The Spanish Journal of Psychology (2017), **20**, e33, 1–12.

doi:10.1017/sjp.2017.31

Presentation of the Coding and Assessment System for Narratives of Trauma (CASNOT): Application in Spanish Battered Women and Preliminary Analyses

Violeta Fernández-Lansac and María Crespo

Universidad Complutense (Spain)

Abstract. This study introduces a new coding system, the Coding and Assessment System for Narratives of Trauma (CASNOT), to analyse several language domains in narratives of autobiographical memories, especially in trauma narratives. The development of the coding system is described. It was applied to assess positive and traumatic/negative narratives in 50 battered women (trauma-exposed group) and 50 nontrauma-exposed women (control group). Three blind raters coded each narrative. Inter-rater reliability analyses were conducted for the CASNOT language categories (multirater K_{free} coefficients) and dimensions (intraclass correlation coefficients). High levels of inter-rater agreement were found for most of the language domains. Categories that did not reach the expected reliability were mainly those related to cognitive processes, which reflects difficulties in operationalizing constructs such as lack of control or helplessness, control or planning, and rationalization or memory elaboration. Applications and limitations of the CASNOT are discussed to enhance narrative measures for autobiographical memories.

Received 15 July 2016; Revised 20 March 2017; Accepted 9 June 2017

Keywords: autobiographical memory, coding system, judge-based thematic content analyses, language domain, traumatic narrative.

The language that we use shapes our perceptions and our way of thinking and facing life. Therefore, language is a window to the cognitive and emotional world of individuals and communities (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007). In psychology, the analysis of language is considered to be the main tool to access the patient's subjectivity; and narrative research has studied the relationships between language use and mental experiences, with particular attention to the linguistic elements involved in identity construction (see Forgas, Vincze, & László, 2014). Consequently, many analysis methods have been developed in order to assess a great number of linguistic aspects, or language domains, in narratives related to psychological processes.

General approaches in the study of language

Given the complexity of language, early narrative research aimed to explore the meaning of texts or utterances in context, by employing *qualitative approaches* (Forgas et al., 2014; Pennebaker, Mehl, & Niederhoffer, 2003). Overall, qualitative research encompasses a wide range of methods whose primary objective is to capture the particular characteristics of human experience through an intensive exploration of the individual reality (Polkinghorne, 2005). Qualitative methods are unique in its flexibility and capacity to capture complex meanings; however, they usually involved a small sample of participants, can be time-consuming, and are often considered to have an inappropriate lack of scientist rigour (Griffin, 2004). Because of these limitations, more recent approaches to language study have sought to use quantitative forms of text analyses (see Popping, 2000). *Quantitative approaches* are based on computerized counts, statistical analyses, and classification of words; and can be more efficient and easy-to-use tools to evaluate narratives at both thematic and semantic levels (Mehl, 2006).

Among quantitative approaches, computerized text analysis methods are becoming increasingly popular. These can be categorized into *word pattern analyses* and *word count strategies* (see Chung & Pennebaker, 2011; Pennebaker et al., 2003). Word pattern analyses explore how words covary across texts by employing a "bottom-up" strategy. For example, the Latent Semantic Analysis (LSA; e.g., Foltz, 1996) measures the semantic similarity between two texts by identifying relationships between the terms used. On the other hand, word count strategies offer information about the frequency in which categories of words are used in a text or narrative. Some examples are the General Inquirer,

Correspondence concerning this article should be addressed to María Crespo. Universidad Complutense de Madrid - Clinical Psychology. Campus de Somosaguas, s/n. 28223. Madrid (Spain).

E-mail: mcrespo@psi.ucm.es

Universidad Complutense de Madrid Predoctoral Scholarship granted to Violeta Fernández-Lansac. Spanish Ministerio de Economía y Competitividad PSI2012–31952.

created by Stone's team in the 1960s (Stone, Dunphy, Smith, & Ogilvie, 1966), and the Linguistic Inquiry and Word Count (LIWC) by Pennebaker, Francis, and Booth (2001), which is currently the most extensively used. The LIWC uses an external dictionary to calculate the percentage of words within a text that belong to several language domains, such as affective processes, cognitive mechanisms, and references to the past, present, and future. It has been proven useful to detect linguistic patterns associated to a great amount of psychological and social dimensions, but like other computer programs, it has the disadvantage of being independent of the semantic context of language (see Pennebaker et al., 2003; Tausczik & Pennebaker, 2010).

A third approach, the *judge-based thematic content analyses*, straddles the border between quantitative and qualitative approaches. It is typically based on the assessment by trained raters of the presence in a text of critical themes, using language categories (usually scored as presence/absence) and language dimensions (scored in numerical scales), previously determined in established coding system (Smith, 2000). A paradigmatic example of this approach is the Thematic Apperception Test (TAT) by Morgan and Murray (1935), for which coding systems have been developed to identify different themes (e.g., need for achievement or power) in histories elaborated by participants about ambiguous pictures. In the same way, several thematic content analyses have been applied in various areas of behaviour and social sciences (see Smith, 1992). The main advantage of this kind of analysis is that, although it allows to handle quantitative data, captures better than quantitative approaches the complexity and contextual organization of human experience (see Smith, 2000). Also, inter-rater comparisons provide the opportunity to explore the reliability of the coding system, to examine its clarity and, if necessary, to revise theoretical bases and operational definitions (Joffe & Yardley, 2004). However, judge-based thematic content analyses have been criticised by proponents of computerized text analyses, arguing that they are highly consuming of time and effort, and are not free of subjective biases (e.g., Chung & Pennebaker, 2011).

Language analyses of traumatic narratives

One of the most common applications of language analysis, from both qualitative and quantitative approaches, is the study of autobiographical memories in clinical contexts, and, more specifically, the study of trauma memories. Papini, Yoon, Rubin, López-Castro, and Hien (2015) state that "in the aftermath of trauma, language provides a distinct window into how survivors process unfathomable events" (Papini et al., 2015, p. 295). Consequently, several linguistic measures have been used to explore structural and content features of narratives about traumatic experiences in order to assess differential linguistic elements related to the adaptation after trauma and the development of posttraumatic stress disorder (PTSD) (e.g., Dekel & Bonanno, 2013; Eid, Johnsen, & Saus, 2005; see Crespo & Fernández-Lansac, 2016; O'Kearney & Perrott, 2006).

