Intrusions in story recall: When over-learned information interferes with episodic memory recall. Evidence from Alzheimer’s disease

Francesca De Anna, Eve Attali, Laurence Freynet, Lucie Foubert, Aurore Laurent, Bruno Dubois, and Gianfranco Dalla Barba

ARTICLE INFO

Article history:
Received 10 August 2005
Reviewed 10 November 2005
Revised 6 April 2006
Accepted 18 August 2006
Action editor Michael Kopelman
Published online 17 November 2007

Keywords:
Alzheimer’s disease
Intrusions
Confabulations

ABSTRACT

Patients with Alzheimer’s disease (AD) suffer from distortions of memory. Among such distortions, intrusions in memory tests are frequently observed.

In this study we describe the performance of a group of mild AD patients and a group of normal controls on the recall of three different types of stories: a previously unknown story, a well-known fairy-tale (Cinderella), and a modified well-known fairy-tale (Little Red Riding Hood is not eaten by the wolf).

The aim of our study was to test the hypothesis that in patients who tend to produce intrusions, over-learned information interferes with episodic recall, i.e., the retrieval of specific, unique past episodes. AD patients produced significantly more intrusions in the recall of the modified fairy-tale compared to the recall of the two other stories. Intrusions in the recall of the modified fairy-tale always consisted of elements of the original version of the story. We suggest that in AD patients intrusions may be traced back to the interference of strongly represented, over-learned information in episodic memory recall.

© 2007 Published by Elsevier Masson Srl.

1. Introduction

Patients with probable Alzheimer’s disease (AD) not only have great difficulty retrieving memories and information but also suffer from distortions of memory (Dalla Barba et al., 1995, 1999; Dalla Barba and Wong, 1995; Budson et al., 2000, 2002; Balota et al., 1999). One of such memory distortions is intrusion. At a general level, intrusion can be defined as the unintentional recall of inappropriate information in a laboratory-learning task such as word-list recall and story recall (Dalla Barba and Wong, 1995). More specifically, following Dalla Barba et al. (2002), intrusion is operationally defined in this study as the production of a word or other story component that deviates from the original, to-be-remembered story.

