IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Assessment battery for communication: Development of two equivalent forms

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/102463 since

Published version:

DOI:10.1016/j.jcomdis.2012.03.002

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

08 September 2024

UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:

Journal of Communication Disorders 45 (2012) 290–303

Bosco F.M., Angeleri R., Zuffranieri, M. Bruno G. Bara, Sacco K. (2012).

Assessment battery for communication: Development of two equivalent forms.:

http://www.journals.elsevier.com/journal-of-communication-disorders/

Assessment Battery for Communication: Development of Two Equivalent Forms

Abstract

The aim of this paper was to develop and test two equivalent forms of the Assessment Battery for Communication (ABaCo), a tool for evaluating pragmatic abilities in patients with neuropsychological and psychiatric disorders. The equivalent forms were created using the data from a sample of 390 children, then tested in a sample of 30 patients with traumatic brain injury. Equivalent forms of the same test are useful in clinical practice and intervention research, when performance needs to be tested on two separate occasions, for example before and after a rehabilitation program. The present results provide additional evidence on the psychometric functioning of the equivalent forms of the ABaCo and their usability in a clinical context.

Keywords: equivalent forms, pragmatic deficit, assessment, communication, rehabilitation

1. Introduction

The benefits of a rehabilitation program can only be determined if there is empirical evidence of its effectiveness. In the current literature there is a real risk of confounding recovery after rehabilitation with practice effects, as the same neuropsychological tests are often administered to patients more than once. This confounding factor may have important consequences for both research and practice since the effectiveness of a clinical intervention cannot be empirically proved (for a discussion in using evidence-based practice to support clinical decision-making in communication disorders, see Dollaghan, 2004). One way to avoid this kind of confound is therefore to develop equivalent alternative forms of the same test, so that patients can be assessed and re-assessed before and after a rehabilitation program, thus obtaining a reliable measure of treatment efficacy. Equivalent forms are interchangeable measures of the same construct (Raykov & Marcoulides, 2009), which can be used independently of each other and be considered alternative measures of the same test.

The need for valid and reliable instruments to assess the success of a rehabilitation program is particularly important in clinical settings, as well as in the context of intervention research.

The purpose of the present paper is to develop and test two equivalent forms of the Assessment Battery for Communication (ABaCo; Sacco, Angeleri, Bosco, Colle, Mate, & Bara, 2008). The ABaCo is a clinical battery for the evaluation of pragmatic abilities in patients affected by neuropsychological disorders following focal and diffuse brain injury - e.g., right hemisphere damage (RHD) and traumatic brain injury (TBI) - and psychiatric disorders such as schizophrenia. Pragmatic ability can be

broadly defined as the use of language to convey meaning in specific contexts (Levinson, 1983); it includes a varied set of skills that are required for communicative competence in naturalistic and functional use of language.

The ability to interpret pragmatic language appropriately in social interactions is key to successful functioning in many aspects of everyday life, and is often disrupted by brain pathology. Indeed, despite differences in the etiology and the clinical profile of these pathologies, patients with acquired brain injury and schizophrenia encounter similar communicative difficulties in the comprehension and production of several pragmatic tasks (for a review see Cummings, 2007; Davis, 2007). More specifically, impairments in *understanding* the literal meaning of utterances - including the interpretation of indirect speech acts, metaphor, irony, or humor - have been observed in patients with RHD (e.g., Kaplan, Brownell, Jacobs, & Gardner, 1990; Brownell, Simpson, Bihrle, Potter, & Gardner, 1990; Winner, Brownell, Happé, Blum, & Pincus, 1998), TBI (e.g., Docking, Murdoch, & Jordan, 2000; Angeleri et al., 2008) and schizophrenia (e.g., Langdon, Coltheart, Ward, & Catts, 2002; Corcoran, 2003). Furthermore, TBI, RHD, and schizophrenic patients all find it difficult to understand prosody (Marquardt, Rios- Brown, Richburg, Seibert, & Cannito, 2001; Meilijson, Kasher, & Elizur, 2004; Pell, 2007).

The *production* of adequate communicative acts is another crucial facet of pragmatic ability; patients with brain injury have difficulty adapting their utterances to specific contexts and may show tangential speech, poor verbal organization, or inadequate turn taking. Both TBI and RHD patients exhibit difficulties in producing requests (McDonald & Van Sommers, 1993) and giving the interlocutor adequate information (Angeleri et al., 2008; McDonald, 1993; Stemmer, Giroux, & Joanette,

1994; Brownell & Stringfellow, 1999). Schizophrenic patients are impaired in their capacity to integrate contextual information (Bazin, Sarfati, Lefrère, Passerieux, & Hardy-Baylè, 2005). Finally, TBI, RHD, and schizophrenic patients all show some difficulty in managing dyadic conversations (Brownell, Pincus, Blum, Rehak, & Winner, 1997; Leroy, Pezard, Nandrino & Beaune, 2005; Russeaux, Vérigneaux & Kozlowski, 2010).

There is widespread agreement that the assessment of pragmatics is a crucial component of the evaluation of communicative competence after brain injury, and that pragmatics should be a critical focus of the rehabilitation process. Indeed, poor communication skills have been shown to be a serious barrier to community reintegration and personal autonomy (Galski, Tompkins, & Johnston, 1998; Sohlberg & Mateer, 2001). At present, the most common methods used in the assessment of pragmatic deficits are *pragmatic assessment* and *functional assessment*.