In most cases, narrative studies about trauma have been based on both qualitative and judge-based thematic content analyses, which have addressed a mix of language domains related to the memory content. In the earliest studies, researchers attempted to identify recurrent topics or themes among trauma memories. For example, Lisak (1994) analysed transcribed interviews by 26 male survivors of childhood sexual abuse according to 15 psychological themes, such as anger, fear, shame and humiliation, betrayal, isolation and alienation, and negative schemas about the self. Other studies have explored the presence of linguistic elements of coherence, finding meaning, self-evaluation, appraisal or coping behaviours, as well as changes in the emotional expression along therapy in trauma narratives (e.g., Murray, Lamnin, & Carver, 1989; Tuval-Mashiach et al., 2004). Additionally, different coding systems have been developed to assess a great number of language categories and dimensions such as emotional valence (e.g., Sutherland & Bryant, 2005); autonomous orientation, other/self ratio, interaction scenario, reflective comments, memory specificity, and self-defining memories (e.g., Jobson & O'Kearney, 2006; Sutherland & Bryant, 2005); perceptual vs. conceptual memory representations (e.g., Buck, Kindt, van den Hout, Steens, & Linders, 2006); dissociation and confusion (e.g., Jones, Harvey, & Brewin, 2007); and vividness and sense of reliving (e.g., Hagenaars, van Minnen, & Hoogduin, 2009).

With regard to the narrative structure, trauma research has mainly focus on the assessment of fragmentation or narrative coherence, because an inability to construct an organized narrative has been traditionally linked to PTSD (e.g., Ehlers & Clark, 2000; Foa & Riggs, 1993). Therefore, complex coding systems have been also developed to quantify heterogeneous indicators of narrative fragmentation, such as repetitions or unfinished thoughts (e.g., Foa, Molnar, & Cashman, 1995), or to obtain total scores for disorganization and coherence (e.g., Freer, Whitt-Woosley, & Sprang, 2010; Hagenaars et al., 2009; Halligan, Michael, Clark, & Ehlers, 2003).

Finally, together with both qualitative and judge-based thematic content analyses, the use of computerized programs (especially, the LIWC) is making headway in traumatic narrative research. In last years, several researchers have explored a number of LIWC domains that have proved to reveal important information about

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

the psychological processes involved in memory elaboration and trauma adaptation (i.e., affective, cognitive, and sensory/perceptual processes, among others), sometimes including additional strategies for thematic content analysis (e.g., D'Andrea, Chiu, Casas, & Deldin, 2012; Römisch, Leban, Habermas, & Döll-Hentschker, 2014).

In sum, narrative research of traumatic memories has combined quantitative and qualitative approaches to assess a broad range of language domains in traumatic narratives. Since both approaches present limitations, the adequacy of the methodology depends on the researcher's objectives. Computerized text analyses can be powerful tools, especially for large samples. Nonetheless, we consider that a comprehensive understanding of the trauma experience needs to take into account the context in which language is produced. For this reason, trauma research needs to incorporate judge-based thematic content strategies, as supported by previous literature (see Crespo & Fernández-Lansac, 2016). To further investigate the main processes related to autobiographical memory construction and trauma elaboration, it is necessary to advance in the development of efficient coding systems that integrate the most important language domains identified in trauma research. These systems should be reliable enough to ensure a minimization of researcher' subjective biases, and should be generalizable to narratives of different memories and participant's samples.

The present study

The objective of the present study is to introduce an operative coding system, the Coding and Assessment System for Narratives of Trauma (CASNOT), that has been designed to evaluate a wide range of linguistic aspects in narratives in Spanish. This system relies on a judge-based thematic content approach aiming to assess the language use in context; and integrates the most relevant language domains identified in memory and trauma research. It seeks to be an easy, efficient, and reliable tool for language analysis in both autobiographical and traumatic narratives.

In order to explore the CASNOT properties, it was applied to assess autobiographical narratives with high emotional content (*positive* and *traumatic/negative*) using two different samples: 50 Spanish battered women (*trauma-exposed group*) and 50 Spanish nontrauma-exposed women (*control group*). In this paper we present the development of the CASNOT and provide information about its reliability coefficients.

Development of the Coding and Assessment System for Narratives of Trauma (CASNOT)

The development and application of the CASNOT was carried out through a series of consecutive but interactive steps according to the recommendations for the design of judge-based thematic content analyses and coding systems in social and clinical psychology (e.g., Bartholomew, Henderson, & Marcia, 2000; Chorney, McMurtry, Chambers, & Bakeman, 2015; Krippendorff, 2004; Smith, 2000). The tasks involved in developing the CASNOT are summarized in Figure 1 and will be described in detail in the following sections.

CASNOT design: Initial version

To select the linguistic domains to include in the CASNOT, we conducted an extensive literature review of previous studies that have used methods for language analyses to explore the content and structure of traumatic memories (see Crespo & Fernández-Lansac, 2016). Particular attention was paid to the existing judge-based thematic strategies developed to date in the trauma narrative area, and to the authors' suggestions for the future improvement of coding systems. LIWC domains traditionally explored in trauma research (e.g., affective processes, cognitive mechanisms, sensory/perceptual processes) were considered also as relevant issues to include. In addition, three psychologists with expertise in PTSD and psychotraumatology assessed the proposed choice of language domains according to its significance for the study of autobiographical memories and trauma research.