Intrusion has been reported in a variety of pathological conditions including dementia (Butters, 1987; Dalla Barba et al., 2002).
et al., 1995; Dalla Barba and Wong, 1995; Granholm and Butter,
1988), amnesia (Butters, 1987; Dalla Barba, 1993b; Dalla
Barba and Wong, 1995; Kopelman, 1987), depression (Dalla
Barba and Wong, 1995; Loewenstein et al., 1991), and also,
on some occasions, in normal subjects (Butters, 1987; Dalla
Barba, 1993b; Dalla Barba et al., 2002; Granholm and Butters,
Fuld et al. (1982) have shown that disruption of cholinergic
systems in normal subjects and in patients with AD plays an
important role in the production of intrusions and concluded
that intrusions were sufficiently characteristic of AD to be
helpful diagnostically. However, Dalla Barba and Wong (1995),
Dalla Barba et al. (1995) have shown that intrusions, al-
though frequently observed in AD, are far from being specific
to this disease since they can be equally observed in depres-
sion and in non-AD amnesias.
Dalla Barba and Wong (1995) found that AD patients made
a high proportion of intrusions when attempting to retrieve
a list of words not related to one another. When AD patients
were asked to retrieve a list of semantically associated words,
the production of intrusions was found to be associated with
the presence of a semantic memory deficit, as defined by ab-
normal performance on tests that utilize semantic memory
stores (e.g., the ability to generate words in a specified cate-
gory over 1 min).
These findings were replicated in another study on AD pa-
tients and patients with depression (Dalla Barba et al., 1995).
In a more recent study, Dalla Barba et al. (2002) found that
normal subjects produced significantly more intrusions in
story recall when they were asked to perform a secondary
task during the encoding of the to-be-remembered story as
compared to when they were asked to perform a secondary
task during the retrieval of the story. The authors concluded
that encoding processes have a greater role in eliciting intru-
sions than retrieval processes. This idea has been formalised
in Dalla Barba’s and co-workers interpretation of confabula-
tion (Dalla Barba et al., 1997, 1999, 2002; Dalla Barba, 2000).
They argued that confabulation in episodic memory is the re-
sult of a condition in which conscious retrieval processes are
no longer able to operate a “fine grain” search in long-term
memory. In this condition, events that do not have a strong
representation in long-term memory cannot be retrieved.
Only more stable representations in long-term memory are
used, with the result that habits or semantic information are
considered in a personal framework. In fact, there is clinical
and experimental evidence that confabulating patients often
mistake their habits or personal semantic information for spe-
cific unique episodes. For example, when asked what they did
the previous day, confabulating patients often report as mem-
ories what they usually do in their daily life (Burgess and
McNeil, 1999; Dalla Barba, 1993a, 2000; Dalla Barba et al.,
1997, 1999). Although admitted to the hospital, they may say
that on the previous day they went to work or they went out
shopping: acts that were part of their routine life.
In keeping with Dalla Barba and co-workers’ previous stud-
ies, we suggest the hypothesis that in patients who produce
intrusions, over-learned information interferes with episodic
recall, i.e., the retrieval of specific, unique past episodes. We
tested this hypothesis in AD patients using an experimental
paradigm aimed to elicit intrusions. We asked AD patients
and normal controls to recall different types of short stories:
one unknown story, (similar to the Logical Memory test in
the Wechsler Memory Scale-Revised (Wechsler, 1987), one
well-known fairy-tale, (Cinderella), and one “modified” well-
known fairy-tale (Little Red Riding Hood is not eaten by the
wolf). According to the above hypothesis, patients should pro-
duce more intrusions for the modified well-known fairy-tale
than for the other types of stories, because original, firmly
established information concerning the fairy-tale interferes
with the recall of its modified version.

2. Methods

2.1. Subjects

A total of 32 subjects participated in the study, 16 AD patients
and 16 normal controls. All subjects gave informed consent
before being tested. Each AD patient was examined according
to the same procedure: physical and neurological examine-
ration, psychiatric status, blood tests, serum tests (including
thyroid hormones, vitamin B12, folate acids), urinalysis, rou-
tine electroencephalogram (ECG) and computed tomography
(CT) brain scan. All AD subjects included in the study met
DSM-IV (American Psychiatric Association, 1994) and NINDS-
ADRDA (McKhann et al., 1984) criteria for probable AD and
they scored four or less on the Hachinski Ischemia Score
(Hachinski et al., 1975). Patients with potentially confounding
neurological and psychiatric disorders, clinically known hear-
ing or vision impairment, a past history of alcohol abuse, psy-
chosis or major depression were excluded. The use of
medication that could interfere with test performance or
diagnosis was considered as further exclusion criteria.
Normal controls were either spouses of patients or other indi-
viduals who volunteered to participate in the research pro-
jects of our laboratory. Table 1 shows the demographic data
of the two groups, and the performance on the Mini Mental
State Examination (MMSE) (Folstein et al., 1975). Analysis of
Variance (ANOVA) revealed that the groups differed signifi-
cantly in terms of MMSE scores (p < .0001) but not in terms
of age and years of education.

2.2. Neuropsychological evaluation

Subjects were tested on the memory impairment screen
(Buschke et al., 1999), the double memory test (Buschke
et al., 1997), digit span, the Modified Card Sorting Test
(MCST), (Nelson, 1976), a verbal fluency test that involves
generating as many words as possible starting with the letters F,

<table>
<thead>
<tr>
<th>Table 1 – Demographic and clinical data</th>
<th>AD patients (n = 16)</th>
<th>Normal controls (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>79.4 (5.2)</td>
<td>78.4 (6.3)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>10.1 (1.2)</td>
<td>11.2 (2.2)</td>
</tr>
<tr>
<td>MMSE</td>
<td>23.9 (2.1)</td>
<td>28.5 (1.0)</td>
</tr>
<tr>
<td>Values are means (SD).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
P, and L in 1 min, a verbal fluency task that involves generating as many words as possible from the categories of animals, fruits and vegetables in 1 min, and the Stroop test (Stroop, 1935). Subjects’ performance on these tests is presented in Table 2. Multiple t-tests revealed that AD patients were significantly impaired on all tasks relative to normal controls (all \( p \leq .001 \)) with the exception of the digit span.

