This is a repository copy of Variation and interactional non-standardization in neuropsychological tests: the case of the Addenbrooke’s Cognitive Examination.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/150040/

Version: Published Version

Article:
Jones, D., Drew, P., Jackson, C. et al. (1 more author) (2019) Variation and interactional non-standardization in neuropsychological tests: the case of the Addenbrooke’s Cognitive Examination. Qualitative Health Research. ISSN 1049-7323

https://doi.org/10.1177/1049732319873052

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don’t have to license any derivative works on the same terms. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Variation and Interactional Non-Standardization in Neuropsychological Tests: The Case of the Addenbrooke’s Cognitive Examination

Danielle Jones1, Ray Wilkinson2, Clare Jackson3, and Paul Drew3

Abstract
The Addenbrooke’s Cognitive Examination (ACE-111) is a neuropsychological test used in clinical practice to inform a dementia diagnosis. The ACE-111 relies on standardized administration so that patients’ scores can be interpreted by comparison with normative scores. The test is delivered and responded to in interaction between clinicians and patients, which places talk-in-interaction at the heart of its administration. In this article, conversation analysis (CA) is used to investigate how the ACE-111 is delivered in clinical practice. Based on analysis of 40 video/audio-recorded memory clinic consultations in which the ACE-111 was used, we have found that administrative standardization is rarely achieved in practice. There was evidence of both (a) interactional variation in the way the clinicians introduce the test and (b) interactional non-standardization during its implementation. We show that variation and interactional non-standardization have implications for patients’ understanding and how they might respond to particular questions.

Keywords
conversation analysis; communication; Addenbrooke’s Cognitive Examination; standardization; administration; qualitative; United Kingdom

Introduction
Amid the challenges of meeting the needs of an increasingly aging population (Steptoe, Breeze, Banks, & Nazroo, 2013), in 2012, the U.K. government set the first national target to increase diagnostic rates for dementia (Department of Health, 2012). In 2012, only 42% of people living with dementia (in United Kingdom) had received a formal diagnosis, with the result that almost half of this population were not accessing appropriate social and health care at a time when it might be most clinically beneficial (Cullen, Neill, Evans, Coen, & Lawlor, 2007; Mukadam, Cooper, Kherani, & Livingston, 2015). Latest figures show that diagnostic rates have risen to 68% (Alzheimer’s Research UK, 2018), and the current government remains committed to further increasing the quality and consistency of dementia diagnosis, care, and awareness (Department of Health, 2015). Accurate and timely diagnosis remains central to achieving social and health policy aims both in the United Kingdom and elsewhere (Ballard, 2015; Borson et al., 2013).

Ballard (2015) suggested that measuring cognitive function is one of the most important assessments clinicians make, particularly within geriatric medicine, as assessments play a key role in determining a dementia diagnosis (Cullen et al., 2007; Larner, 2017; National Institute for Health and Care Excellence [NICE], 2018). Cognitive assessments cover a broad range of activities, take place in a number of settings (including primary care, specialist memory clinics, acute care, care homes, and in the community), and are administered for a variety of reasons (including screening, diagnosing, and measuring change; Ballard, 2015). The NICE (2018) recommends that practitioners should use validated brief structured cognitive instruments during initial assessments within non-specialist settings, for example, the six-item cognitive impairment test (6CIT; Brooke & Bullock, 1999) for those with suspected dementia. There is also a wide range of assessment tools designed to measure different aspects of...
functionality within specialist memory services, for example, Addenbrooke’s Cognitive Examination (ACE-111) and Montreal Cognitive Assessment (MOCA). Guidance designed to assist clinicians in identifying and appropriately using these tools (e.g., Ballard, 2015) often incorrectly assumes that practitioners have a high level of specialist clinical knowledge about measures of cognitive functioning, as well as how to administer and interpret the tests (Cabana et al., 1999; Murphy et al., 2014). Hence, appropriate training is not always provided (Boise, 2006). Although the NICE (2018) guidelines state that all health and social care professionals involved in diagnosis should be trained in starting and holding “difficult and emotionally challenging conversations,” there is no clarification of the specific training needs for administering cognitive examinations (p. 34). Furthermore, General Practitioners (GPs) involved in the screening and assessment of dementia have reported that they feel ill equipped to use and interpret cognitive assessment tools in accordance with guidelines provided (Smith, 2015). Specialist clinicians working in a U.K. memory service (as a part of this study) report having received no formal training on the administration of the ACE-111, which they used for initial assessments.

ACE-111 is recognized as the most appropriate validated tool for use in specialist memory services in the United Kingdom (Hodges & Larner, 2017). It measures cognitive functioning across five different domains: attention, memory, verbal fluency, language, and visuospatial abilities. It is scored out of 100, with the highest score denoting better cognitive function; the cut-off for dementia is 82–88/100 (Crawford, Whitnall, Robertson, & Evans, 2012). The ACE-111 has been validated against other standard neuropsychological tests and has been shown to be a valid cognitive screening tool for dementia syndromes (Hsieh, Schubert, Hoon, Mooshi, & Hodges, 2013; Matias-Guiu et al., 2017). Furthermore, it has been found to distinguish early-onset dementia from healthy controls with high sensitivity and specificity (Elamin, Holloway, Bak, & Pal, 2016). The ACE-111 relies on the accuracy and consistency of test delivery, so that patients’ scores can be interpreted by comparison with normative scores.