For the operationalization of each domain, three information sources were consulted: (a) descriptions provided in the previous literature to define psychological processes linked to traumatic memories, (b) standard definitions provided in commonly used dictionaries (e.g., Dictionary of the Royal Academy of Spanish Language –Diccionario de la Real Academia de la Lengua Española, RAE, in Spanish–; WordReference, synonym dictionaries), and (c) the judgement of the three PTSD experts involved in the language domains selection. The aim was to incorporate common and broadly understandable variable' descriptions. Moreover, according to the recommendations by Bartholomew et al. (2000), these variables must be sufficiently detailed for its identification but abstract enough to be applicable to different types of narratives.

As a result of this process, a preliminary coding system was developed for language analysis of autobiographical memories. It was designed to be applied by trained raters using both audio and transcribed versions of the participants' narratives. As detailed in instructions, raters should identify along the narratives the presence of utterances or language expressions that implied the experiencing of different emotions, sensations, and psychological processes (*categories*). Also, raters should assess the degree or intensity of various language domains (*dimensions*). Specifically, the initial CASNOT

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

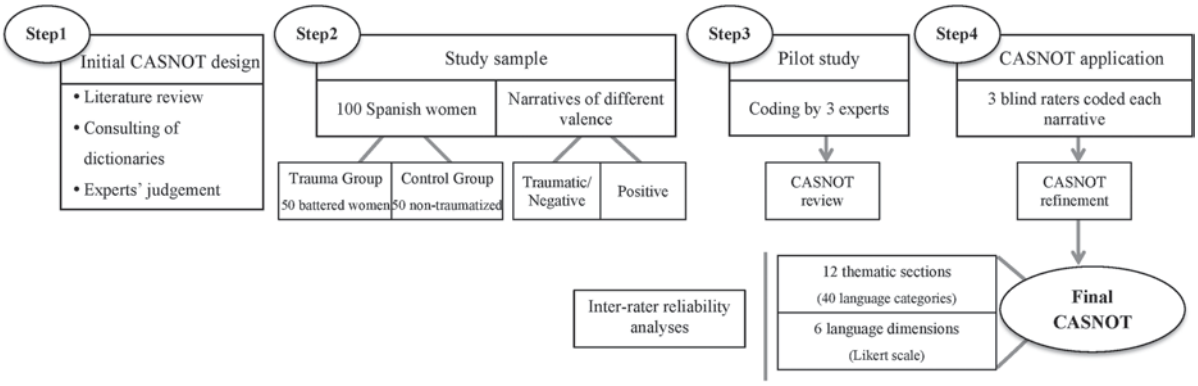

Figure 1. Steps in developing and implementing the CASNOT.

assessed both structural and content narrative aspects according to: (a) 11 thematic sections composed by 43 language categories (i.e., a section could be comprised of more than one category) that represented discrete dichotomous variables (present vs. absent; e.g., raters should identify the presence or absence of the "sadness" category, belonging to the "emotional processes" section); and (b) 8 language dimensions that implied a global evaluation of the whole narrative into an analogical Likert-type scale (e.g., raters should assess the emotional tone of the overall narrative in a continuous from 0 = completely positive to 4 = completely negative). Categories were not mutually exclusive; that is, the same language expression could be coded in different categories. The CASNOT encompassed an application handbook (with detailed instructions, and operational definitions and several examples for each category and dimension) and a number of record sheets.

Final version of the CASNOT

Pilot study

The three PTSD experts involved in the initial CASNOT design were trained and applied it to evaluate each set of narratives (positive, and traumatic or negative) by six participants from the study sample (three from the trauma-exposed group and three from the control group; the study sample is further described below). Narratives were randomly selected and were constant for the three raters. Raters were blind to the study objectives and to the other judges' assessments. In addition to the coding system, raters received a feedback form instructing them to assess the clarity of the instructions and operational definitions of categories and dimensions, identifying any problem during the task and suggesting potential language domains to include or remove. The three raters and the lead researchers met to discuss any disagreement to reach consensus about which categories and dimensions captured the narrative content and to refine its definitions. Feedback from the raters resulted in a slight reduction of the number of categories (from 43 to 41), although the total number of thematic sections as well as the number of dimensions remained the same.

CASNOT application and refinement

Subsequent application of the CASNOT was carried out in two phases. In a first phase, thirteen blind raters participated. All of them were graduated Psychology students and underwent intensive training in the use of the CASNOT. As part of this training, they had to use the system in a test application. Then, the difficulties encountered were solved, and narratives from the trauma-exposed and control groups were randomly assigned to the raters. Each set of narratives (that is, positive, and traumatic/negative narratives by the same participant) was coded by three different raters. Raters received approximately half of narratives from the trauma-exposed group and half from the control group (i.e., 23–24 narrative sets), distributed in three batches. Each rater received both the audio and the transcribed version of the narratives to assess, as well as the CASNOT manual and several record sheets. The order of positive and traumatic/negative narrative for each participant was set randomly and marked on the corresponding record sheet. Narratives were identified with numerical codes to ensure participants' anonymity, and raters signed a confidentiality commitment.

The CASNOT application was conducted between February and May 2015. Weekly meetings were established, where raters shared their difficulties and doubts. Subsequently, a further CASNOT refinement was carried out to rewrite operational definitions for problematic variables (i.e., variables with poor inter-rater reliability values according to criteria described in the data analysis section). To conduct this refinement, the three experts involved in the pilot study met with the lead researchers to reach a consensus about the changes to be made. Specifically, language categories

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

related to the "cognitive processes" section were the most problematic and, therefore, needed to be redefined in depth. In addition, the "integration" section, composed of only one category related to the impact of the event, replaced the initial "centrality" section, composed of different categories that reflected positive and negative changes in the perception of oneself, the world, and others. Among language dimensions, "wealth of details" was redefined, and "structure" and "coherence" were integrated into a single coherence dimension. Finally, "specificity", previously a dimension, was treated as a thematic section, which comprised four different categories. The redefined language categories and dimensions were tested in a second application phase (from December 2015 to January 2016), in which nine out of the thirteen initial raters participated. After a training session, narratives were randomly distributed among raters to recode only the refined categories and dimensions.

