ELSEVIER

International Journal of Psychophysiology 38 (2000) 21–41

INTERNATIONAL JOURNAL OF PSYCHOPHYSIOLOGY

www.elsevier.nl/locate/ijpsycho

The objective assessment of amnesia in dissociative identity disorder using event-related potentials

John J.B. Allen*, Hallam L. Movius, II

Department of Psychology, University of Arizona, P.O.B. 210068, Tucson, AZ 85721-0068, USA

Received 1 September 1999; received in revised form 2 April 2000; accepted 2 May 2000

Abstract

Assessment of amnesia in Dissociative Identity Disorder (DID) typically relies on self-report, the veracity of which cannot often be independently verified. Memory in DID was therefore assessed using an objective method that involved event-related potentials (ERPs) as well as indirect behavioral measures of memory, and that provided statistically supported assessments for each participant. Four participants who met DSM-IV criteria for DID participated in an ERP memory assessment task, in which words learned by one identity (identity A) were then presented to a second identity (identity B). All four participants — tested as identity B — produced ERP and behavioral evidence consistent with recognition of the material learned by identity A. While it would be premature to generalize all cases of DID, the results suggest that there may be reasons to question the veracity of reports by individuals who meet diagnostic criteria for DID on the basis of a structured clinical interview. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Amnesia; Dissociative identity disorder (DID); Memory

1. Introduction

Dissociative Identity Disorder (DID) — formerly termed Multiple Personality Disorder (MPD) — is characterized by the presence of two or more distinct identities or personality states that recurrently take control of the individual's behavior, accompanied by an inability to recall important personal information that is too extensive to be explained by ordinary forgetfulness (American Psychiatric Association, 1994).[1] The prevalence of DID has fluctuated widely over the last century, which some researchers have argued

* Corresponding author.

[1]The disturbance cannot be due to the direct physiological effects of a substance (e.g. alcohol) or medical condition (e.g. complex partial seizures).

PII: S0167-8760(00)00128-8

reflects the influence of media and culture on suggestible individuals who are already distressed (Spanos, 1996). The nature of DID has also changed over the last century, with early cases usually involving no more than three personalities and transitions between personalities occurring only after a period of sleep; modern patients may display an average of 15 personalities, and may have up to 100 (Ross et al., 1989b).

Perhaps the most controversial part of this diagnosis has been the debate over whether different identities or 'alters' might truly be unaware of each other. Prior to 1994 intrapsychic amnesia was not *required* as a criterion for MPD. In 1994 — with the publication of the DSM-IV — the name of the disorder was changed from MPD to DID, and amnesia 'too extensive to be explained by ordinary forgetfulness' became necessary for diagnosis. While some therapists have claimed that a DID case could have several personalities who share consciousness, the ability to determine whether personalities are truly amnestic for each other has become a key theoretical and diagnostic issue. The present study therefore used an objective and reliable method of assessing memory that is able to identify recognized material, even in cases where individuals deny knowledge of such material (Allen et al., 1992).

1.1. Previous studies of MPD[2] using 'objective' measures

Anecdotal and case studies comprise the majority of publications to date on MPD. Behavioral studies of multiple identities date back to 1817, and psychophysiological studies date back nearly a century, to an investigation of galvanic skin responses in a patient who displayed multiple identities (Greaves, 1980). There have been, however, surprisingly few controlled studies using putative MPD cases (Miller and Triggiano, 1992; Kihlstrom et al., 1993). Reported behavioral correlates of MPD include changes in handwriting across personalities (Thigpen and Cleckley, 1957; Coons, 1984), changes in vocal expression (Putnam, 1984), and changes in handedness (Putnam, 1984). Most, if not all, of these behaviors fall under voluntary control, however, and are therefore possible to simulate (Spanos, 1996). In addition to doubts about the reliability and validity of the reported observations, the lack of adequate controls or comparison subjects in almost all cases makes firm conclusions difficult, if not impossible (Miller and Triggiano, 1992; Brown, 1994).