If standard administration procedures are not followed, although results may be informative as to a patient’s maximal residual function, they are not useful in indicating whether their score falls in the normal or pathological range. Scores obtained following non-standard administration procedures are not comparable to norms. (Venneri, 2005, p. 97)

An administration and scoring guide accompanies the assessment tool, and aims to support practitioners’ understanding and delivery of the test. However, within this guide, advice is not consistent; perhaps as a consequence (see the “Analysis” section and as noted above; Smith, 2015), there is significant variance in the implementation of this guidance. Other important elements are missing from the guidance altogether; for example, there is no instruction on how to introduce the test in clinical settings with patients, which creates a condition for variation in delivery. As we demonstrate, this has interactional significance and can lead to patients’ confusion. Furthermore, there is inconsistency in the way that questions are presented; some questions have verbatim instruction, for example, “say to the participant: ‘Now tell me what you remember of that name and address we were repeating at the beginning,’” while other questions are not scripted in this manner, for example, “Ask the participant for the day, date, month, year, season . . .” These latter quasi-scripted questions do not require the practitioner to use the specific wording from the guidance but instead allows for interactional variation, resulting in a lack of administrative or interactional standardization.

In line with most cognitive assessments, the ACE-111 is delivered and responded to in talk-in-interaction (Drew, Raymond, & Weinberg, 2006). Despite the efforts made to ensure the standardization of administration, and hence the validity and reliability (Bentvelzen, Aerts, Seeher, Wesson, & Brodaty, 2017) of such tools, the unavoidable contingencies associated with talk add fundamentally social, and crucially, non-standardized elements to the testing process (Marlaire & Maynard, 1990; Maynard & Marlaire, 1992; Maynard & Turowetz, 2017). This, as we will show, is not without significance.

Data and Method

The data are video recordings made between October 2012 and October 2014 of 105 initial assessment consultations in a specialist neurology-led memory clinic in the United Kingdom. Patients are usually referred to the memory clinic by their GP. These initial assessments typically comprise history-taking, followed by a cognitive examination using a screening tool, and then a brief physical examination. The data were collected as part of a study on patient assessment for differential diagnosis of dementia in the memory clinic (Elsey et al., 2015; Jones et al., 2016; Reuber et al., 2018; the administration of the Addenbrooke’s memory test was not however included in analysis for that study, the results of which are therefore not pertinent here). The current research focuses on the administration of the ACE-111, which the majority of patients (n = 98) undertook. Of these, 92 were recorded (providing a corpus of 23 hours), from which a sample of 40 cases (10 hours) were randomly selected for detailed analysis (given the detailed nature of CA transcription and analysis, a sample of a total corpus is generally taken; in one study related to this, that sample was 30; Reuber et al.,

Marlaire, 1992; Maynard & Turowetz, 2017). This, as we will show, is not without significance.
The administration of the test takes on average 15 minutes (the full initial consultation lasts on average approximately 35 minutes). The interactions in the randomly selected sample were transcribed in detail, according to the conventions used in conversation analysis (CA). The data collection for the study were approved by the NHS Research Ethics Committee (NRES Committee Yorkshire & The Humber—South Yorkshire; Ref 12/YH/0205). Written informed consent was obtained from both patients and clinicians.

CA is increasingly used in research in medical settings (Heritage & Maynard, 2006; Maynard & Heritage, 2005; Robinson & Heritage, 2014; Stivers, 2007) to identify patterns of language and interaction that inform practice (Heritage, Robinson, Elliott, Beckett, & Wilkes, 2007; Heritage & Stivers, 1999; Wilkinson, 2013), medical assessment (Heritage & Stivers, 1999; Monzoni, Duncan, Grünwald, & Reuber, 2011), and diagnosis (Heath, 1992; Maynard, 2017; Peräkylä, 1998). CA’s method is to examine in close detail the various communicative formats used to “deliver” medically relevant actions, such as treatment recommendations (Chappell, Toerien, Jackson, & Reuber, 2018; Stivers et al., 2018) and examination of dementia diagnoses (Dooley, Bass, & McCabe, 2018)—and to examine the varying interactional consequences systematically associated with the different formats involved (Heritage & Robinson, 2006; Heritage et al., 2007).

Previous CA research into the administration of clinical tests has shown that questions that are expected to be asked in a standardized manner are, in fact, recurrently asked in diverging and diverse formats. It has further shown that divergence from the standardized forms can influence test outcomes (Marlaire & Maynard, 1990; Maynard & Turowetz, 2017), which are thereby the collaborative products of the interaction between testers and tested—rather than reflecting the intrinsic quality that is taken to be measured through what should be a neutral instrument. Employing the same CA perspective and method, our inquiry into precisely how the ACE-111 test is delivered focuses primarily on the variation in administration that differs from standard administrative procedures. However, the method also enables us to consider the implications for how patients understand the cognitive task at hand, and the association between certain non-standard question formats and patient responses, including their apparent confusion.

Analysis

Our examination of the interactional accomplishment of the ACE-111 showed, first, interactional variations—when elements of the test do not appear on the guidance, clinicians vary the way they deliver the instrument, for example, in the way they introduce the test (see section “Interactional Variation: Introduction of the Memory Assessment”). This variability is evident between different clinicians, and within each clinician’s individual conduct in any given consultation. Although it is not possible to quantify the frequency of variance in the delivery of ACE-111 questions, we noted that such variations occurred at least once in each of the 40 cases sampled. Second, there is evidence of interactional non-standardization when neurologists deviate from the (quasi) scripted guidance designed to ensure standardized administration (see section “Interactional Non-Standardization: Question Design”). This could again result from inconsistencies within the guidance itself, or originate from individual clinicians’ interactional style, among other things. We discuss the broader implications of these findings for cognitive assessment procedures and their outcomes (see “Discussion” section).

Interactional Variation: Introduction of the Memory Assessment

The first feature which demonstrates variation in administration is the manner in which clinicians transition from the history-taking phase of the consultation to administering the formal memory assessment. During history taking, the clinicians have typically asked a series of questions about the patient’s personal information, what concerns they have about their memory, and requested full descriptions about their competency in performing activities in daily living. The average length of the history-taking phase of the consultation is 19.6 minutes. Direction on how to initiate the transition into “testing” is not included in the guidance and is managed differently by each neurologist and by the same neurologist on different occasions. Here is one example.