The final CASNOT version consisted of: (a) 40 dichotomous language categories, distributed in 12 thematic sections (see Table 1); and (b) 6 continuous language dimensions (see Table 2). Among categories, only those that belonged to the "specificity" section were mutually exclusive. A detailed description of the CASNOT categories and dimensions can be seen in the Appendix[1].

RELIABILITY OF THE CASNOT

Method

Study sample

The sample (n = 100) consisted of 50 battered women (trauma-exposed group; ages 21-60) recruited through clinics and centres for assistance to women in the Madrid area (Spain), and 50 nontrauma-exposed women (control group; ages 20–73) from the general population. All of the participants were fluent in Spanish. Women from the trauma-exposed group had experienced violence from their intimate partner for at least 1 month. Further details about the sample have been previously published (Fernández-Lansac & Crespo, 2015).

Procedure

Trauma-exposed participants were asked to narrate their most distressing episode of abuse (traumatic event), and control group participants their most stressful life event (negative event). Both groups also described the happiest event of their lives (positive event). Previously, to familiarize the participants with the narrative task, they had to narrate a normal day in their life as a neutral event. The instructions were as follows:

> Please tell me about your daily activities for today/the happiest event/the worst event (or worst episode of violence). Tell me as you remember it, in the most detailed way possible. I would like you to tell me what happened just as it occurs in your mind. Tell me like I was a friend, and try to describe everything that you can remember: what took place, what you did, what you felt, what you thought… During your description, I will not interrupt you. You can say whatever you want, with absolute freedom, and you can take as much time as you need. We will start and finish whenever you want.

The order of episodes (traumatic/negative and positive) was randomly counterbalanced across participants, although the neutral narrative was always the first. Narratives were uninterrupted, audio-recorded, and then transcribed verbatim. This procedure was previously tested in a pilot study (n = 5).

The first author interviewed all of the women from the trauma-exposed group. Women from the control group were interviewed by female Master's psychology students, who were intensively trained for this purpose. Participation in the study was voluntary and was carried out after participants were informed about the procedure and data confidentiality, and after their written consent was obtained. The Ethics Committee of the university approved this study.

Within the trauma-exposed group, 62.0% of the women had experienced physical aggressions in the worst episode of abuse (i.e., traumatic event), 12.0% sexual abuse, and all of them psychological abuse (100.0%). The control group selected as the worst life events the loved one's death (58.0%) or illness (16.0%), an experience of family abuse (6.0%), or job problems (4.0%), among others. Positive events that were selected included the pregnancy or birth of a child (50.0% of trauma-exposed group participants and 32.0% of controls), a happy moment shared with children or other loved ones (28.0% and 16.0%, respectively), the wedding day (4.0% and 10.0%), or a professional or academic achievement (2.0% and 22.0%). Among the trauma-exposed participants, 8.0% of them selected as their happiest moment the separation or divorce from their abuser.

Data analysis

SPSS version 20 was used to perform statistical analysis. Because three independent raters coded each narrative, final scores were calculated using the mean for language

[1]Application handbook and record sheets of the final Spanish version of the CASNOT are freely available from the authors upon request.

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

Table 1. *Inter-rater reliability of the CASNOT language categories (n = 100)*

		Kappa	
Section	Category	P (+)	T/N (-)
Emotional processes	Sadness	.89	.51
	Fear	.93	.71
	Guilt	.97	.77
	Humiliation/Shame	.96	.69
	Anger towards oneself	.99	.81
	Anger outward	.99	.68
	Other negative emotions	.72	.29
	Positive emotions	.85	.71
	Surprise	.68	.44
	Empathy/Concern for the aggressor	—	.69
	Empathy/Concern for other/s	.44	.47
Bodily states and symptoms	Anxiety	.93	.81
	Bodily states	.69	.51
	Dissociative experiences	.99	.83
Sensory/ perceptual processes	Visual perception	.69	.52
	Auditory perception	.79	.45
	Other sensations	.71	.71
Cognitive processes	Lack of control/Helplessness	.77	.37
	Control/Planning	.81	.37
	Insight	.63	.48
	Rationalization/Elaboration	.36	.28
References to other people	Social support	.48	.53
	Lack of social support	.92	.68
Self-evaluations	Negative self-evaluation	.93	.71
	Positive self-evaluation	.75	.87
Meaning	Spirituality	.96	.85
	Meaning	.84	.76
Integration	Impact	.56	.49
Overcoming	Overcoming/Hope	—	.83
	Forgiveness	—	.92
Memory consistency	Uncertainty	.73	.52
	Certainty	.61	.56
Other aspects	Threats to physical integrity or life	.95	.65
	Other references to death	.93	.60
	Escapes/avoidance	.95	.84
	Allusions to the negative event	.83	—
Specificity	Extended event	.64	.43
	Categorical event	.72	.97
	Different specific events	.79	.81
	One specific event	.65	.51

Note: P (+) = positive narratives; T/N (-) = traumatic/negative narratives.

The categories empathy/concern for the aggressor, overcoming/hope, and forgiveness only were assessed in traumatic/negative narratives; the category allusions to the negative event only was assessed in positive narratives.

dimensions (continuous variables), and the score (0 for absence and 1 for presence) most commonly assigned for language categories (i.e., final score assigned to a category was based on the score by at least two raters).

To explore psychometric properties of the CASNOT, three inter-rater reliability measures were used for both positive [P (+)] and negative or traumatic narratives [T/N (-)], separately. First, exploratory analyses were conducted to assess the percentage of absolute agreements for language categories. Absolute agreement was defined as the percentage of times that the three raters provided identical evaluations (i.e., 0 or 1) divided by the total number of ratings.