### 2.3. Experiment

Three different types of short stories were used in this experiment: one unknown story, i.e., similar to the first Logical Memory sub-test in the Wechsler Memory Scale-Revised (Wechsler, 1987), one well-known fairy-tale, (Cinderella) and one “modified” well-known fairy-tale (Little Red Riding Hood). Each story was composed of 35 to-be-remembered elements. In the unknown story, 10 to-be-remembered elements were added to the first Logical Memory sub-test in the Wechsler Memory Scale-Revised. For example, some added elements were: /Anne Boiron/could pay/a meal/to her children. In the two other stories the to-be-remembered elements were, as far as possible, matched for length, emotional content and amount of perceptual details to the elements of the unknown story. The three types of stories were read aloud to the subjects in a counterbalanced order. Immediately after the presentation of each story, which lasted approximately 60 sec, the subjects were required to free recall as many elements as they could of the story they had just heard. They were allowed 120 sec for the recall. The subjects’ recall was tape recorded and subsequently scored by three different raters according to the number of elements correctly recalled and the number of intrusions. Inter-raters’ reliability was 100%. Intrusions were defined as the production of elements that deviate in any way from the original to-be-remembered material. For example, in the report “Anne Boiron... working as a teacher...” the word “teacher” was considered as an intrusion since it deviates clearly from the to-be-remembered element “Anne Boiron... working as a cook...”

Statistical analyses were conducted on the number of elements recalled, on the number of correct responses and the number of intrusions, and on the number of correct responses and of intrusions over the total number of elements recalled. The rationale for this latter analysis (the comparison of the number of correct responses and of intrusions over the total number of elements recalled) was conducted in order to know whether, regardless the total number of intrusions, patients produced more intrusions than correct responses, in particular in the “modified well known fairy-tale” condition. In other words, a subject may recall only a few elements of the to-be-remembered story, but most, or all of these elements are intrusions. This result is much more informative about the tendency to make intrusions than the total number of intrusions alone.

### 3. Results

#### 3.1. Total recall

The total number of elements, i.e., the sum of correct responses and intrusions, recalled in the three experimental conditions were entered in a 2 \( \times \) 3 repeated measure ANOVA with Group (AD patients vs normal controls) as between-subjects factor and Type of Story (unknown vs well-known vs modified well-known) as within-group factor. Results indicated a main effect for Group \( [F(1,30) = 40.2, \, p < .0001] \), and for Type of Story \( [F(1,2) = 12.6, \, p < .0001] \). There was also a Group by Type of Story interaction \( [F(1,2) = 4.7, \, p < .05] \).

Post hoc analysis (Scheffe’s F) corrected for multiple non-independent comparisons) revealed that AD patients recalled significantly fewer elements than normal controls in any experimental condition (all \( p < .0001 \)).

Normal controls recalled significantly fewer elements in the first experimental condition (unknown story) compared to the second experimental condition, i.e., well-known fairy-tale \( (t = 3.8, \, p < .05) \). They also produced significantly fewer elements in the third experimental condition, i.e., modified well-known fairy-tale, compared to the second experimental condition \( (t = 3.0, \, p < .05) \). No significant difference emerged between Condition 1 and Condition 3.

AD patients recalled significantly fewer elements in the first experimental condition (unknown story) compared to the second experimental condition, i.e., well-known fairy-tale \( (t = 2.3, \, p < .05) \). They also produced significantly fewer elements in the third condition (modified well-known fairy-tale) compared to the second experimental condition \( (t; \, p < .05) \). No significant difference emerged between Condition 1 and Condition 3.