1.2. Psychophysiological correlates of DID

Previous studies have reported a number of changes in physiology that parallel reported changes in personalities. Such changes include a change in respiration rate (Bahnson and Smith, 1975; Larmore et al., 1977), differences in patterns of galvanic skin responses (Ludwig et al., 1972; Brende, 1984), differences in visual acuity (Miller, 1989), and differences in patterns of EEG across alleged personalities (Ludwig et al., 1972; Larmore et al., 1977; Braun, 1983). While it might initially seem that such responses are outside the realm of voluntary control, it is critical to keep in mind that substantial changes in physiological measures can covary with changes in behavior. If one is to appropriately infer that some physiological change is a marker of an identity change in DID, it would be necessary to show not only that this change necessarily occurs with putative identity change in DID, but also that there is *not* a similar change across 'normal' states among non-DID individuals (cf. Cacioppo and Tassinary, 1990). Moreover, many such changes could in theory be simulated; EEG can vary markedly as a function of eye movement, for example (Plotkin, 1979; Perlini and Spanos, 1991), and merely imagining different moods can induce a number of autonomic shifts (Sheikh and Kunzendorf, 1984). In general, the strategy of looking for psychophysiological differences across alleged personalities as markers of distinct identities is a poor one, as

[2] Because the amnesia criterion was not present among the DSM-III and DSM-III-R criteria for MPD, we use the acronym MPD when reviewing studies that relied on DSM-III-R or earlier criterial for diagnosis. We reserve the acronym DID to refer to cases that meet DSM-IV criteria, and therefore necessarlily involve reported amnesia.

such differences: (1) could represent nothing more than normal change associated with changes in emotional or cognitive state; (2) could be due to a number of factors that fall under voluntary control such as muscle tension (Coons, 1988); or (3) could be due simply to capitalization on chance. In order to draw more convincing conclusions about behavioral correlates, it is necessary at a minimum to show that such changes are significantly different from those displayed by reasonable control groups or during control conditions. Unfortunately, studies that have employed either control subjects or control conditions are few in number.[3]

1.3. Studies of psychophysiological correlates with control groups

Putnam et al. (1990) examined electrodermal and cardiovascular responses to tones in nine MPD patients and five control subjects instructed to fake changes in personality, finding that there was a non-significant *trend* for MPD patients to show less habituation. Coons et al. (1982) found that a simulator/control case exhibited larger changes in various EEG frequency bands when simulating a personality change than did two putative MPD cases. The researchers concluded that the pattern of changes in EEG were more consistent with changes in anxiety, muscle tension, and concentration — perhaps reflecting changes in emotional states, expectancies, muscle tension, or task demand — than with complete transformation in personalities. As Brown (1994 p. 107) concluded, to date "there is neither consistent evidence for a distinctive EEG abnormality in MPD patients (trait) nor a shift in EEG activity (state) when a patient experiences a 'shift' between personalities".

In two studies, Miller and colleagues (Miller, 1989; Miller et al., 1991) found that MPD patients exhibited more variability in opthalmalogical responses, including visual acuity, manifest refraction, eye muscle balance, and visual fields. As Miller and Triggiano (1992, p. 54) point out, however, the obtained between-group differences may have been due to greater overall response variability within the MPD group, independent from variability associated with specific changes in personalities, and that these measures fail to 'provide evidence for the distinctiveness or uniqueness of individual personality states'.

In a study of cerebral blood flow patterns, Mathew et al. (1985) found increased blood flow to the right temporal lobe during changes in personality in one MPD patient when compared with three control subjects. No statistical tests were performed that would allow inferences to be made about the significance of reported changes in blood flow.