Tests based on pragmatic assessment aim to identify and measure individual cognitive processes underlying a range of communicative behaviors, in order to describe different impairment profiles. Examples of these include: the Pragmatic Protocol (PP; Prutting & Kirchner, 1987) and the Profile of Communicative Appropriateness (PCA; Penn, 1985), two checklists based on speech act theory designed to evaluate the appropriateness of specific pragmatic components during naturally occurring conversation; the Right Hemisphere Communication Battery (RHCB; Gardner & Brownell, 1986) and the Right Hemisphere Language Battery (Bryan, 1995), two tests of pragmatic skills designed for patients with RHD; and the Awareness of Social Inference Test (TASIT; McDonald, Flanagan, Rollins, & Kinch, 2003), which focuses on pragmatic deficits in TBI patients.

Tests based on functional assessment aim to measure the ability to communicate efficiently in real-life situations, without directly identifying the basic abilities underlying communication. Examples of these include: the Functional Communication Profile (FCP; Sarno, 1969), which rates the effectiveness of communicative behavior during informal conversations; the American speech-language hearing association Functional Assessment of Communication skills for adults (ASHA FACS; Frattali, Thompson, Holland, Wohl, & Ferketic, 1995), which uses functional communication tasks such as naming familiar people, expressing feeling and so on; the Communicative Abilities in Daily Living (CADL; Holland, 1980; Holland, Frattali, & Fromm, 1998), which makes use of role-playing, reproducing everyday social situations; and the Scenario Test (van der Meulen, van de Sandt-Koenderman, Duivenvoorden & Ribbers, 2010), which examines the effectiveness of verbal and non-verbal communication in an interactive setting with a supportive communication partner.

In a previous paper (Sacco et al., 2008) we discussed why, notwithstanding the merits of all the above-mentioned assessment tools, we consider them unsatisfactory for different reasons, motivating the creation of the ABaCo. The ABaCo was developed in order to allow a comprehensive assessment of pragmatic abilities, and encompasses a wide range of pragmatic abilities and communicative modalities (e.g., linguistic, extralinguistic and paralinguistic), evaluated with an objective coding system. The theoretical basis of the ABaCo is Cognitive Pragmatics (Bara, 2010), a theory on the cognitive processes underlying human communication. Cognitive Pragmatics has been used to explain the development of communication in typically developing children (Bosco, Bucciarelli, & Bara, 2004; 2006; Bosco & Bucciarelli,

2008, Bosco, Angeleri, Colle, Sacco & Bara, 2012), as well as communicative deficits in patients with different patterns of brain impairment (Bara, Tirassa, & Zettin, 1997; Bara, Bosco, & Bucciarelli, 1999; Bara, Cutica, & Tirassa, 2001).

1.1 The Assessment Battery for Communication (ABaCo)

The ABaCo (Sacco et al., 2008) includes a total of 180 items: 72 *in vivo* items and 108 items based on videotaped scenes. Each videotaped scene lasts 20-25 seconds and the number of words in each sentence is controlled (range: 5 to 9 words). For the *in vivo* items, the examiner directly asks questions to the participant assuming the role of an interlocutor. For the items based on videotaped scenes, the examiner shows the participant a video clip, then asks him/her a number of questions concerning the communicative interaction depicted in the clip.

The ABaCo is organized in five evaluation scales: (a) Linguistic, (b) Extralinguistic, (c) Paralinguistic, (d) Context, and (e) Conversational[1]. A detailed description of the pragmatic phenomena included in the ABaCo, together with sample items and test questions, is provided in Appendix A. We will now summarize the most salient features of each scale.

1.1.1 Linguistic and extralinguistic scales. The linguistic scale assesses the comprehension and production of communication acts expressed primarily through

[1] The ABaCo was originally created for use with Italian speakers. However, several parts of the battery (i.e. all parts that do not include spoken video-clips) have been already adapted for English speakers: specifically, the linguistic comprehension and production of Basic speech acts, the extralinguistic scale, the paralinguistic scale, and the conversational scale. All these parts of the Battery are obviously adaptable in others languages.

linguistic means. The extralinguistic scale also assesses the comprehension and production of communication acts, but only expressed through extralinguistic means. It includes the same communication acts included in the linguistic scale. For this reason, the tasks used in these two scales are described together.

We used the following tasks to assess the *comprehension* of linguistic and extralinguistic communication acts:

- *Basic speech acts (BSAs,* Kasher 1991*)*: the examiner asks the participants to evaluate the truthfulness of assertions, to answer simple questions, to perform actions on request, to execute orders. In the extralinguistic scale, the examiner shows the participants short videos where an actor makes an assertion, asks a question, makes a request or issues a command through the use of gestures. The participant has to understand the act produced by the actor.
- *Standard (direct and indirect) communication acts, deceit and irony*: the examiner shows the subject short videos where two agents are engaged in a communicative interaction: the actor asks his partner a question and the partner replies. The participant has to understand the communication act produced by the partner. On the linguistic scale the actors communicate verbally, whereas on the extralinguistic scale they only communicate through gestures.

We used the following tasks to assess the *production* of linguistic and extralinguistic communication acts:

- *Basic speech acts (BSAs)*: the examiner asks the participants to produce assertions, questions, requests and commands (Kasher, 1991) the

examiner provides the semantic content of the requested act. For example, the examiner asks the participant "Tell me that you are cold", or "Order me to be quiet". In the linguistic scale the participant has to produce linguistic acts, while in the extralinguistic scale only gestural acts.