#### 3.2. Number of correct responses and intrusions

The number of correct responses (Fig. 1) and intrusions (Fig. 2) produced in the three experimental conditions were entered in a 2 \( \times \) 3 \( \times \) 2 repeated measure ANOVA with Group (AD patients vs normal controls) as between-subjects factor,
and Type of Story (unknown vs well-known vs modified well-known) and Type of Response (correct responses vs intrusions) as within-group factor. Results indicated a main effect for Group \(F(1,30) = 40.2, p < .0001\), for Type of Story \(F(1,2) = 12.6, p < .0001\), and for Type of Response \(F(1,1) = 104.4, p < .0001\). There was also a Group by Type of Story interaction \(F(1,2) = 4.7, p < .05\), a Group by Type of Response interaction \(F(1,1) = 75.0, p < .0001\), and a Type of Story by Type of Response interaction \(F(2,1) = 15.6, p < .0001\).

Post hoc analysis (Scheffe’s F corrected for multiple nonindependent comparisons) revealed that AD patients produced significantly fewer correct responses and more intrusions than normal controls in any experimental condition (all \(p < .0001\)).

The number of correct responses did not significantly change across experimental conditions in normal controls. In contrast, AD patients produced significantly fewer correct responses in the modified fairy-tale condition compared to the first (\(t = 2.1; p < .05\)) and to the second (\(t = 3.7; p < .01\)) condition.

The number of intrusions did not significantly change across experimental conditions in normal controls. In contrast, AD patients produced significantly more intrusions in the modified fairy-tale condition compared to the first (\(t = 4.9; p < .0001\)) and to the second (\(t = 2.2; p < .05\)) condition.

### 3.3. Percentage of correct responses and intrusions

As stated in Section 2, a further analysis was conducted on the percentage of correct responses and intrusions (see Section 2 for the rationale of this analysis).

The percentage of correct responses and of intrusions (Fig. 3) produced in the three experimental conditions were entered in a \(2 \times 3 \times 2\) repeated measure ANOVA with Group (AD patients vs normal controls) as between-subjects factor, and Type of Story (unknown vs well-known vs modified well-known) and Type of Response as within-group factor. Results indicated a main effect for Group \(F(1,30) = 37.3, p < .0001\), for Type of Story \(F(1,2) = 8.5, p < .0001\), and for Type of Response \(F(1,1) = 40.6, p < .0001\). There was also a Type of Story by Type of Response interaction \(F(2,1) = 8.5, p < .0005\).

Post hoc analysis (Scheffe’s F corrected for multiple nonindependent comparisons) revealed that AD patients produced significantly fewer correct responses and more intrusions than normal controls in any experimental condition (all \(p < .0001\)).

The percentage of correct responses did not significantly change across experimental conditions in normal controls. In contrast, AD patients produced a significantly lower percentage of correct responses in the modified fairy-tale condition compared to the two other conditions (\(t\) test; \(p < .05\)).

The percentage of intrusions did not significantly change across experimental conditions in normal controls. In contrast, AD patients produced a significantly higher percentage of intrusions in the modified fairy-tale condition compared to the first (\(t = 4.8; p < .0001\)) and to the second (\(t = 4.5; p < .005\)) condition.

On both normal controls and patients, and within each group of subjects correlation coefficients were calculated...
between intrusions and correct responses in the three experimental conditions. Significant, negative correlations were observed in the three experimental conditions (all $p = -1$, $p < .0001$). On both normal controls and patients, and within each group of subjects, correlation coefficients were also calculated between intrusions in the three experimental conditions and the performance on verbal fluency, on the Stroop test and on the MCST. The only significant, negative correlation was observed in both normal controls and patients between intrusions in the first experimental condition and the number of errors in the MCST ($p = -5$, $p < .05$).