Extract 1

01 DOC: Alright. We’ll jus- we’ll run through a few
02 quick questions a- then I’ll examine ya-
03 I’ll just er::,
04 (22.5) {(Doc retrieves test papers from behind him)}
05 DOC: Okay, Erm: (1.0) >Could you jus tell
06 me the< da::y: today:, 
“Alright” (line 01) initiates closure of the history-taking sequence and simultaneously projects a movement to a new activity (Beach, 1995); the clinician then introduces the assessment (for the first time in this consultation), “We’ll jus- We’ll run through a few quick questions a- then I’ll examine ya” (lines 01–02). The “few quick questions” here refers to conducting the ACE-111 and the mention of an examination refers to a short physical examination (important for differential diagnosis; Chui et al., 1992). During the start of this turn, and within the 22.5 second silence (line 04), the clinician turns his back on the patient to gather the assessment paperwork from the cupboard behind him, returns to the desk, picks up his pen, and is still attaching the patient’s information label to the test documents at the point the first test question is asked (line 05). Nonverbal behavior is integral to the accomplishment of transition (Robinson & Stivers, 2001) and this embodied action (detailed above), whereby the clinician physically removes himself from the desk to locate the test papers, indicates a shift in activity. Immediately after the clinician has informed the patient that he will “examine” him, the patient’s facial expression changes markedly—He quickly looks from left up to right, his mouth is pursed and his brow furrowed. This expression appears to embody a negative stance in response to a prior turn, a stance that is recognizable as displaying “confusion” in response to a prior turn, potentially due to the ambiguous and unspecified introduction of a “medical examination,” or confusion arising from the difference between the “few quick questions” the practitioner will now ask as compared with the 21 minutes of questions asked during history-taking. Variations of the introduction by different clinicians appear to have implications for patient’s understanding and uncertainty as displayed within either their embodied or verbal reactions. Here is another variation of this introduction.

**Extract 2**

01 DOC: .hh Okay. Well we’ll move on to doing the:
02    formal memory assessment. Erm (.) the: the:
03    memory assessment tool that I use, (.) was
tested in a variety of different people,
different ages, different educational
9 backgrounds. .h Some of the questions are
tricky, .h Don’t worry if there’s anything
that you ca[n’t do:.]
13 PAT:               [No. be- ] before, oh: a month
14 or two ago when I went with this .h,=

This demonstrates a significant variation in the way the clinician transitions to the testing phase of the consultation. Here the clinician introduces it as a “formal memory assessment” (lines 01–02) and goes into detail specifying its origins (line 04), naming the test (lines 05–06), and how it was developed (lines 06–09), as well as preparing the patient for the types of questions she is about to be asked (lines 09–12). Neither clinician (in Extracts 1 and 2) had prepared the patient with information about the schedule or tasks involved in the full consultation at the beginning of the meeting. The different methods the practitioners use to prepare the patients for the task ahead appear to be consequential for how patients receive and understand the activity they are being asked to complete. The first example appears to engender some confusion, as evident in the patient’s embodied stance; whereas the second patient, upon topicalizing the recollection of a previous test (Extract 3), appears to be fully aware of the expectations for the next phase of the consultation.

**Extract 3 (continuation from Extract 2)**

15 DOC: =Yes.=
16 PAT: =they:, they gave me a memory test to
17   d[o:], ]
18 DOC:  [I th]ink if y[ou,]
19 PAT:                [And] I got thirty-out-of-thirty.=
In other physician–patient encounters, one source of patients’ uncertainty can be this transition from the activity of history-taking to that of the examination (Robinson & Stivers, 2001; Sheer & Cline, 1995). Neuropsychological testing in particular can produce a sense of anxiety and threat (Cahill, Gibb, Bruce, Headon, & Drury, 2008; Cheston, Bender, & Byatt, 2000). Keady and Gilliard (2002) note that patients reported feeling “trapped” or “caught out” by the process of assessment. Similarly, Cahill et al. (2008) report that patients considered assessment processes to be “probing, demoralizing and frightening” (p. 184). These authors suggested that the provision of more detailed information about neuropsychological assessment might be useful in improving the patient experience (Cahill et al., 2008; Keady & Gilliard, 2002). Giving people more information about these unfamiliar events and clearly explaining this transition might reduce this anxiety and uncertainty (Berger, 1986; Berger & Calabrese, 1974), and can work to secure their acceptance of the transition (Levinson, Roter, Mullooly, Dull, & Frankel, 1997).

We are beginning to find that when there is a lack of guidance, there is evidence of interactional variation in the administration during the introduction of the cognitive assessment tool. This variation in interactional style between clinicians affects the patient experience and responses as displayed in both their verbal and non-verbal conduct. However, we have discovered that deviation also occurs when guidelines are more prescriptive.

**Interactional Non-Standardization: Question Design**

There are other points of disparity in delivery styles between clinicians, even when guidance is provided. We have found evidence of interactional non-standardization, when practitioners deviate from both the quasi scripted and the full verbatim guidance designed to ensure standard administration. Furthermore, there is evidence of interactional non-standardization between the clinicians when each one designs aspects of the assessment differently for different patients. One feature of non-standardization is turn design (see Drew, 2012), in this setting how practitioners design the questions they ask as part of the test sequence. During the first set of questions, intended to measure a patient’s cognitive function concerning their attention, clinicians are instructed to “Ask the participant for the day, date, month, year,” and slots for these answers appear in this order on the response sheet. There is, however, no verbatim script for how the questions should be formulated; hence, practitioners use a variety of question designs, which include, “Could you just tell me the day today”: (Extract 3), “What day of the week is it today?” (Extract 4), and others (not shown here) include, “Do you know what day it is?” and “Can you tell me what day it is”—all have different implications for the response (see Curl & Drew, 2008). Furthermore, following the sequence of questions in the order prescribed on the test can pose potential difficulties when responding. The next example illustrates this potential difficulty.
Extract 4