- *Standard (direct and indirect) communication acts, deceit and irony*: the examiner shows the subject short videos where two agents are engaged in a communicative interaction: the actor says something to the partner, the video stops and the subject is requested to assume the partner's perspective in answering the actor. On the linguistic scale the communicative interaction occurs in the linguistic modality and the subject has to reply verbally. On the extralinguistic scale the actor performs communicative gestures without any language support, and the subject has to reply using gestures alone.

1.1.2 Paralinguistic scale. The paralinguistic scale assesses the comprehension and production of those aspects that generally accompany a communication act, such as proxemics and prosody. The paralinguistic scale comprises:

- *Basic Communication Acts* (assertions, questions, requests, and commands)
- *Basic Emotions* (anger, sadness, happiness, and fear)
- Acts characterized by a *paralinguistic contradiction*

We used the following tasks to assess the *comprehension* of paralinguistic aspects:

- *Basic Communication Acts*: This task investigates the BSAs originally theorized by Kasher (1991) in the paralinguistic modality. BSAs are very basic and prototypical types of Speech acts. A BSA is generally performed by uttering a specific kind of sentence which is linguistically marked as

appropriate for it.

The examiner shows the subject short videos in which an actor, speaking an invented language, makes an assertion, asks a question, makes a request, or gives a command. The subject has to understand the type of act produced by the actor, by relying on paralinguistic aspects.

- *Basic Emotions*: the examiner shows the subject short videos in which an actor, speaking an invented language, expresses a basic emotion. The participant has to recognize the emotion by relying on paralinguistic aspects.

Paralinguistic Contradiction: the examiner shows the participant short videos in which two agents are engaged in a communicative interaction: the actor verbally expresses something that is in contrast with the paralinguistic indicators (e.g., the actor says "I like that cake very much!" while his voice and attitude reveal that he doesn't like it at all). The participant has to understand what the actor actually thinks, in the example that the actor doesn't like the cake; this information is detectable through the paralinguistic indicators.

We used the following tasks to assess the *production* of paralinguistic aspects:

- *Basic Communication Acts*: the examiner asks the subject to produce assertions, questions, requests, and commands, paying special attention to paralinguistic aspects.
- *Basic Emotions*: the examiner asks the subject to produce communication acts colored by a specific emotion or mood; the examiner provides the semantic content of the requested act and the emotion with which it has to be expressed.

1.1.3 Context scale. The context scale assesses the adequacy/inadequacy of a communication act -assessed only in comprehension - with respect to Grice's maxims (Grice, 1989). Furthermore it assesses: (a) the adequacy/inadequacy of a communication act with respect to social norms (Lakoff, 1973), that is, the ability to recognize whether and why a communication act is appropriate in a given context or situation (comprehension); and (b) the ability to produce communication acts appropriate to a given context or situation, according to rules of formality and informality (production).

We used the following tasks to assess the *comprehension* of discourse and social norms:

- *Grice's Maxims* (Grice, 1989): the examiner shows the subject short videos where two agents are engaged in a communicative interaction; the actor asks his partner a question; the partner replies either according to or violating the norms of discourse, giving a generic, false, irrelevant or ambiguous answer. The participant has to detect and explain the adequacy/inadequacy of the partner's reply.
- *Social Norms*: the examiner shows the subject short videos in which two agents are engaged in a communicative interaction; the actor asks his partner a question; the partner replies either according to the norms of social appropriateness (Lakoff, 1973) or in a manner which is not appropriate in the given social context.

We used the following tasks to assess *production*:

- *Social norms*. The examiner asks participants to produce communication acts requiring different levels of formality/informality; the examiner provides the semantic content of the requested act.

1.1.4 Conversational scale. This last scale assesses the ability to participate appropriately in a conversation, complying with the topics of the discourse and turn-taking. The examiner and the participant are engaged in four short conversations, where the examiner introduces four different topics, for a total duration of 4-6 minutes each. The scale evaluates the following aspects:

- *Topic management*: (1) topic maintenance, (2) new topic introduction, and (3) change in topic.
- *Turn-Taking*: (1) taking turns and (2) respecting the partner's turn during exchanges with the examiner.

Note that, in the present study, we will only consider the first four scales; the Conversational scale comprises only four items, and is therefore too short to be split into two equivalent versions.

The Linguistic scale, some parts of the Paralinguistic comprehension scale, and the Context comprehension scale contain items in Italian (i.e., videotaped scenes in which the actors speak Italian) and thus they are presently suitable only for Italian speakers; the other sections of the battery (i.e., the Extralinguistic scale, some parts of the Paralinguistic scale, and the Context production scale) are also suitable for speakers of other languages, since they do not involve language but only gestures, prosody, facial expressions, and face-to-face interactions with the examiner.

In Sacco et al. (2008), we reported descriptive and validation data of the ABaCo.

With the sole exception of the production of items on the Context scale, which had an $\alpha = .52$, the internal consistency of the scales that make up the ABaCo ranged from $\alpha = .63$ to $\alpha = .91$, thus showing satisfactory to excellent internal consistency (De Vellis, 1991; Nunnally, 1978). The ABaCo also showed high inter-rater agreement: in particular, following the criteria defined by Landis and Koch (1977), inter-rater correlations showed excellent agreement ($k > .75$) within each sub-scale, ranging from $k = .76$ to $k = .96$.

So far, the ABaCo has been used to assess communicative competence in a group of 21 TBI patients, and has proved useful in identifying specific pragmatic deficits in their communicative performance (Angeleri et al., 2008). Moreover, we recently collected a normative sample of adult Italian speakers (stratified by age and education), so as to obtain a normative distribution against which to compare the pragmatic performance of patients (Angeleri, Bosco, Gabbatore, Bara, & Sacco, 2011).