Second, the Free-marginal Multirater Kappa (multirater $K_{free;}$ Randolph, 2005) was also calculated for each of the language categories. The multirater K_{free} is

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

Table 2. *Inter-rater reliability of the CASNOT language dimensions (n = 100)*

	ICC	
Dimension	P (+)	T/N (-)
Emotional tone	.70	.75
Emotional valence	.53	.60
Wealth of details	.61	.56
Spatial orientation	.85	.78
Temporal orientation	.76	.69
Coherence	.56	.60

Note: P (+) = positive narratives; T/N (-) = traumatic/negative narratives.

an alternative to Fleiss's Kappa, a variant of Cohen's Kappa, which is a chance-adjusted index of agreement for nominal variables and any number of raters (Fleiss, 1971). Since Fleiss' Kappa is affected by prevalence, low kappa and high inter-rater agreement could paradoxically occur. Alternatively, multirater K_{free} is appropriate for studies that have free-marginal distributions, such as the present study, in which there are no restrictions in the distribution of judgements over categories. Values of kappa range from –1.0 (perfect disagreement) to 1.0 (perfect agreement about chance). According to Landis and Koch (1977), kappa coefficients above .40 are considered acceptable, and values above .75 are considered excellent.

Finally, inter-rater reliabilities for language dimensions were calculated using one-way random, average-measures Intraclass Correlation Coefficients (ICC) for both P (+) and T/N (-) narratives. ICC is suitable for ordinal and interval ratio variables and for two or more raters (Shrout & Fleiss, 1979). Cicchetti (1994) cutoffs were adopted, with values above .40 considered to be fair, values between .60 and .74 considered to be good, and values above .75 considered to be excellent. It should be note that usually in narrative studies greater reliability values are expected, however some authors claim that no language assessment method can reliably capture psychological qualities at rates higher than 80.0% (e.g., Chung & Pennebaker, 2011).

Results

Overall, the percentage of absolute agreement for language categories was 78.2%. Considering both P (+) and T/N (-) narratives together, absolute agreement was 79.2% for "emotional processes", 84.5% for "bodily states and symptoms", 73.3% for "sensory/perceptual processes", 63.2% for "cognitive processes", 74.0% for "references to other people", 86.0% for "self-evaluations", 89.0% for "meaning", 64.5% for "integration", 90.5% for "overcoming", 70.5% for "memory consistency", 86.6% for "other aspects", and 76.7% for "event specificity". Absolute agreement was in general lower for T/N (-) narratives than for P (+) narratives. For T/N (-) and P (+) narratives, the percentage of identical ratings was 73.4% and 82.1%, respectively. Additionally, by comparing scores for trauma-exposed and control groups, raters showed a lower agreement when assessing trauma narratives. Absolute agreement was 69.7% for T (-) narratives and 77.2% for N (-) narratives.

Table 1 shows the multirater K_{free} coefficients (Kappa) for all of the language categories, and Table 2 the ICCs for the language dimensions. As displayed in Table 1, raters achieved acceptable to excellent levels of agreement for most of the categories in both P (+) and T/N (-) narratives. Kappa coefficients were less than .40 only for the residual "other negative emotions" category and for the categories included in the "cognitive processes" section. However, the analyses showed that agreement in "other negative emotions" was poor only when raters coded T/N (-) narratives but not when they coded P (+). To examine whether inter-rater reliability for this category was affected by the group (trauma-exposed vs. control), kappa coefficients were calculated separately for T (-) and N (-) narratives. Unfortunately, the results showed that values were poor for narratives from both groups; kappa was .23 for T (-) narratives and .36 for N (-) narratives. Regarding the "cognitive processes" section, agreement between raters was satisfactory when assessing "lack of control/helplessness" and "control/planning" categories for narratives from the control group: kappa values were .25 for T (-) and .49 for N (-) for "lack of control/helplessness", and .28 and .47, respectively, for "control/planning". Nevertheless, agreement for "rationalization/elaboration" was unsatisfactory for both groups: kappa coefficients were .31 for T (-) and .25 for N (-).

Finally, as shown in Table 2, the ICCs of all of the language dimensions, for both P (+) and T/N (-) narratives, were above .50, indicating at least a moderate inter-rater agreement. Moreover, values were good for most of the dimensions, and excellent or almost excellent for "emotional tone" and "spatial" and "temporal orientation".

Discussion

The present study proposes an innovative and comprehensive coding system for analysing several structural and content aspects of narratives about autobiographical experiences, which is especially recommended for the assessment of traumatic memories. This coding system, named CASNOT, relies on the judge-based thematic content analysis. Therefore, it allows exploration of

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

a wide range of linguistic domains within the natural language context, but like methods based on quantitative approaches seeks to be an efficient and reliable tool for narrative analysis. Since the CASNOT includes a great number of language categories and dimensions related to psychological processes, it would be potentially useful in research and in clinical psychology; it can be adapted to the specific requirements of researchers and clinicians, and could become a reference for language coding systems in other idioms.

The need to develop operative and integrative measures to evaluate narrative content and structure has been a recurrent challenge in the study of autobiographical memories and trauma research. For example, some authors have noted that a mix of criteria to define narrative organization prevents definitive conclusions about the hypothesized fragmentation of traumatic memories, and terms such as coherence or cohesion are often confused (see Crespo & Fernández-Lansac, 2016; O'Kearney & Perrot, 2006). To select and define the language categories and dimensions of the CASNOT, we have taken into account authors' recommendations aiming to establish a solid and unified theoretical framework. However, one of our primary objectives has been to capture shared and common meanings (i.e., it addresses the typical way in which language is understood in its context). This goal has prevailed in the decision-making processes. For instance, we initially introduced two separate fragmentation dimensions (i.e., coherence and structure) in response to the researchers' concerns. Nevertheless, after noticing the difficulties encountered by raters in distinguishing between these two dimensions, a single coherence dimension that integrated both constructs was created. In the same way, other overlapping constructs have been merged in order to minimize discrepancies and to adopt helpful and operational definitions. On the other hand, due to the great complexity in measuring narrative aspects, dichotomous values were considered to be a good option to value language categories, since they reduce the risk of subjective bias, and the time and effort to invest in the application of the system.