01 DOC: .hh Okay. Erm:. (0.2) >Could you
02 jus- tell me the< day:, today:,
03 (3.0)
04 PAT: A- Uh- It’s Monday,
05 DOC: An- the date,
06 (1.2)
07 PAT: [{° was twenty - six°} ]
08 [Patient gestures backward]
09 (3.7) ((Patient has his eyes closed and
10 is mounting words))
11 PAT: Twentieth of the twelfth.
12 DOC: And the::,(1.6) er: month- oh sorry
13 the: ,eh the y- the ye-
14 (1.4)
15 PAT: .hhhh Twelfth.=
16 DOC: =Sorry do y- when you sai- did
17 you say twentieth of the twelfth,=
18 PAT: =Ye[ah.]
19 DOC: [Wh-] What month is it,
20 PAT: [{(Ah-) >we- it i-< hhh aww:: hh } tch
21 [Patient shakes and scratches head]
22 it’s- huh hh (1.7) October.
23 DOC: Okay. And the year,
24 (1.9)
25 PAT: Twelfth.
26 DOC: Okay.
27 (0.4)
28 PAT: We- y- I m[ean h]hh
29 DOC: [°Yeah°]
30 COM: Yo[u mean two: thousand and ]twelve.
31 PAT: [Two thousand and twelve. Yeah.]  
32 DOC: And what season are we in,

After establishing the day (of the week) (line 04), the clinician proceeds to ask questions in accordance with the guidance, and therefore asks next for the date (line 05). Asking for the date sets up a number of possible relevant responses. Respondents might provide the date in its full form by including the date, month, and year, for example, 14th July 2017 or variants thereof—for example, 14th of the 7th, 2017. They could also legitimately offer something in less than full form—for example, 14th of July, or most minimally, just the 14th. In asking for the date as part of an expected sequence for the purpose of this assessment, a practitioner would only require the patient at that stage to produce the most limited form of response, that is, the 14th. Nevertheless, if the patient responds quite correctly in the full form, the response could be classed as correct for all parts of the expected sequence and the practitioner could move onto the next set of questions. However, this potential for variation also poses certain challenges for the patient in understanding what exactly is being required, especially in the absence of fuller explication of the terms of the question. Also, problems occur in this interaction when the patient actually produces “20th of the 12th” (line 11), when asked for the date, when in fact the date was October 22, 2012. Here the patient gets the year correct (albeit oddly formulated by abbreviating 2012 to “12th”), does not produce the month, and gets the date wrong (although this is within 2 days of the correct date so is considered an acceptable response in terms of scoring for a point on the test). This complexity is compounded by the fact that the clinician is sorting his papers and not paying full attention to the patient’s response. From lines 1–12, the clinician is flicking through the patient’s records to find the name label to stick onto the assessment form. He only breaks off from this activity in line 12 to initiate repair (Schegloff, Jefferson, & Sacks, 1977) on his own line of questioning.
when he realizes there was some problem with the patient’s prior response. This takes extra interactional work and is not straightforward. We can start to see a potential difficulty or confusion caused by the design of this question. The next extract demonstrates an alternative way to ask about the date by a different clinician.

Extract 5

01 DOC: .hh So first of all what day of the week is it today?
02 (0.8)
03 PAT: Er:, Tuesday.
04 DOC: .h And what month are we in now?
05 PAT: October.
06 DOC: And what date in October [is it?]
07 PAT: [Twenty]-three.
08 DOC: .hh An what year is it now,
10 (0.2)
11 PAT: U- Twelve.
12 DOC: And what season of the year is it,

The clinician asks the questions in a different order, deconstructing the date into components, and thereby removing the indeterminacy or ambiguity for the patient. As the clinician starts by asking for the day of the week, then month, she is able to then design the question regarding the date differently. By asking “and what date in October is it?” (line 07), the required answer is much clearer. It is important to note that both patients here scored above the higher cut-off for dementia and were diagnosed with functional memory disorder (Schmidtke, Pohlmann, & Metternich, 2008). Despite being clearer for the patient, this sequence of questions does not follow the order indicated on the administrative guidance for this part of the test. The different design of such questions can have implications for the responses they achieve and on how the interaction unfolds.

In another case, however, the clinician deviates from the standard instruction for the question.

Extract 6

01 DOC: Can you subtract seven from a hundred?
02 PAT: Hhh
03 (1.2)
04 PAT: Ninety-three:=
05 DOC: =Keep going subtracting seven.
06 PAT: Huuh hhh
07 COM: Huh huh
08 PAT: hhh (0.9) Ninety-thee: .h kch kch so that’d be- (3.2) Eighty-four, (4.3) °( ) Eighty-four,”
10 (2.0) S::: (1.4) Seventy-seven. tch (2.0)
11 Seventy. (1.3) .hhh (0.2) S::ixty-three, (1.2)
12 DOC: Okay. Can you . . .

In another case, however, the clinician deviates from the standard instruction for the question.

Extract 7

01 DOC: Now a bit of er mental arithmetic.= Can you
02 take seven away from one hundred for me?
The clinician introduces the task by indicating the type of activity which will take place, “now a bit of mental arithmetic” (line 01); then after each subtraction, the clinician repeats the answer as the framework for the next sum, meaning the number of origin is repeated back to the patient, for example, “and seven away from ninety-three” (line 05). There is evidence of this type of co-construction throughout the test for some of the clinicians, where they appear to be helping the patients by adding addition information into the question (see also Extract 9). By contrast, in Example 6, the patient was “going solo” and was required to remember each of the answers he had established before subtracting another seven. This places greater demands on the attention needed to complete the task. The different design not only marks a divergence from the standard test requirements given in the guidance but also places a differential “cognitive load” (Chandler & Sweller, 1991) on patients, with possible consequences for their scores.

**Interactional Non-Standardization: Providing Additional Help**

The last aspect of non-standardization appears when clinicians (sometimes) deviate from the scripted guidance to co-construct, or to “help” the patients to respond to the questions, which relates to the relationship between the tester and recipient of the test. This offer of help is inconsistent, with different clinicians administering the test differently for different patients. The next question requires the patient to identify the season of the year and ordinarily runs off like this.