Aim of the Study

The aim of the present study was to develop and test two equivalent forms of the ABaCo. We developed the two forms of the battery using data from a large sample of children aged 5 to 8.5 years; then, we further tested the equivalence and psychometric characteristics of the two forms in a sample of TBI patients.

2. Material and Method

2.1 Participants

2.1.1 Children sample. Three hundred and ninety children took part in the

study. They were aged 5 to 8.5 (M = 6.6 yrs; SD = 1.2). The children sample was the same already described by Sacco and colleagues (2008). We decided to use a large sample of children of different ages so as to increase score variability. Indeed, children's scores were normally distributed, while those of adults exhibited substantial skewness (for further details, see Angeleri et al., 2012). Moreover, Bosco and Bucciarelli (2008) found that 5-year-old children do not show ceiling performance even on the easiest kind of tasks included in the battery, and that 8.5-year-olds do not show floor performance on the most difficult ones. Following the indications by Bosco and Bucciarelli (2008), we divided the age range into three intervals of six months each, thus obtaining three age groups of equal size: 130 children ranging from 5 to 5.5 years of age, 130 children ranging from 6.5 to 7 years, and 130 children ranging from 8 to 8.5 years. Children in each age group were balanced by gender.

Children were recruited from kindergarten and elementary schools in the Turin area (northern Italy). The children came from families ranging from the working to the middle and upper classes; socio-economic status (SES) was measured by family composition, parental education level and occupation, and was obtained using a questionnaire filled in by the participants' parents. The SES index was derived from the Two-Factor Index of Social Status (Hollingshead, 1975); we updated the employment categories included in the Hollingshead procedure with reference to the current Italian social context. In our sample, the modal SES corresponded to a middle-class family (43% of the children). Children with physical disabilities and neurological/psychiatric disorders were excluded from the study; all children were native Italian speakers. Parents were given details about the study and gave informed consent for their children to participate in the study.

2.1.2 TBI sample. The TBI sample consisted of 30 TBI patients (7 female, 23 male) ranging in age from 20 to 68 years (*M* = 37.1; *SD* = 11.4). Education ranged from 5 to 18 years of schooling (*M* = 11.1; *SD* = 3.3). The TBI sample used in this study includes 21 patients already described by Angeleri and colleagues (2008). TBI patients were recruited from three different neuropsychological rehabilitation centers located in Genoa and Turin. The time after onset ranged from 3 to 252 months (*M* = 52.8; *SD* = 56.2). At the time of the study, all patients were living at home; all were in a post-acute phase, and none lived independently (i.e., without a partner or a parent). Patients were included if they met the following criteria: (1) be at least 18 years old; (2) be at least three months post-brain injury; (3) be native Italian speakers; and (4) have sufficient cognitive and linguistic skills, assessed through the following neuropsychological tests using a predetermined cutoff score: Mini Mental State Examination (MMSE, Folstein, Folstein, & McHugh, 1975; cutoff: 24/30); denomination scale of the Aachener Aphasie Test (AAT, cutoff: no deficit); Token Test short version (De Renzi & Vignolo, 1962; cutoff: 5/6). Exclusion criteria were a prior history of brain injury or other neurological disease, and pre-morbid alcohol or drug addiction. All the participants gave written informed consent.

2.2 Material and Procedure

In the present study we used the first four scales of the ABaCo, as described in the section *"Assessment Battery for Communication (ABaCo)"*. The four scales (Linguistic, Extralinguistic, Paralinguistic, and Context) comprised a total of 176 items (the 180 items of the whole battery minus the four items of the Conversational

scale).

One of the authors or a research assistant individually administered the battery to children, in a quiet room at their school. TBI patients completed the battery at their rehabilitation centers. The administration time varied from 60 to 90 minutes; participants were allowed to take a pause, if needed. Both samples (children and TBI patients) were video-recorded during administration to permit offline scoring. Each response was assigned a score of either 0 (incorrect answer) or 1 (correct answer). Examples of test questions, participants' answers, and coding are reported in Appendix B.

3. Psychometric Analysis

3.1 Construction of Two Equivalent Forms

Two equivalent forms of ABaCo (form A and form B) were constructed following a hierarchical procedure, based on the data from the children sample. First, items within each pragmatic phenomenon in the Linguistic, Extralinguistic, Paralinguistic, and Context scales (see Appendix A) were matched to form item pairs of similar difficulty (i.e., with a similar proportion of correct answers). The similarity criterion we adopted was a difference less than or equal to .10 between the proportions of correct answers of the two items. The average difference between paired items was .04 (*SD* = .03). Sixteen items that did not meet the similarity criterion were excluded at this stage. Next, whenever multiple pairings between the remaining items within a phenomenon were possible, pairs were selected so as to maximize the tetrachoric correlation coefficients within each pair. Finally, item pairs were randomly split between the two forms. At the end of the process, each form

comprised 68 items. More specifically, the Linguistic scale of each form comprised 28 items, the Extralinguistic scale comprised 23 items, the Paralinguistic scale comprised 12 items, and the Context scale comprised 5 items. The items of the ABaCo Conversational scale were too few to construct two equivalent versions of sufficient length, and were thus excluded from the equivalent forms.