The development of the CASNOT followed a series of consecutive steps to guarantee its reliability and applicability. First, together with an extended literature review, we consulted a wide variety of information resources, including experts' views. In addition, regular meetings with raters were set up to discuss any application problems and to collect any suggestions for improvement. This strategy has helped to reach an in-depth understanding of the coding system in all of our researchers, to refine the operationalization of categories and dimensions, and to ensure the standardization of the coding procedures. Second, the CASNOT has been applied on a clinical sample composed of battered women who were exposed to long lasting traumatic experiences. Further, a control group was established, and the coding system was tested on narratives about highly intense autobiographical memories with different emotional valence. Finally, inter-rater agreement analyses have been crucial to ensure the reliability of the coding system. Because agreement coefficients were provided for all the language categories and dimensions separately, and according to different narratives, data point which specific domains are more prone to subjective bias and are less generalizable to distinct types of narratives.

Results showed that the vast majority of categories and dimensions that composed the CASNOT were sufficiently reliable. Exceptions were the categories related to the identification of "other negative emotions" and of several cognitive processes (e.g., processes that reflect lack of control or helplessness feelings, or rationalization or elaboration of traumatic memories). Regarding the former, results are not surprising because it includes negative emotions that were not addressed in the previously established emotion categories; actually, "other negative emotions" is a residual category that could act as a catchall. On the other hand, identifying the presence of different cognitive processes is likely the most difficult task for the raters. Cognitive processing words have been considered to reveal narrative aspects related to conceptual organization, but language measures have often failed in capturing these processes involved in the memory construction. For example, the LIWC makes no distinction between language expressions that indicate understanding (i.e., organized thoughts) and confusion (i.e., disorganized thoughts) (Jelinek et al., 2010). Because we analysed language use in its natural context, our goal was to determine which specific thoughts emerge in trauma narratives, paying attention to cognitive processes that reflect attempts to cope with and understand the traumatic situation. In the first application phase of the CASNOT, categories from the "cognitive processes" section obtained low levels of agreement among raters and were consequently redefined. Specifically, categories such as lack of control and helplessness should merge, and categories related to internal and external attributions should be removed. Nonetheless, cognitive processes remained the most problematic section even after refinement. This result evidences the difficulty in capturing cognitive processes via narrative analysis because they refer to complex constructs and are less familiar to external raters than other language domains. Future efforts should aim to provide more operational definitions for the different cognitive processes, even though this proposal is a good starting point for further developing of improved narrative measures.

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

Notwithstanding its contributions, the present study has some limitations. First, despite the above-mentioned advantages, the use of dichotomous variables for language categories provides restricted information. Regarding the inter-rater reliability analyses, we chose the multirater K_{free} coefficient, which does not have restrictions in the number of cases that should be distributed into each category and does not distinguish between agreements in coding "presence" and "absence". Consequently, we could assume that high levels of inter-rater agreement for positive narratives might be due to the great number of absences. In fact, exploratory analyses confirm that high levels of agreement were more probable in relation to absences. In the same way, some categories with high kappa values (e.g., anger towards oneself) were also those that were less frequently identified in the narrative analysis. It is worth noting, however, that none of the categories were absent in all the narratives, and, therefore, all of them reflect psychological processes that could be recognized. Another aspect that could have affected to the CASNOT reliability is the participation of a large number of raters. According to the recommendations (Chorney et al., 2015; Krippendorff, 2004), all the raters had similar levels of professional expertise and appropriate backgrounds, and were properly trained. However, the variety of raters with different personal features would increase the heterogeneity in coding; even though the CASNOT demonstrated good psychometric properties overall.

In sum, the present study, within its limitations, introduces an efficient and reliable method to assess language domains related to psychological processes involved in the narrative construction of meaningful autobiographical memories. In particular, the CASNOT could be a powerful tool for the analysis of traumatic memories and, therefore, for the analysis of how individuals elaborate and cope with their most significant experiences.

References

Bartholomew K., Henderson A. J. Z., & Marcia J. E. (2000). Coded semi-structured interviews in social psychological research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 286–312). Cambridge, MA: Cambridge University Press.

Buck N., Kindt M., van den Hout M., Steens L., & Linders C. (2006). Perceptual memory representations and memory fragmentation as predictors of post-trauma symptoms. *Behavioural and Cognitive Psychotherapy*, *35*, 259–272. https://doi.org/10.1017/S1352465806003468

Chorney J. M., McMurtry C. M., Chambers C. T., & Bakeman R. (2015). Developing and modifying behavioral coding schemes in pediatric psychology: A practical guide. *Journal of Pediatric Psychology*, *40*, 154–164. https://doi.org/10.1093/jpepsy/jsu099

Chung C., & Pennebaker J. W. (2011). Using computerized text analysis to assess threatening communications and behavior. In C. Chauvin (Ed.), *Threatening communications and behavior: Perspectives on the pursuit of public figures* (pp. 3–32). Washington, DC: National Academies Press.

Cicchetti D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Crespo M., & Fernández-Lansac V. (2016). Memory and narrative of traumatic events: A literature review. *Psychological Trauma: Theory, Research, Practice, and Policy*, *8*, 149–156. https://doi.org/10.1037/tra0000041

D'Andrea W., Chiu P. H., Casas B., & Deldin P. (2012). Linguistic predictors of post-traumatic stress disorder symptoms following 11 September 2001. *Applied Cognitive Psychology*, *26*, 316–323. https://doi.org/10.1002/acp.1830

Dekel S., & Bonanno G. A. (2013). Changes in trauma memory and patterns of posttraumatic stress. *Psychological Trauma: Theory, Research, Practice, and Policy*, *5*, 26–34. https://doi.org/10.1037/a0022750

Ehlers A., & Clark D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, *38*, 319–345. https://doi.org/10.1016/S0005-7967(99)00123-0

Eid J., Johnsen B. H., & Saus E. R. (2005). Trauma narratives and emotional processing. *Scandinavian Journal of Psychology*, *46*, 503–510. https://doi.org/10.1111/j.1467-9450.2005.00482.x

Fernández-Lansac V., & Crespo M. (2015). Narrative length and speech rate in battered women. *PLos ONE*, *10*, e0142651. https://doi.org/10.1371/journal.pone.0142651

Fleiss J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*, 378–382. https://doi.org/10.1037/h0031619

Foa E. B., Molnar C., & Cashman L. (1995). Change in rape narratives during exposure therapy for posttraumatic stress disorder. *Journal of Traumatic Stress*, *8*, 675–690. https://doi.org/10.1002/jts.2490080409

Foa E. B., & Riggs D. (1993). Posttraumatic stress disorder in rape victims. In J. Oldham, M. Riba, & A. Tasman (Eds.), *American psychiatric press review of psychiatry* (Vol. 12., pp. 273–303). Washington, DC: American Psychiatric Press.