**Extract 8**

01 DOC: And what season of the year is it, (0.2)  
02 (0.2)  
03 PAT: Autumn. (2.3)  
04 (2.3)

Prior to this question about the season (line 1), the patient (Extract 8) did not know the date and month; she was also unable to answer questions about her location (where they are), which the clinician has already (atypically) asked. The patient scored 36 on the ACE-111 and was diagnosed with Frontotemporal Dementia. Despite this, the clinician did not do any additional interactional work to help the patient to establish the correct answer (which would have been summer), and follows the standard procedure for the test. In the next extract, the patient also has an extremely high level of cognitive decline (ACE-111 score of 31), and during history-taking did not know his age, where he lives, or why he was at the clinic. Here the clinician is again asking, “what season of the year are we in” (lines 01 and 02), but on this occasion amends the standard administrative procedure by producing options from which the patient could choose (line 02). This kind of anticipatory work—anticipating trouble and explicating possible answers for the patient—is seldom done in other assessments (as shown in Extract 7).

**Extract 9**

01 DOC: Erm, what erm, what season of the year are we in? Is it spring, summer, autumn, winter? (0.4)  
02 (0.4)  
03 PAT: Erm. (0.4)
After a hesitation marker and pause, indicating the patient’s difficulty in responding (line 04), the clinician looks out of the window and states, “I know it’s hard to tell at the moment. Huh huh huh huh huh” (lines 05–06). This implies that the current weather condition, which is visible from the window, is atypical for the season they are in. It also works to excuse the patient for his displayed difficulty in answering. It could function to aid the patient in using the weather as an indication of season; for example, if it was raining, cold, and dark, and the weather was atypical, one could deduce that it was perhaps spring or summer. However, this is not the case for this patient who, after further prompting, suggests it is autumn, when in fact it is spring. On this occasion, this divergent administration of this part of the test did not appear to have any implication for the patient’s response, but it is important to appreciate how different practitioners can depart from standard procedure and guide and alter their practices to design the assessment differently. There is a key tension here in delivery of a cognitive instrument, between the standardized procedure and the different administrative designs employed. This adds an interactionally unique dimension to the test.

Discussion

It is recognized that if standard administration procedures are not followed, although results may be informative, they are not useful in determining a diagnosis (Venneri, 2005). Previous CA research into the administration of clinical tests has shown that questions that are expected to be asked in a standardized manner are, in fact, recurrently asked in diverging and diverse formats which influence test outcomes (Marlaire & Maynard, 1990; Maynard & Turowetz, 2017). We are not claiming that these variations in delivery affect the test outcomes, nor are we questioning the conduct of the clinicians or patients within these data. Indeed, the neurologists may be altering their communication to help the patients with their tasks. Krohne, Torres, Slettebo, and Bergland (2013) explore the experiences of health care professionals acting as standardized test administrators within acute geriatric care assessments. They note that this role as administrator places restrictions on health professionals that, “reduce the relational aspects of patient interaction.” They illustrate how therapists navigate between adherence to the test standard and meeting what they consider to be the individual patient’s needs in the test situation. It is further acknowledged that “the negative affects associated with these tools are felt by both the person being assessed as well as the professional administering the test” (Swallow & Hillman, 2019, p. 233). We also recognize that there may be other cognitive or social explanations for why patients may perform differently under different circumstances at any given time; for example, sleep deprivation (Rauchs et al., 2008), medication (Nevado-Holgado, Kim, Winchester, Gallacher, & Lovestone, 2016), language barriers, and cultural issues (Mirza, Panagioti, Waheed, & Waheed, 2017) may all affect test performance. However, we have demonstrated that some of the variations exhibited by clinicians can result in confusing patients (see Extract 1; also see Cahill et al., 2008; Keady & Gilliard, 2002) and there are links between high levels of confusion and anxiety, and reduced cognitive and brain functioning (Dotson et al., 2014), as well as negative effects on working memory (Williams et al., 2017), all of which could have an implication for poorer test performance (Kivimäki, 1995).

If the clinical priority is to ensure strictly standardized administration procedures during the conduct of these assessments, then clinical guidance needs to be clearer. All questions should be pre-scripted and guidance provided on other important elements of the interaction surrounding the test, for example, on how to introduce it, which currently does not feature. In addition, adequate training should be given to specialists who are required to use the tests in practice, in part to enable them to handle the contingencies that can arise in the administration of the test. It is worth reiterating that only one of these clinicians in these data has received any formal training on how to deliver this test and that was not part of their formal medical education. The absence of this provision within medical education, as well as the lack of clarity regarding specialist knowledge acquisition within policy (NICE, 2018), points to it being, as Maynard and Turowetz (2017, p. 485) termed it, a “domain of skill that is under-appreciated in the study of diagnosis.” Despite these suggestions, it is recognized that “clinical practice guidelines have had limited effect on changing physician behavior” (Cabana et al., 1999, p. 1458). It has also been demonstrated here that even when instruction regarding questioning is provided in the guidance, there is evidence of interactional non-standardization in the delivery of the assessments, and even when clinicians receive training (in the case of one of the practitioners in the data), variation in
administrative communication still remains. Such variations, one could suggest, are inevitable when assessments are carried out in interaction. If clinicians seek to establish a “true” measure of cognitive ability within this initial consultation, one that is not influenced by the interaction in which the test takes place, then one solution would be to remove the human or social element, employing computerized cognitive assessments (Newman et al., 2018). An alternative solution would be to remove the reliance on formal cognitive assessments in these initial consultations and instead use conversational markers. Previous research has demonstrated how language and communication during history-taking can be a useful tool to help clinicians determine differential diagnosis (Elsey et al., 2015; Jones et al., 2016; Reuber et al., 2018). If clinicians were able to form a working diagnosis through the conversations they have with patients, then this would make formal cognitive testing at this stage redundant. Furthermore, to alleviate patient anxieties and concerns, and promote a positive patient experience, the social and interactional elements of these consultations are essential. Clinicians helping during assessment practices (see Extract 8; also see Sacks, 1992) is one manner in which interaction can be crafted to contribute to establishing a positive relationship between tester and tested (Swallow & Hillman, 2019), and while this variation in administration may undermine standard assessment procedures, it could be seen to be an important component for enhancing patient experience.

