Study. *PLoS One*. 2016;11:e0147602.
- [28] Roth R, Isquith P, Gioia G. *Behavior rating inventory of executive function - adult version*. Lutz (FL): Psychological Assessment Resources; 2005.
- [29] Oczkowski C, O'Donnell M. Reliability of proxy respondents for patients with stroke: a systematic review. *J Stroke Cerebrovasc Dis*. 2010;19:410–416.
- [30] Hilari K, Owen S, Farrelly SJ. Proxy and self-report agreement on the Stroke and Aphasia Quality of Life Scale-39. *J Neurol Neurosurg Psychiatry*. 2007;78:1072–1075.
- [31] Hilari K, Cruice M, Sorin-Peters R, et al. Quality of life in aphasia: state of the art. *Folia Phoniater Logop*. 2015;67:114–118.
- [32] Olsson C, Arvidsson P, Blom Johansson M. Measuring executive function in people with severe aphasia: comparing neuropsychological tests and informant ratings. *NeuroRehabilitation*. 2020;46:299–310.
- [33] Wright J, Moghaddam N, Dawson DL. Cognitive interviewing in patient-reported outcome measures: a systematic review of methodological processes. *Qual Psychol*. 2019.
- [34] Miller K, Chepp V, Padilla JL, et al. *Cognitive interviewing methodology*. Hoboken (NJ): John Wiley & Sons; 2014.
- [35] Boeije H, Willis G. *The Cognitive Interviewing Reporting Framework (CIRF): towards the harmonization of cognitive testing reports*. *Methodology*. 2013;9:87–95.
- [36] Tourangeau R, et al. Cognitive science in survey methods. In: Jabine T, Straf ML, Tanur JM, editors. *Cognitive aspects of survey design: building a bridge between disciplines*. Washington (DC): National Academy Press; 1984. p. 73–100.
- [37] Hill NL, Mogle J, Whitaker EB, et al. Sources of response bias in cognitive self-report items: “which memory are you talking about?” *Gerontologist*. 2018;59:912–924.

- [38] Willis G. Analysis of the cognitive interview in questionnaire design. New York (NY): Oxford university press; 2015.
- [39] Ridolfo H, Schoua-Glusberg A. Analyzing cognitive interview data using the constant comparative method of analysis to understand cross-cultural patterns in survey data. *Field Methods*. 2011;23:420–438.
- [40] Chepp V, Gray C, et al. Foundations and new directions. In: Miller K, Willson S, Chepp V, editors. *Cognitive interviewing methodology*. Hoboken (NJ): John Wiley & Sons; 2014. p. 7–14.
- [41] Conrad FG, Blair J. Sources of error in cognitive interviews. *Public Opin Q*. 2009;73:32–55.
- [42] Willis G. *Cognitive interviewing: a tool for improving questionnaire design*. Thousand Oaks (CA): Sage Publications; 2005.
- [43] Miller K, Willson S, Chepp V, et al. Analysis. In: Miller K, Willson S, Chepp V, et al. editors. *Cognitive interviewing methodology*. Hoboken (NJ): John Wiley & Sons; 2014. p. 35–50.
- [44] Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*. 2018;52:1893–1907.
- [45] Ziemnik RE, Suchy Y. Ecological validity of performance-based measures of executive functions: is face validity necessary for prediction of daily functioning? *Psychol Assess*. 2019;31:1307–1318.
- [46] Suchy Y. Executive functioning: overview, assessment, and research issues for non-neuropsychologists. *Ann Behav Med*. 2009;37:106–116.
- [47] Byeon H, Koh HW. Health science students' perceptions of motor and sensory aphasia caused by stroke. *J Phys Ther Sci*. 2016;28:1772–1774.
- [48] Allard ER, Williams DF. Listeners' perceptions of speech and language disorders. *J Commun Disord*. 2008;41:108–123.
- [49] Kagan A. Supported conversation for adults with aphasia: methods and resources for training conversation partners. *Aphasiology*. 1998;12:816–830.