3.1.1 Internal consistency. Internal consistency (Cronbach's α) was computed for each scale, as well as for the global score (all 68 items) of the two forms. For individual scales, α ranged from .49 to .73 in form A and from .52 to .70 in form B. The internal consistency of the global score was $\alpha = .88$ in both forms (Table 1). The lowest internal consistency was shown by the Context scale; this is unsurprising, given the comparatively small number of items in this scale. In contrast, the Linguistic, Extralinguistic, and Paralinguistic scales had satisfactory values of α in both forms, and the global score had very good consistency. The average item-total correlation for individual scales ranged from .20 to .29 in form A and from .19 to .33 in form B. The average item-total correlation for the global score was .29 in form A and .28 in form B.

3.1.2 Form equivalence. The equivalence between form A and B was evaluated in five steps. To begin with, paired t-tests were performed to evaluate mean score equivalence between the two forms. The results are shown in Table 1: the difference was statistically significant only for the extralinguistic scale; however, children showed very similar mean scores (.70 in Form A and .72 in Form B) and standard deviations (.14 in Form A and .13 in Form B). Second, the mean and *SD* of each scale

and of the global score was calculated separately for the two forms. The standard error of measurement (SEM) of each score was computed as . As can be seen in Table 1, the mean difference between form A and form B scores was always smaller than the SEM; in other words, the size of the difference was below the resolution of the measurement instrument. Thus, on average, the scores obtained with form A and form B can be considered equivalent in magnitude. To further assess score equivalence at the level of individuals, we plotted score differences between the two forms against average scores (Figure 1; see Bland & Altman, 1986). Score differences were centered around zero at all ability levels, indicating the absence of any systematic distortion. Furthermore, few participants showed substantial (i.e., larger than ±2 *SD*, indicated by dashed lines in the figure) discrepancies between form A and form B scores.

We then computed the correlations between scores in form A and B (diagonal cells of Table 2). Pearson's correlations between individual scales ranged from .63 to .81, thus indicating a satisfactory to good level of convergence between the two versions; all correlations were significant ($p = .01$). Predictably, the lowest correlations were those between the two versions of the Context scale. The correlation between the global scores of form A and B was .88. Finally, we computed correlations between scales within each of the two forms (Table 2). Correlations were remarkably similar across form A and B; in all but two cases, the absolute difference was less than .10, and it never exceeded .17. This suggests that the overall structure of the ABaCo scales was roughly equivalent in the two forms we constructed.

3.2 Evaluation of the Equivalent Forms in the TBI Sample

After sorting the ABaCo items in two equivalent forms and assessing their equivalence with the data of the children sample, we tested the two forms in a sample of 30 TBI patients (see the Methods section). The performance of TBI patients was (on average poorer) than that of children; as a result, we expected the psychometric characteristics of the two forms (e.g., internal consistency, between-form correlation) to improve in the TBI sample compared with the children sample, reflecting the lower number of items eliciting ceiling performance from participants.

3.2.1 Internal consistency. In the TBI sample, α ranged from .35 to .83 in the scales of form A and from .59 to .88 in the scales of form B. The internal consistency of the global score was .92 in form A and .95 in form B (Table 3). As in the children sample, the lowest internal consistency was shown by the Context scale. The internal consistency of the other three scales, however, was satisfactory to good, and that of the global score was excellent. The average item-total correlation ranged from .18 to .39 for the scales of form A, and from .30 to .46 for the scales of form B. The average item-total correlation for the global score was .36 in form A and .44 in form B.

3.2.2 Form equivalence. As shown in Table 3, none of the t-tests was statistically significant. Moreover, the mean difference between form A and form B scores was always smaller than the SEM (Table 3). Bland-Altman plots of individual-level discrepancies between the two forms are shown in Figure 2. Again, score differences were centered around zero at all ability levels, indicating the absence of any systematic distortion with dashed lines indicating a difference of 2 standard deviations. Substantial discrepancies between form A and form B scores were absent,

with the exception of four participants on the Context scale.

Pearson's correlations between individual scales in form A and B ranged from .48 to .90 (Table 4), indicating a good level of convergence with the only exception of the Context scale; however all correlations were significant ($p = .05$). The correlation between the global scores of form A and B was .92. Again, correlations between scales within each of the two forms (Table 4) were quite similar across form A and B (the largest absolute difference was .11), despite the small number of TBI participants. In other words, the overall structure of the ABaCo scales was remarkably stable across the two forms.

4. Discussion

The aim of the present paper was to develop and test two equivalent forms of the Assessment Battery for Communication (ABaCo; Sacco et al., 2008), a clinical battery for the evaluation of pragmatic abilities in patients affected by neuropsychological and psychiatric disorders.

The two forms were constructed using the data from a large sample of children, and showed good psychometric performance in a sample of 30 patients with traumatic brain injury (TBI). Specifically, the Linguistic and Extralinguistic scales had good internal consistency and showed high correlations between forms; the global score (obtained by averaging all 68 items of the equivalent forms) also had excellent psychometric properties. The Paralinguistic scale showed a satisfactory performance, whereas the Context scale (a short scale composed of 5 items) had comparatively low consistency and correlations between forms. The means and standard deviations of the four scales and the global score were equivalent across the two forms, well within

the measurement precision of the scales.