Intrusions in the first and second experimental conditions always consisted of elements not belonging to the to-be-remembered material, but coherent with the semantic structure of the story. An example of intrusion in the first experimental condition is reported in Section 2. As far as the second experimental condition is concerned, several patients reported, for example, that “Cinderella was beautiful”, an element which didn’t belong to the to-be-remembered material, but was coherent with the semantic gist of the story.

In the modified fairy-tale condition, AD patients’ intrusions always consisted of elements of the original version of the story. The following is an example of a modified version of the Little Red Riding Hood fairy-tale read by the examiner to the patient: “Once upon a time there was a little girl called Little Red Riding Hood because of her little red hood. On her way to visit her grandmother, Little Red Riding Hood found a nice wolf cub lost in the forest. Little Red Riding Hood decided to take the wolf cub to her grandmother’s to feed it. Very scared, the grandmother implored Little Red Riding Hood to put the animal back in its den. Little Red Riding Hood was very sad when she left her grandmother’s house with the wolf cub, but then, fortunately, in a clearing of the forest she found the mother wolf.” This narration was recounted in the following way by one AD patient: “Once upon a time there was a little girl called Little Red Riding Hood because she was wearing a red hood. She goes in to the forest and there she encounters the wolf. But the wolf doesn’t eat Little Red Riding Hood... not immediately. Little Red Riding Hood takes the wolf to her grandmother’s and there the wolf eats Little Red Riding Hood. It ends like that. But there is something in between... I can’t remember...” As it emerges from this example, the original version of the story clearly interferes with the recall of its modified version. In other words, strongly represented, over-learned information interferes with episodic recall, even though some episodic elements are present in the patient’s report (e.g., Little Red Riding Hood takes the wolf to her grandmother’s house.)

Not surprisingly and consistent with previous reports, the results have shown that AD patients recalled a lower number of elements, i.e., the sum of correct responses and intrusions, than normal controls. This result supports the idea that, even in the early stage of the disease, AD patients show a deficit of retrieval that may be traced back to an impairment of executive functions (Perry and Hodges, 2000). Actually our patients performed poorly on some tests considered sensitive to executive dysfunction (see Table 2).

Both groups of subjects recalled more elements in the second experimental condition than in the two other conditions. This result may be interpreted as reflecting the fact that in the second condition the to-be-remembered material matched the original version of the fairy-tale. Therefore, both AD patients and normal controls could rely on a firmly established memory of the original version of the fairy-tale in order to retrieve both to-be-remembered elements and intrusions coherent with the original version of the fairy-tale.

AD patients produced a lower number of correct responses and a greater number of intrusions than normal controls in the three experimental conditions, but, whereas in normal controls the number of correct responses and of intrusions did not change across experimental conditions, AD patients produced a lower number of correct responses and a greater number of intrusions in the third experimental condition (modified well-known fairy-tale). The performance of normal controls is difficult to interpret because they showed ceiling effects for correct responses in the three conditions. This is the result of a methodological limitation of the study. In fact, normal controls were given the same task as patients and the level of difficulty was not controlled. However, the results are consistent with the hypothesis that in patients who show a tendency to produce intrusions, firmly established, over-learned information interferes with episodic recall, i.e., the retrieval of specific, unique past episodes.

As far as the percentage of correct responses and of intrusions over the total number of elements recalled (both correct and intrusions), normal controls performed at ceiling for correct responses in the three experimental conditions and produced very few intrusions in all conditions (see Fig. 3). In contrast, AD patients produced more correct responses in Conditions 1 and 2 than in Condition 3 and produced fewer intrusions in Conditions 1 and 2, than in Condition 3.

It has been proposed that false remembering is sensitive to emotional information and to the amount of perceptual details story contain (Johnson, 1997). In addition, at least some confabulating patients have been shown to have positive emotional bias in their memory (Fotopoulou et al., 2004). However, as far as possible, in this study the three stories were matched for emotional information and amount of perceptual details. In addition, there was no evidence that the memories produced by our subjects were the result of a positive emotional bias.