Conclusion
In sum, we have shown that when there is a lack of guidance, there is evidence of interactional variation during the introduction of a cognitive assessment tool. Furthermore, we have presented evidence of a lack of standardization in administration of the ACE-111. Clinicians do not always follow the scripted instruction that is provided in the guidance; clinicians may use different delivery or administration procedures, and clinicians also vary their approach for different patients. We can see that these interactional modifications have potential implications for how the patient understands the task at hand, their level of confusion, and how they respond to certain questions. The interactional complexity within the delivery of the ACE-111 means that administrative standardization is rarely achieved in practice.

Acknowledgments
The authors would like to thank the patients, companions, and health care professionals whose consultations were filmed. All authors were involved in the study design, analysis, and writing of the article.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was partially funded by a grant by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Program (Grant Reference Number PB-PG-0211-24079). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

ORCID iD
Danielle Jones https://orcid.org/0000-0002-2875-781X

Notes
1. Transcription conventions
   DOC/PAT Speakers labels (DOC = Clinician; PAT = Patient)
   COM (COM = companion)
   [overlap] Brackets: Onset and offset of overlapping talk.
   = Equals sign: Utterances are latched or ran together, with no gap of silence.
   - Hyphen: Preceding sound is cut off/self-interrupted.
   (0.0) Time pause: Silence measured in seconds and tenths of seconds.
   (.) Parentheses with a period: A micropause of less than 0.2 s.
   : Colon(s): Preceding sound is extended or stretched; the more the longer.
   . Period: Falling or terminal intonation.
   , Comma: Continuing or slightly rising intonation.
   ? Question mark: Rising intonation.
   _ Underline: Increased volume relative to surrounding talk.
   °soft° Degree signs: Talk with decreased volume relative to surrounding talk.
   >fast< Greater-than/less-than signs: Talk with increased pace relative to surrounding talk.
   <slow> Less-than/greater-than signs: Talk with decreased pace relative to surrounding talk.
   .h Superscripted periods preceding h’s: Inbreaths; the more the longer.
   h H’s: Outbreaths (sometimes indicating laughter); the more the longer.
   hah/heh Laugh token: Relative open or closed position of laughter.
   (that)/(hat) Filled single parentheses: Transcriptionist doubt about talk. Alternative hearings.
   (. . .) Empty single parentheses: Transcriptionist cannot identify talk.
   ((Cough)) Filled double parentheses: Additional details or an event/sound not easily transcribed.
References

Alzheimer’s Research UK. (2018). Diagnoses in the UK. Dementia Statistics Hub. Retrieved from https://www.dementia-statistics.org/statistics/diagnoses-in-the-uk/

Ballard, C. (2015). Helping you to assess cognition: A practical toolkit for clinicians. Retrieved from https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/alzheimers_society_cognitive_assessment_toolkit.pdf

Beach, W. A. (1995). Preserving and constraining options: “Okays” and “official” priorities in medical interviews. In B. Morris & R. Chenail (Eds.), Talk of the clinic: Explorations in the analysis of medical and therapeutic discourse (pp. 259–290). Hillsdale, NJ: Lawrence Erlbaum.

Bentvelzen, A., Aerts, L., Seeher, K., Wesson, J., & Brodaty, H. (2017). A comprehensive review of the quality and feasibility of dementia assessment measures: The dementia outcomes measurement suite. Journal of the American Medical Directors Association, 18, 826–837. doi:10.1016/j.jamda.2017.01.006

Berger, C. R. (1974). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. Human Communication Research, 1, 99–112. doi:10.1111/j.1468-9581.1975.tb00258.x

Berger, C. R. (1986). Uncertain outcome values in pre -dicted relationships: Uncertainty reduction theory then and now. Human Communication Research, 13, 34–38. doi:10.1111/j.1468-2958.1986.tb00093.x

Berger, C. R., & Calabrese, R. J. (1974). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. Human Communication Research, 1, 99–112. doi:10.1111/j.1468-9581.1975.tb00258.x

Boise, L. (2006). Improving dementia care through physician education: Some challenges. Clinical Gerontologist, 29(2), 3–10. doi:10.1300/J108v29n02_02

Borson, S., Frank, L., Boustanli, M., Dean, M., Lin, P.-J., . . . Ashford, J. W. (2013). Improving dementia care: The role of screening and detection of cognitive impairment. Alzheimer’s & Dementia, 9, 151–159. doi:10.1016/j.jalz.2012.08.008

Brooke, P., & Bullock, R. (1999). Validation of a 6 Item Cognitive Impairment Test. International Journal of Geriatric Psychiatry, 14, 936–940. doi:10.1002/(SICI)1099-1166 (199911)14:11<936::AID-GPS39>3.0.CO;2-1

Cabana, M. D., Rand, C. S., Powe, N. R., Wu, A. W., Wilson, M. H., Abboud, P.-A., & Rubin, H. R. (1999). Why don’t physicians follow clinical practice guidelines? A framework for improvement. Journal of the American Medical Association, 282, 1458–1465. doi:10.1001/jama.282.15.1458

Cahill, S. M., Gibb, M., Bruce, I., Headon, M., & Drury, M. (2008). “I was worried coming in because I don’t really know why it was arranged”: The subjective experience of new patients and their primary caregivers attending a memory clinic. Dementia, 7, 175–189. doi:10.1177/1471301208091157

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8, 293–332. doi:10.1207/s1532690xci0804_2

Chappell, P., Toerien, M., Jackson, C., & Reuber, M. (2018). Following the patient’s orders? Recommending vs. offering choice in neurology outpatient consultations. Social Science & Medicine, 205, 8–16. doi:10.1016/j.socscimed.2018.03.036