Foltz P. W. (1996). Latent semantic analysis for text-based research. *Behavior Research, Methods, Instruments and Computers*, *28*, 197–202. https://doi.org/10.3758/BF03204765

Forgas J. P., Vincze O., & László J. (2014). Social cognition and communication: Background, theories, and research. In J. P. Forgas, O. Vincze, & J. László (Eds.), *Social cognition and communication. The Sydney Symposium of Social Psychology* (pp. 1–22). New York, NY: Psychology Press.

Freer B. D., Whitt-Woosley A., & Sprang G. (2010). Narrative coherence and the trauma experience: An exploratory mixed-method analysis. *Violence and Victims*, *25*, 742–754. https://doi.org/10.1891/0886-6708.25.6.742

Griffin C. (2004). The advantages and limitations of qualitative research in psychology and education. *Psychological Society of Northern Greece*, *2*, 3–15.

Hagenaars M. A., van Minnen A., & Hoogduin K. A. L. (2009). Reliving and disorganization in posttraumatic

stress disorder and panic disorder memories. *Journal of Nervous and Mental Disease*, *197*, 627–630. https://doi.org/10.1097/NMD.0b013e3181b08bdf

Halligan S. L., Michael T., Clark D. M., & Ehlers A. (2003). Posttraumatic stress disorder following assault: The role of cognitive processing, trauma memory, and appraisals. *Journal of Consulting and Clinical Psychology*, *71*, 419–431. https://doi.org/10.1037/0022-006X.71.3.419

Jelinek L., Stockbauer C., Randjbar S., Kellner M., Ehring T., & Moritz S. (2010). Characteristics and organization of the worst moment of trauma memories in posttraumatic stress disorder. *Behaviour Research and Therapy*, *48*, 680–685. https://doi.org/10.1016/j.brat.2010.03.014

Jobson L., & O'Kearney R. (2006). Cultural differences in autobiographical memory of trauma. *Clinical Psychologist*, *10*, 89–98. doi:10.1080/13284200600939892.

Joffe H., & Yardley L. (2004). Content and thematic analysis. In D. F. Marks & L. Yardley (Eds.), *Research methods for clinical and health psychology* (pp. 56–68). London, UK: SAGE Publications.

Jones C., Harvey A. G., & Brewin C. R. (2007). The organisation and content of trauma memories in survivors of road traffic accidents. *Behaviour Research and Therapy*, *45*, 151–162. https://doi.org/10.1016/j.brat.2006.02.004

Krippendorff K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Landis J. R., & Koch G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

Lisak D. (1994). The psychological impact of sexual abuse: Content analysis of interviews with male survivors. *Journal of Traumatic Stress*, *7*, 525–548. https://doi.org/10.1002/jts.2490070403

Mehl M. R. (2006). Quantitative text analysis. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 141–156). Washington, DC: American Psychological Association.

Morgan C. D., & Murray H. A. (1935). A method for investigating fantasies: The thematic apperception test. *Archives of Neurology and Psychiatry*, *34*, 289–306. https://doi.org/10.1001/archneurpsyc.1935.02250200049005

Murray E. J., Lamnin A. D., & Carver C. S. (1989). Emotional expression in written essays and psychotherapy. *Journal of Social and Clinical Psychology*, *8*, 414–429. https://doi.org/10.1521/jscp.1989.8.4.414

O'Kearney R., & Perrott K. (2006). Trauma narratives in posttraumatic stress disorder: A review. *Journal of Traumatic Stress*, *19*, 81–93. https://doi.org/10.1002/jts.20099

Papini S., Yoon P., Rubin M., López-Castro T., & Hien D. A. (2015). Linguistic characteristics in a non-trauma-related narrative task are associated with PTSD diagnosis and symptom severity. *Psychological Trauma: Theory, Research, Practice, and Policy*, *7*, 295–302. https://doi.org/10.1037/tra0000019

Pennebaker J. W., Chung C. K., Ireland M. E., Gonzales A. L., & Booth R. J. (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: LIWC.net.

Pennebaker J. W., Francis M. E., & Booth R. J. (2001). *Linguistic inquiry and word count*. Mahwah, NJ: Erlbaum.

Pennebaker J. W., Mehl M. R., & Niederhoffer K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547–577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Polkinghorne D. E. (2005). Language and meaning: Data collection in qualitative research. *Journal of Counseling Psychology*, *52*, 137–145. https://doi.org/10.1037/0022-0167.52.2.137

Popping R. (2000). *Computer-assisted text analysis*. London, UK: Sage.

Randolph J. J. (2005, October). Free-marginal multirater kappa: An alternative to Fleiss' fixed-marginal multirater kappa. *Paper presented at the Joensuu University Learning and Instruction Symposium*, 14–15th. Joensuu, Finland. (ERIC Document Reproduction Service No. ED490661).

Römisch S., Leban E., Habermas T., & Döll-Hentschker S. (2014). Evaluation, immersion, and fragmentation in emotion narratives from traumatized and nontraumatized women. *Psychological Trauma: Theory, Research, Practice, and Policy*, *6*, 465–472. https://doi.org/10.1037/a0035169

Shrout P. E., & Fleiss J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Smith C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. New York, NY: Cambridge University Press.

Smith C. P. (2000). Content analysis and narrative analysis. In H. T. Reis & C. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 313–335). New York, NY: Cambridge University Press.

Stone P. J, Dunphy D. C., Smith M. S., & Ogilvie D. M. (1966). *The general inquirer: A computer approach to content analysis*. Cambridge, MA: MIT Press.