Consistent with our hypothesis, in the present study patients produced significantly more intrusions when asked to recall the modified well-known fairy-tale than when they recalled the two other stories. The modified well-known fairy-tale is composed of elements in common with the original, over-learned version of the fairy-tale, (Little Red Riding Hood went to visit her grandmother), and of elements which
differ considerably from the original version of the fairy-tale, (Little Red Riding Hood is not eaten by the wolf). Therefore it is likely that the representation of the original version of the fairy-tale is activated by elements in common with the modified version. The activation of the representation of the original, over-learned version of the fairy-tale interferes with the recall of the episodic representation of the elements of the modified version, which differ from the original version. Interference of over-learned information with episodic recall may reflect intensive effort to recall poorly encoded material. It is possible that, while trying to retrieve the fairy-tales, AD patients are at higher risk of producing intrusions from the original version of the fairy-tale while recalling its modified version than when recalling an unknown story. In that sense, intrusions would represent a normal reaction to defective learning (Kopelman, 1987). This explanation is partially supported by the significant, negative correlation between the number of intrusions and the number of correct responses. This correlation may in fact reflect the tendency to “replace” poorly encoded material with over-learned information.

Our study didn’t provide evidence for a correlation between intrusions and subjects’ performance on the Stoop test. A paradoxical negative correlation was observed between intrusions in the first experimental condition and the number of errors in the MCST. Accordingly, although defective inhibition of over-learned material during episodic recall remains a potential explanation of intrusions, it is not directly supported by the results of this study. It must be noted, however, that the Stoop test tap on the inhibition of visually presented automatic responses, which is different from the inhibition of verbal over-learned memory material involved in our experimental task. In addition, in the Stoop test subjects are instructed to inhibit the reading of the word, whereas in our story recall experiment they are instructed to remember a story. In other words, the Stoop and other standard neuropsychological tests cannot tell us exactly which part of the higher order functions that regulate memory is impaired in intrusions. Our results do not support directly any specific interpretation of intrusions and other types of false memories, but are compatible with a number of interpretations. For example, the pattern of intrusions observed in this study may be the result of a ‘source monitoring’ deficit (Johnson, 1997) or of a deficit of ‘filtering’ (Schnider and Ptak, 1999). Intrusions may also occur because ‘strategic retrieval’ is impaired and information can only be retrieved associatively, as suggested by Moscovitch and Melo (1997). It is also possible that patients in this study produced intrusions because they could not make the ‘fine grain’ distinction between more recent (and thus less well-learned, harder to retrieve) and once learned and now well-known information, and thus they went for the later, consciously or not, as suggested by Dalla Barba et al. (1997) and Dalla Barba (2000).

Our study did not address the potential role of an encoding and consolidation deficit in the production of intrusions. It is conceivable, however, that the recall of poorly encoded and consolidated episodic material is more sensitive to the interference of strong representations. Accordingly, both a retrieval as well as an encoding and consolidation deficit are likely to be involved in the origin of intrusions and other confabulatory-like behaviours. This pattern of deficits may explain why remote episodic autobiographical memories, which are better encoded and consolidated, are less likely to be contaminated by confabulation (Dalla Barba et al., 1998). The contribution of an encoding deficit in the origin of intrusions is also consistent with data showing that poor encoding yields an increase of intrusions in story recall tasks in normal subjects (Dalla Barba et al., 2002).

In conclusion, our findings indicate that the interference of strongly represented, over-learned information in episodic memory recall is implicated in the production of intrusions. This phenomenon has been known, but this study is the first to demonstrate it experimentally. In addition, our work suggests the need to address the question as to whether firmly established episodic representation interferes with newly-encoded episodic information. Future research may answer this question.

REFERENCES

Dalla Barba G, Mantovan MC, Traykov L, Rieu D, Laurent A, and Devouche E. The functional locus of intrusions: encoding or