Chesnot, R., Bender, M., & Byatt, S. (2000). Involving people who have dementia in the evaluation of services: A review. Journal of Mental Health, 9, 471–479. doi:10.1080/09638230020005200

Chui, H. C., Victoroff, J. I., Margolin, D., Jagust, W., Shankle, R., & Katzman, R. (1992). Criteria for the diagnosis of ischemic vascular dementia proposed by the State of California Alzheimer’s Disease Diagnostic and Treatment Centers. Neurology, 42, 473–480. doi:10.1212/WNL.42.3.473

Crawford, S., Whitnall, L., Robertson, J., & Evans, J. J. (2012). A systematic review of the accuracy and clinical utility of the Addenbrooke’s Cognitive Examination and the Addenbrooke’s Cognitive Examination–Revised in the diagnosis of dementia. International Journal of Geriatric Psychiatry, 27, 659–669. doi:10.1002/gps.2771

Cullen, B., O’Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2007). A review of screening tests for cognitive impairment. Journal of Neurology, Neurosurgery, and Psychiatry, 78, 790–799. doi:10.1136/jnnp.2006.095414

Curl, T., & Drew, P. (2008). Contingency and action: A comparison of two forms of requesting. Research on Language and Social Interaction, 41, 129–153. doi:10.1080/08351810802028613

Department of Health. (2012). Prime minister’s challenge on dementia: Delivering major improvements in dementia care and research by 2015. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215101/dh_133176.pdf

Department of Health. (2015). Prime minister’s challenge on dementia 2020. Retrieved from https://www.gov.uk/government/publications/prime-ministers-challenge-on-dementia-2020

Dooley, J., Bass, N., & McCabe, R. (2018). How do doctors deliver a diagnosis of dementia in memory clinics? The British Journal of Psychiatry, 212, 239–245. doi:10.1192/bjp.2017.64

Dotson, V. M., Szymkowicz, S. M., Kirton, J. W., McLaren, M. E., Green, M. L., & Rohani, J. Y. (2014). Unique and interactive effect of anxiety and depressive symptoms on cognitive and brain function in young and older adults. Journal of Depression & Anxiety, Suppl 1, 22565. doi:10.4172/2167-1044.S1-003

Drew, P. (2012). Turn design. In J. Sidnell & T. Stivers (Eds.), The handbook of conversation analysis (pp. 131–149). Sussex, UK: Wiley-Blackwell. doi:10.1002/9781118325001.ch7

Drew, P., Raymond, G., & Weinberg, D. (Eds.). (2006). Talk and interaction in social research methods. London: SAGE.

Elamin, M., Holloway, G., Bak, T. H., & Pal, S. (2016). The utility of the Addenbrooke’s Cognitive Examination version three in early-onset dementia. Dementia and Geriatric Cognitive Disorders, 41(1–2), 9–15. doi:10.1159/000439248

Elsey, C., Drew, P., Jones, D., Blackburn, D., Wakefield, S., Harkness, K., . . . Reuber, M. (2015). Towards diagnostic conversational profiles of patients presenting with dementia or functional memory disorders to memory clinics. Patient Education and Counseling, 98, 1071–1077. doi:10.1016/j.pec.2015.05.021
Heritage, J., Robinson, J. D., Elliott, M. N., Beckett, M., & Heritage, J., & Maynard, D. W. (2006). Communication in medical care: Interaction between primary care physicians and patients. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511607172

Heritage, J., & Robinson, J. D. (2006). The structure of patients’ presenting concerns: Physicians’ opening questions. Health Communication, 19, 89–102. doi:10.1207/s15327027hc1902_1

Heritage, J., Robinson, J. D., Elliott, M. N., Beckett, M., & Wilkes, M. (2007). Reducing patients’ unmet concerns in primary care: The difference one word can make. Journal of General Internal Medicine, 22, 1429–1433. doi:10.1007/s11606-007-0279-0

Heritage, J., & Stivers, T. (1999). Online commentary in acute medicine visits: A method of shaping patient expectations. Social Science & Medicine, 49, 1501–1517. doi:10.1016/S0277-9536(99)00219-1

Larner, A. J. (2017). Introduction to cognitive screening instruments. In P. Harris (Ed.), Talk at work (pp. 235–267). Cambridge, UK: Cambridge University Press.

Levinson, W., Roter, D. L., Mullolooy, J. P., Dull, V. T., & Frankel, R. M. (1997). Physician-patient communication: The relationship with malpractice claims among primary care physicians and surgeons. Journal of the American Medical Association, 277, 553–559. doi:10.1001/jama.1997.03540310051034

Marlaire, C. L., & Maynard, D. W. (1990). Standardized testing as an interactional phenomenon. Sociology of Education, 63, 83–101. doi:10.2307/2112856

Mattias-Guiu, J. A., Cortés-Martínez, A., Valles-Salgado, M., Rognoni, T., Fernández-Matarrubia, M., Moreno-Ramos, T., & Mattias-Guiu, J. (2017). Addenbrooke’s Cognitive Examination III: Diagnostic utility for mild cognitive impairment and dementia and correlation with standardized neuropsychological tests. International Psychogeriatrics, 29, 105–113. doi:10.1017/S1041610116001496

Maynard, D. W. (2017). Delivering bad news in emergency care medicine. Acute Medicine & Surgery, 4, 3–11. doi:10.1002/ams2.210

Maynard, D. W., & Heritage, J. (2005). Conversation analysis, doctor-patient interaction and medical communication. Medical Education, 39, 428–435. doi:10.1111/j.1365-2929.2005.02111.x

Maynard, D. W., & Marlaire, C. L. (1992). Good reasons for bad testing performance: The interactional substrate of educational exams. Qualitative Sociology, 15, 177–202. doi:10.1007/BF00989493