Based on this pattern of results, we suggest that the equivalent forms of the ABaCo should be used in a flexible, context-sensitive way. The Context scale should only be used in large samples, and in research contexts that do not focus on individual cases. In small samples, its limited reliability and equivalence across forms would make it considerably less useful. In contrast, the Linguistic and Extralinguistic scales are suitable for use in small samples, and can provide meaningful information about individual cases as well. Finally, when researchers/clinicians are not interested in fine-grained differentiation between pragmatic domains, and are dealing with small samples or even single cases, we recommend using the global score of the ABaCo as a psychometrically robust summary of pragmatic ability. It is important to note at this point that the clinical use of the ABaCo is twofold. First, it can be treated as a tool for the evaluation of communicative rehabilitation programs, by using the equivalent forms at two different points in time, typically before and after the clinical intervention. Second, the ABaCo can be used as a comprehensive battery for the evaluation of pragmatic impairment in patients with communicative disorders during the clinical assessment phase. In this contexts, it is obviously preferable to employ the whole battery, for which normative data are available (Angeleri et al., 2012).

In summary, the equivalent forms of the ABaCo have the potential to become a useful tool in intervention studies. The ABaCo was originally created for use with Italian speakers. However, several parts of the battery (i.e. all parts that do not include spoken video-clips) have been already adapted for English speakers: specifically, the linguistic comprehension and production of BSAs, the extralinguistic scale, the paralinguistic scale, and the conversational scale.

We are currently using the forms in a pretest-posttest study of a pragmatic rehabilitation intervention in a group of traumatic brain injured and schizophrenic patients; the results will provide additional evidence concerning the psychometric functioning of the equivalent forms, as well as their usability in a clinical context. Finally, we hope that our effort will stimulate other research groups to develop equivalent forms of the most widely adopted measures of communicative and cognitive abilities.

Acknowledgements

This research was supported by Regione Piemonte, in particular Francesca Bosco and Bruno Bara were supported by the project “Institutions, Behaviour and Markets in Local and Global Settings (IIINBEMA)”, and Romina Angeleri and Katiuscia Sacco by the project “The Invisible Grammar of Social Relations”. We thank Silvia Testa for her help with the revision of the manuscript. Special thanks go to Marco Del Giudice for his helpful suggestions and for his critical reading of the manuscript.

References

Angeleri, R., Bosco, F. M., Zettin, M., Sacco, K., Colle, L., & Bara, B. G. (2008). Communicative impairment in traumatic brain injury: A complete pragmatic assessment. *Brain and Language, 107,* 229-245.

Angeleri, R., Bosco, F. M., Gabbatore, I., Bara B. G., & Sacco, K. (2012). Assessment Battery for Communication (ABaCo): Normative data. *Behavior Research Methods*. Advance on line publication. DOI: 10.3758/s13428-011-0174-9.

Bara, B. G. (2010). *Cognitive Pragmatics.* Cambridge, MA: MIT Press.

Bara, B. G., Bosco, F. M., & Bucciarelli, M. (1999). Developmental pragmatics in normal and abnormal children. *Brain and Language, 68,* 507-528.

Bara, B. G., Cutica, I., & Tirassa, M. (2001). Neuropragmatics: Extralinguistic communication after closed head injury. *Brain and Language, 77,* 72-94.

Bara, B. G., Tirassa, M., & Zettin, M. (1997). Neuropragmatics: Neuropsychological constraints on formal theories of dialogue. *Brain and Language, 59,* 7-49.

Bazin, N., Sarfati, Y., Lefrère, F., Passerieux, C., & Hardy-Baylè, M. (2005). Scale for the evaluation of communicative disorders in patients with schizophrenia: A validation study. *Schizophrenia Research, 77,* 75-84.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, i,* 307-310.

Bosco, F. M., & Bucciarelli, M. (2008). Simple and complex deceits and ironies. *Journal of Pragmatics, 40,* 583-607.

Bosco, F. M., Angeleri, R., Colle, L., Sacco, K. & Bara, B. G. (2012). *Communicative abilities in children: An assessment through different phenomena and*

expressive means. Manuscript submitted for publication.

Bosco, F. M., Bucciarelli, M., & Bara, B. G. (2004). The fundamental context categories in understanding communicative intention. *Journal of Pragmatics, 36*, 467- 488.

Bosco, F. M., Bucciarelli, M., & Bara, B. G. (2006). Recognition and repair of communicative failures: A developmental perspective. *Journal of Pragmatics, 38*, 1398-1429.

Brownell, H. H., Simpson, T. L., Bihrle, A. M., Potter, H. H., & Gardner, H. (1990). Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia, 28*, 375-383.

Brownell, H., & Stringfellow, A. (1999). Making requests: Illustrations of how right-hemisphere brain damage can affect discourse production. *Brain and Language, 68*, 442-465.

Brownell, H., Pincus, D., Blum, A., Rehak, A., & Winner, E. (1997). The effects of right-hemisphere brain damage on patients' use of terms of personal reference. *Brain and Language, 57*, 60-79.

Bryan, K. (1995). *The right hemisphere language battery*. London, UK: Whurr Publishers.

Corcoran, R. (2003). Inductive reasoning and the understanding of intention in schizophrenia. *Cognitive Neuropsychiatry, 8*, 223-235.

Cummings, L. (2007). Pragmatics and adult language disorders: Past achievements and future directions. *Seminars in Speech and Language, 28*, 96-110.

Davis, G. A. (2007). Cognitive pragmatics of language disorders in adults. *Seminars in Speech and Language, 28*, 111-121.

De Renzi, E., & Vignolo, L. A. (1962). The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain, 85,* 665–678.

De Vellis, R. F. (1991). *Scale development: Theory and applications.* Thousand Oaks, CA: Sage Publications.

Docking, K., Murdoch, B. E., & Jordan, F. M. (2000). Interpretation and comprehension of linguistic humor by adolescents with head injury: A group analysis. *Brain Injury, 14,* 89-108.