1.4. Experimental tests of memory in MPD cases

Nissen et al. (1988) investigated transfer of memory across identities in an individual who reported intrapsychic amnesia. On several tests there was little evidence of memory transfer, while on others there was evidence of such transfer. Nissen et al. (1988) concluded that assessing memory indirectly was a necessary, but not a sufficient, condition for demonstrating transfer of memory across personalities. Moreover, they suggested that implicit memory was seen in tasks where the interpretation of stimuli was not likely to differ between personalities, but no implicit memory was seen where individual interpretation of the stimuli was involved. Eich et al. (1997) replicated some of these findings, again finding that direct assessments of memory (e.g. recall) reveal inter-identity amnesia, and indirect methods (picture-fragment completion) show evidence of inter-identity transfer of knowledge.

Silberman et al. (1985) adopted an interesting approach to assessing memory, assuming that inter-personality amnesia should facilitate 'compartmentalizing' memories, and thereby facilitate correctly discriminating lists of words learned as different personalities. Two identities first learned

[3]Even studies of DID symptoms have suffered from the absence of comparison groups. For example, Ross et al. (1989a) found that loss of hearing, reported by 18% of DID cases in one sample (Bliss, 1980), turned out to be equally prevalent among control groups with eating disorders, panic disorder, and schizophrenia.

a list of words, and then both identities were subsequently tested after a 2-h interval. Although their overall recall and recognition performances were similar to controls, MPD subjects did no better than controls in correctly assigning learned words to the personality that had learned them, suggesting that despite MPD subjects' subjective reports of compartmentalized memories, such memories had substantially 'leaked' between personalities.

1.5. Event related potentials as a method for assessing memory

The use of event-related potentials (ERPs) with respect to MPD has been very limited. Moreover, to the extent that ERPs have been utilized in MPD (e.g. Braun, 1983), their use has not been motivated by a theoretical rationale to assess memory or other symptomatology of MPD; rather, such investigations seem motivated by the erroneous assumption that psychophysiological differences between identities can provide evidence in support of their veracity. ERPs, however, represent a promising method by which to assess memory and amnesia, a central symptom of DID.

1.5.1. ERPs and the assessment of memory

While several components of the ERP have been shown to be sensitive to memory (e.g. Sanquist et al., 1980; Neville et al., 1986; Rugg and Nagy, 1989; Smith and Halgren, 1989; Bentin and Moscovitch, 1990; Bentin et al., 1992; Paller and Kutas, 1992; Rugg and Doyle, 1992, Rugg and Doyle, 1994; Smith, 1993; for a recent review see Rugg, 1995), most research examining the relationship of these components to memory have relied on the assessment of groups of subjects, comparing components of ERPs for previously-seen and new items using conventional parametric statistical procedures. While this approach is essential for understanding the theoretical significance of these components with respect to memory, such procedures do not directly inform clinical assessment. For clinical assessment, reliable and valid decisions need to be rendered for individual cases. For this purpose, a variety of statistical methods have been used with event-related potentials.

In using ERPs to assess memory in individual cases, investigators are often interested in using ERPs to assess memory in situations where subjects are unable, or unwilling, to report their recollection for learned material. These assessment procedures are essentially variants of the Guilty Knowledge Technique (Lykken, 1959) and have most often, but not exclusively, been applied in the detection of deception (Rosenfeld et al., 1988, 1991, 1998; Boaz et al., 1991; Farwell and Donchin, 1991; Allen et al., 1992). In most of these paradigms, two classes of previously seen items are presented, only one of which a subject acknowledges recognizing. These previously seen items appear infrequently among frequent items that have not been previously seen. The previously seen items will therefore stand out as rare and significant by virtue of having been learned previously. This experimental design makes it likely that the previously seen items, regardless of whether they are acknowledged as familiar, will elicit a larger P3 amplitude in the ERP than will the frequent unfamiliar items.

Our previous work using this experimental design (Allen, 1992; Allen et al., 1992; Allen and Iacono, 1997) has shown that statistically supported decisions can be rendered for individual subjects with acceptably high rates of accuracy. Using a Bayesian classification procedure (see Appendix A) in three groups of 20 subjects each (Allen et al., 1992), the ERP procedure was successful in correctly identifying previously-learned material — *regardless of whether a person explicitly acknowledged having learned it* — 94% of the time. It was similarly successful in classifying unfamiliar material 96% of the time.