Monzoni, C. M., Duncan, R., Grünwald, R., & Reuber, M. (2016). Are there interactional reasons why doctors may have emotional causes? A conversation analytic study in neurology outpatients. Patient Education and Counseling, 85, 189–200. doi:10.1016/j.pec.2011.07.014

Mukadam, N., Cooper, C., Kherani, N., & Livingston, G. (2015). A systematic review of interventions to detect dementia or cognitive impairment. International Journal of Geriatric Psychiatry, 30, 32–45. doi:10.1002/gps.4184

Murphy, K., O’Connor, D. A., Browning, C. J., French, S. D., Michie, S., Francis, J. J., . . . Green, S. E. (2014). Understanding diagnosis and management of dementia and guideline implementation in general practice: A qualitative study using the theoretical domains framework. Implementation Science, 9(1), Article 31. doi:10.1186/1748-5908-9-31

National Institute for Health and Care Excellence. (2018). Dementia: Assessment, management and support for people living with dementia and their carers (NICE guideline [NG97]). Retrieved from https://www.nice.org.uk/guidance/ng97

Nevado-Holgado, A. J., Kim, C.-H., Winchester, L., Gallacher, J., & Lovestone, S. (2016). Commonly prescribed drugs associate with cognitive function: A cross-sectional study
in UK Biobank. *BMJ Open*, 6(11), e012177. doi:10.1136/bmjopen-2016-012177

Newman, C. G. J., Bevins, A. D., Zajicek, J. P., Hodges, J. R., Vuillermoz, E., Dickenson, J. M., . . . Noad, R. F. (2018). Improving the quality of cognitive screening assessments: ACOMobile, an iPad-based version of the Addenbrooke’s Cognitive Examination-III. *Alzheimer’s & Dementia*, 10, 182–187. doi:10.1016/j.dadm.2017.12.003

Peräkylä, A. (1998). Authority and accountability: The delivery of diagnosis in primary health care. *Social Psychology Quarterly*, 61, 301–320. doi:10.2307/2787032

Rauchs, G., Schabus, M., Parapatics, S., Bertran, F., Clochon, P., Hot, P., . . . Anderer, P. (2008). Is there a link between sleep changes and memory in Alzheimer’s disease? *Neuroreport*, 19, 1159–1162. doi:10.1097/WNR.0b013e32830867c4

Reuber, M., Blackburn, D. J., Elsey, C., Wakefield, S., Ardern, K. A., Harkness, K., . . . Drew, P. (2018). An interactional profile to assist the differential diagnosis of neurodegenerative and functional memory disorders. *Alzheimer Disease & Associated Disorders*, 32, 197–206. doi:10.1097/WAD.0000000000000231

Robinson, J. D., & Heritage, J. (2014). Intervening with conversation analysis: The case of medicine. *Research on Language and Social Interaction*, 47, 201–218. doi:10.1080/08351813.2014.925658

Robinson, J. D., & Stivers, T. (2001). Achieving activity transitions in physician-patient encounters: From history taking to physical examination. *Human Communication Research*, 27, 253–298. doi:10.1111/j.1468-2958.2001.tb00782.x

Sacks, H. (1992). *Lectures on conversation* (G. Jefferson, Ed., with Introduction by Schegloff, E. A.) (2 vols.). Oxford, UK: Basil Blackwell.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361–382. doi:10.1353/lan.1977.0041

Schmidtke, K., Pohlmann, S., & Metternich, B. (2008). The syndrome of functional memory disorder: Definition, etiology, and natural course. *The American Journal of Geriatric Psychiatry*, 16, 981–988. doi:10.1097/JGP.0b013e318187df19

Sheer, V. C., & Cline, R. J. (1995). Testing a model of perceived information adequacy and uncertainty reduction in physician patient interactions. *Journal of Applied Communication Research*, 23, 44–59. doi:10.1080/00909889509365413

Smith, S. (2015). *GP’s competencies in assessment and diagnosis* (Unpublished presentation). University of Bradford, UK.

Steptoe, A., Breeze, E., Banks, J., & Nazroo, J. (2013). Cohort profile: The English longitudinal study of ageing. *International Journal of Epidemiology*, 42, 1640–1648. doi:10.1093/ije/dys168

Stivers, T. (2007). *Prescribing under pressure: Parent-physician conversations and antibiotics*. New York: Oxford University Press.

Stivers, T., Heritage, J., Barnes, R., McCabe, R., Thompson, L., & Toerien, M. (2018). Treatment recommendations as actions. *Health Communication*, 33, 1335–1344. doi:10.1080/10410236.2017.1350913

Swallow, J., & Hillman, A. (2019). Fear and anxiety: Affects, emotions and care practices in the memory clinic. *Social Studies of Science*, 49, 227–244. doi:10.1177/0306312718820965

Venneri, A. (2005). The promised land: The blooming business of neuropsychological assessment guidance books. *Cortex*, 41, 96–98. doi:10.1016/S0010-9452(08)70184-9

Wilkinson, R. (2013). The interactional organization of aphasia naming testing. *Clinical Linguistics & Phonetics*, 27, 805–822. doi:10.3109/02699206.2013.815279

Williams, M. W., Kueider, A. M., Dmitrieva, N. O., Manly, J. J., Pieper, C. F., Verney, S. P., & Gibbons, L. E. (2017). Anxiety symptoms bias memory assessment in older adults. *International Journal of Geriatric Psychiatry*, 32, 983–990. doi:10.1002/gps.4557

**Author Biographies**

**Danielle Jones** is a lecturer in Dementia Studies in the Centre for Applied Dementia Studies at the University of Bradford, Bradford, United Kingdom.

**Ray Wilkinson** is a professor of human communication in the Department of Human Communication Sciences at the University of Sheffield, Sheffield, United Kingdom.

**Clare Jackson** is a senior lecturer in the Department of Sociology at the University of York, York, United Kingdom.

**Paul Drew** is a professor in the Department of Language and Linguistic Science at the University of York, York, United Kingdom.