Dollaghan, C. A. (2004). Evidence-based practice in communication disorders: What do we know, and when do we know it? *Journal of Communication Disorders, 37,* 391-400.

Folstein, M., Folstein, S., & McHugh, P. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinicians. *Journal of Psychiatric Research, 12,* 189–198.

Frattali, C. M., Thompson, C. K., Holland, A. L., Wohl, C. B., & Ferketic, M. M. (1995). *American speech-language hearing association functional assessment of communication skills for adults*. Rockville, MD: ASHA Fulfillment Operations.

Galski, T., Tompkins, C., & Johnston, M. V. (1998). Competence in discourse as a measure of social integration and quality of life in persons with traumatic brain injury. *Brain Injury, 12,* 769-782.

Gardner, H., & Brownell, H. H. (1986). *Right hemisphere communication battery.* Boston, MA: Psychology Service, VAMC.

Grice, H. P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Holland, A. L. (1980). *Communicative abilities in daily living.* Baltimore, MD: University Park Press.

Holland, A. L., Frattali, C., & Fromm, D. (1998). *Communication activities of daily living (CADL-2).* Austin, TX: Pro-Editor.

Hollingshead, A. B. (1975). *Four factors index of social status*. New Haven, CT: Yale University Press.

Kaplan, J. A., Brownell, H. H., Jacobs, J. R., & Gardner, H. (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain and Language, 38,* 315-333.

Kasher, A. (1991). Pragmatics and Chomsky's research program. In A. Kasher (Ed.), *The Chomskyan turn* (pp. 122–149). Oxford, UK: Blackwell.

Lakoff, R. (1973). The logic of politeness; or, minding your P's and Q's. In C. Corum, T. C. Smith-Stark, and A. Weiser (Eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, Chicago, IL: University of Chicago Press.

Landis, J. R. & Koch, G. G. (1977). The Measurement of Observed Agreement for Categorical Data. *Biometrics*, 33, 159-174.

Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2002). Disturbed communication in schizophrenia: The role of poor pragmatics and poor mind-reading. *Psychological Medicine, 32*, 1273-1284.

Leroy, F., Pezard L., Nandrino, J. L., & Beaune, D. (2005). Dynamical quantification of schizophrenic speech 2005. *Psychiatry Research, 133*, 159-171.

Levinson F. C. (1983). *Pragmatics*. Cambridge, UK: Cambridge University Press.

Marquardt, T. P., Rios-Brown, M., Richburg, T., Seibert, L. K., & Cannito, M. P.

(2001). Comprehension and expression of affective sentences in traumatic brain injury. *Aphasiology, 15,* 1091–1101.

McDonald, S. (1993). Viewing the brain sideways? Right hemisphere versus anterior models of non-aphasic language disorders. *Aphasiology, 7,* 535-549.

McDonald, S., & Van Sommers, P. (1993). Pragmatic language skills after closed head injury: Ability to negotiate requests. *Cognitive Neuropsychology, 10,* 297-315.

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *Journal of Head Trauma Rehabilitation, 18,* 219-238.

Meilijson, S. R., Kasher, A., & Elizur, A. (2004). Language performance in chronic schizophrenia: A pragmatic approach. *Journal of Speech, Language, and Hearing Research, 47,* 695-713.

Nunnaly, J. (1978). *Psychometric theory*. New York, NY: McGrow-Hill.

Pell, M. D. (2007). Reduced sensitivity to prosodic attitudes in adults with focal right hemisphere brain damage. *Brain and Language, 101,* 64-79.

Penn, C. (1985). The profile of communicative appropriateness: A clinical tool for the assessment of pragmatics. *South African Journal of Communication Disorders, 32,* 18-23.

Prutting, C. A., & Kirchner, D. M. (1987). A clinical appraisal of the pragmatic aspects of language. *Journal of Speech and Hearing Disorders, 52,* 105-119.

Raykow, T., & Marculides, G. A. (2009). *Introduction to Psychometric Theory*. New York, USA/London, UK: Routledge

Russeaux, M., Vérigneaux, C., & Kozlowski, O. (2010). An analysis of

communication in conversation after severe traumatic brain injury. *European Journal of Neurology, 17,* 922-929.

Sacco, K., Angeleri, R., Bosco, F. M., Colle, L., Mate, D., & Bara, B. G. (2008). Assessment Battery for Communication - ABaCo: A new instrument for the evaluation of pragmatic abilities. *Journal of Cognitive Science, 92,* 111-157.

Sarno, M. T. (1969). *The functional communication profile*. New York, NY: Institute of Rehabilitation Medicine, NYU Medical Center.

Sohlberg, M. M., & Mateer, C. A. (2001). *Cognitive rehabilitation: An integrative neuropsychological approach*. New York, NY: The Guilford Press.

Stemmer, B., Giroux, F., & Joanette, Y. (1994). Production and evaluation of requests by right hemisphere brain-damaged individuals. *Brain and Language, 47,* 1-31.

van der Meulen, I., van de Sandt-Koenderman, W. M. E., Duivenvoorden, H. J., & Ribbers, G. M. (2010). Measuring verbal and non verbal communication in aphasia: Reliability, validity, and sensitivity to change of Scenario Test. *International Journal of Language and Communication Disorders, 45,* 424-435.

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language, 62,* 89-106.