Sutherland K., & Bryant R. A. (2005). Self-defining memories in post-traumatic stress disorder. *British Journal of Clinical Psychology*, *44*, 591–598. https://doi.org/10.1348/014466505X64081

Tausczik Y. R., & Pennebaker J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 24–54. https://doi.org/10.1177/0261927X09351676

Tuval-Mashiach R., Freedman S., Bargai N., Boker R., Hadar H., & Shalev A. Y. (2004). Coping with trauma: Narrative and cognitive perspectives. *Psychiatry*, *67*, 280–293. https://doi.org/10.1521/psyc.67.3.280.48977

Appendix

Description of the CASNOT language categories and dimensions

LANGUAGE CATEGORIES

Category	Description	Example
Section: Emotional processes		
Sadness	Sadness (melancholy, unhappiness, depression…)	*"I felt depressed and blue"*
Fear	Fear (terror, horror, panic…)	*"I was terrified"*
Guilt	Guilt or regret for an action, feelings or thoughts	*"I should have done more than I did"*
Humiliation/Shame	Humiliation or shame for an action, feelings or thoughts	*"He compelled me to humble myself"*
Anger towards oneself	Anger (irritation, fury…) in relation to own actions, feelings or thoughts	*"I am angry at myself"*
Anger outward	Anger (irritation, fury…) in relation to other persons or situations	*"I was furious with everybody"*
Other negative emotions	Negative emotions not previously included	*"I miss him"*
Positive emotions	Positive emotions primary (e.g., joy) and secondary (e.g., happiness)	*"I feel realized"*
Surprise	Surprise in relation to oneself, other persons or situations	*"I was surprised to see him like that"*
Empathy/Concern for the aggressor	Understanding or concern for the aggressor	*"I do not want to hurt him"*
Empathy/Concern for other/s	Understanding or concern for other persons or animals	*"She was really upset"*
Section: Body states and symptoms		
Anxiety	Symptoms of anxiety, panic, agitation…	*"I was nervous"*
Bodily states	Somatosensory or physical sensations not related to anxiety	*"My stomach hurt"*
Dissociative experiences	Dissociative experiences (depersonalization, derealization…) peritraumatic and posttraumatic	*"I do not know how I could getting there"*
Section: Sensory/ perceptual processes		
Visual perception	Visual sensations or attributes perceived through sight	*"I looked at him"*
Auditory perception	Auditory sensations or attributes perceived through hearing	*"There was a song in the radio"*
Other sensations	Olfactory, gustatory, or tactile sensations	*"He wore him usual perfume"*
Section: Cognitive processes		
Lack of control/Helplessness	Feelings of lack of control in relation to oneself, other persons, or situations	*"There was nothing I could do"*
Control/Planning	Feelings of control in relation to oneself, other persons, or situations, and development of action plans	*"I knew I was able to stop it"*
Insight	Processes that reflect knowledge or understanding of oneself, other persons, or situations	*"I see it clearly now"*
Rationalization/Elaboration	Processes that reflect a conscious effort for elaborating and explaining the event	*"Talking about it helping me to understand"*
Section: References to other people		
Social support	Perception of social support from other persons or animals	*"A neighbor called the police"*
Lack of social support	Perception of lack of social support from other persons or animals	*"Nobody did anything"*

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press

Appendix *(Continued)*

LANGUAGE CATEGORIES

Category	Description	Example
Section: Self-evaluation		
Negative self-evaluation	Negative assessment of oneself	*"I know I am weak"*
Positive self-evaluation	Positive assessment of oneself	*"I think I am worthwhile"*
Section: Meaning		
Spirituality	Religious, transcendental, or spiritual experiences or practices	*"God intended it this way"*
Meaning	Endowment or search of meaning or sense to the event	*"It happened because of a reason"*
Section: Integration		
Impact	Impact of the event for the life, the vision of oneself, the others or the world	*"I never thought that it could happen to me"*
Section: Overcoming		
Overcoming/Hope	Overcoming, or attempts or desire to overcome, the event and its consequences	*"I have already overcome it"*
Forgiveness	Feelings, or attempts or desire, to forgive the aggressor, oneself, or other persons	*"I try to forgive him"*
Section: Memory consistency		
Uncertainty	Insecurity or confusion in the memory of the event	*"I do not know what I did"*
Certainty	Security in the memory of the event	*"I am sure that was what happened"*
Section: Other aspects		
Threats to physical integrity or life	Perception of danger in relation to the own life or body	*"I thought he was going to kill me"*
Other references to death	References to death not previously included	*"I wanted him to die"*
Escapes/avoidance	Expressions that divert the attention of the narrative	*"I am talking a lot…"*
Allusions to the negative event	In positive narratives, references to the traumatic/negative event	*"This day was happy although he already hit me"*
Section: Specificity		
Extended event	Events that last more than one day	*General abuse experience*
Categorical event	Repeated actions or categories of similar events	*Summer vacation, repeated abuse episodes…*
Different specific events	Two or more specific events	*Birth of two children, two specific abuse episodes…*
One specific event	One single event that lasts less than one day	*The weeding day, an abuse episode…*

LANGUAGE DIMENSIONS

Dimension	Description	Score range
Emotional tone	General assessment of emotional tone of voice	0 (Completely positive) – 4 (Completely negative)
Emotional valence	Degree to which the narrative contains positive or negative emotions	0 (Completely positive) – 4 (Completely negative)
Wealth of details	Degree to which qualities and attributes are described	0 (Absence of details) – 3 (Very detailed)
Spatial orientation	Degree to which the event is placed in a specific spatial context	0 (Lack context) – 3 (Very orientated)
Temporal orientation	Degree to which the event is placed in a specific temporal context	0 (Lack context) – 3 (Very orientated)
Coherence	Degree to which the narrative is expressed in a clear and understandable manner	0 (Absence of coherence) – 3 (Very coherent)

https://doi.org/10.1017/sjp.2017.31 Published online by Cambridge University Press