Appendix A

Structure and sample items of the equivalent forms of ABaCo

Scale	Sub-scale	Pragmatic task	Example	Test questions	No. of items in each form
Linguistic	Comprehension	Basic Speech Act	In vivo item	Where do you live?	8
		Standard	Videotaped scene - Man: "This pasta is very good. Who cooked it?" Woman: "I did."	What did the woman say?	2
		Deceit	Videotaped scene -The child knocks a vase over. Mum: "Who knocked the vase over?" Child: "It was the doggie."	What did the child say?	2
		Irony	Videotaped scene - In a shop, Lara tries on a dress that is clearly too tight and asks Simone: "Does this dress fit me?" Simone answers: "Your diet is working well!"	What did the boy say?	2
	Production	Basic Speech Act	In vivo item	Ask me whether I've got children.	8

		Standard	Videotaped scene - Husband and wife are sitting on the sofa. Wife: "What would you like to do this afternoon?"	What could the man answer?	2
		Deceit	Videotaped scene - Matteo enter to the dining room, sees an orange juice on the table, and drinks it. Elisa arrives and asks: "Who drank my orange juice?"	The boy doesn't want to be scolded. What can he reply?	2
		Irony	Videotaped scene - Brother and sister are having breakfast. He's put is elbow in the jam. Brother: "Can you pass me the jam, please?"	What could the girl answer to make fun of the boy?	2
Extralinguistic	Comprehension	Basic Speech Act	Videotaped scene - The guy performs a gesture meaning "Go out!"	What did he tell you?	6
		Standard	Videotaped scene - The wife has got a dish of steaming soup. She nods to her husband with a gesture meaning "Are you coming?" The husband nods yes.	What did the man say?	1
		Deceit	Videotaped scene - The boy performs a gesture with which he asks for some candies. The girl doesn't want to give him any	What did the girl say?	1

			candy. So she looks at the candies with disgusted expression, which means "They are awful!"		
		Irony	Videotaped scene - Boy and girl are tasting some disgusting soup. The boy smacks his lips with a gesture meaning "It's very good!"	What did the boy say?	2
	Production	Basic Speech Act	In vivo item	Order me to be quite.	8
		Standard	Videotaped scene - A man needs help in the street. He sees a car coming.	What gesture can the man perform?	2
		Deceit	Videotaped scene - The boy throws a dish of vegetables in the bin. Mum comes in and performs a gesture meaning "Have you already finished your vegetables?"	The boy doesn't want to be scolded. What gesture can he perform?	1
		Irony	Videotaped scene - Lisa and Giovanni are in the kitchen emptying their shopping bags. Giovanni absent-mindedly drops an egg he was about to put away. The egg breaks, making a mess on the table…	Imagine the girl wants to make fun of the boy. What gesture can she make?	2
Paralinguistic	Comprehension	Basic Communication Act	Videotaped scene - The man makes a request in an invented language	The man • said what he thinks • made a request	2

				• gave an order • told a lie	
		Basic Emotions	Videotaped scene - The woman screams and gesticulates angrily	The woman is • angry • sad • happy • embarassed	2
		Contradiction	Videotaped scene - Girl: “Did you like the cake?” Boy, with disgusted expression: “Yes, very good.”	What did the boy say?	2
	Production	Basic Communication Act	In vivo item	Ask me to give you a pen.	2
		Basic Emotion	In vivo item	Ask me where the doctor is • acting sad • acting happy	4
Context	Comprehension	Grice’s Norms	Videotaped scene - Sister: “Where did you put my diary?” Brother, in front of a red chest of drawers: “In the red drawer.”	Is the answer ok?	2
		Social Norms	Videotaped scene - Head office: “Miss, can you type this letter, please?” Secretary, angrily: “I cannot do it now! I’ve got so much work!”	Was the secretary polite?	1

Production	Social Norms	In vivo item	Imagine you are late for an appointment: • with your lawyer • with a friend of yours How would you apologize for being late?	2

Appendix B
Examples of Participants' Answers

LINGUISTIC SCALE - COMPREHENSION

[1] DECEIT

Videotaped scene content: Andrew is eating some biscuits. He hears Julia arriving, and then he pushed away the empty plate in front of him. Julia sees the empty plate and asks: "Who has finished my biscuits?".
Andrew answers: "I don't have the slightest idea."

Test question: In your opinion, why did the boy answer to the girl this way?

Example of correct answer: "He's telling a lie to his sister."

Example of wrong answer: "Because he doesn't know."

EXTRALINGUISTIC SCALE - COMPREHENSION

[2] IRONY

Videotaped scene content: Piero and Alice are in the kitchen. Alice gets up to fetch a pan, which she brings to the table, and pours a ladle of soup into the dishes. They taste a spoonful and both pull a disgusted face. Alice looks questioningly at Piero that takes his fingers to his mouth and kisses his fingertips with an expression as if to say "Delicious!"

Test question: In your opinion, why did the boy answer to the girl with that gesture?

Example of correct answer: "He was joking! He doesn't like the soup!"

Example of wrong answer: "Because he like the soup."

PARALINGUISTIC SCALE - PRODUCTION

[3] BASIC EMOTION

Examiner: Ask me where the teacher is. Ask me as if you were sad.
Examiner: Ask me where the teacher is. Ask me as if you were scared.

Example of correct answer: The subject uses the appropriate intonation in proffering the utterance.

Example of wrong answer: The subject does not use the appropriate intonation in

proffering the utterance.

CONTEXT SCALE - COMPREHENSION
[4] GRICE'S NORMS
Videotaped scene content: Sara asks Giorgio: "What time is it?" Giorgio replies: "I think today is Monday."

Test question: Is the answer ok?

Example of correct answer: "Of course not, she was talking about another thing…"

Example of wrong answer: "Yes, it was Monday."