This experimental procedure, and associated analytical strategy, lends itself readily to the study of memory in DID. If one were to identify two distinct identities (identity A and identity B), each of whom reported mutual amnesia for recollections of the other identity, then each could learn a set of material that should be familiar to only that identity and not the other. Because this ERP memory assessment procedure identifies familiar material regardless of whether a subject explicitly acknowledges having learned it, the ERP could serve as an objective way to assess the report of amnesia in DID. If identity B were amnestic for items learned by identity A, then the ERP procedure should only classify ERPs to identity B's items as familiar, while classifying ERPs to identity A's material as unfamiliar. If, by contrast, identity B in fact denied recognizing identity A's items despite being aware of these items, this ERP procedure should be sensitive to that as well. In other words, as a clinical assessment tool, either outcome would prove informative. The former outcome would provide objective evidence of amnesia, while the latter outcome would suggest that the reported amnesia was, in fact, fabricated.

2. Method

2.1. Participants

Five potential participants volunteered to participate. Participants were referred from community practitioners, or responded to an advertisement — in a monthly patient newsletter — describing a study of memory and Dissociative Identity Disorder being conducted at the University of Arizona.

Participants agreed to come to two separate appointments, one for diagnostic interviewing, and a second a few weeks later involving 'brainwave' recording. During the first visit, all participants completed the Structured Clinical Interview for DSM-IV Dissociative Disorders (SCID-D; Steinberg, 1993), the Dissociative Experiences Scale (Bernstein and Putnam, 1986), the Folstein Mini Mental State Examination (MMSE; Folstein et al., 1975), and several other measures (earliest memories, digit span, various personality measures) not reported here.

From among these five potential participants, four were able to complete testing and are described here. All four met DSM-IV diagnostic criteria for DID according to their responses to questions from the SCID-D. The fifth, who had a rather remarkable claim of over 600 personalities and was judged not to provide reliable information during the interview, did not complete the testing during session 2 (see footnote 4) and did not keep several rescheduled appointments. Each of the remaining four cases is described in summary format in Table 1. As can be seen in the table, each of the four participants is rather typical of cases reported in the literature (Spanos, 1994) in that they are female, they have a history of other psychiatric problems that precede their diagnosis of DID, and that they report on average about eight identities.

2.1.1. College student comparison sample

Sixty college student controls previously reported in Allen et al. (1992) were used as a comparison sample for the present study. Although these college student controls were not specifically asked to simulate DID, they were given instructions to deny knowledge of a recently learned list in the context of an ERP assessment paradigm that was virtually identical to that used in the present study. The college students were asked to deny and conceal knowledge of a recently learned list in order to determine whether ERPs could detect recognized items even when participants denied recognizing them as previously learned.

2.2. Procedure

DID participants made two visits to the University. The first visit was to a clinical interviewing room within the Psychology Clinic, and entailed the diagnostic interviewing and the administration of the instruments listed above. The second visit, scheduled 2–4 weeks later, was to the Psychophysiology Laboratory. At the end of the first visit, the procedure for the second visit was described in order to help participants prepare for the procedure, and to determine whether they wished to bring an assistant of their choosing. (Only participant M03 did so.) Participants were told that one of the difficulties in Dissociative Identity Disorder, as they themselves had already described in the diagnostic interview, was that some identities had no awareness of the activities and experiences of other personalities. They were further told that we were specifically interested in

Table 1
Demographic, diagnostic, and clinical information for DID Participants[a]

M02	M03	M04	M05
Demographics			
□ P.K. 39-year-old female □ 2 years college □ Divorced × 4 □ Unemployed × 2 years □ DES score = 70 □ MMSE = 29	□ M.C. 36-year-old female □ 2 years college □ Separated from only marriage □ Unemployed × 2 years; now student □ DES score = 47 □ MMSE = 29	□ M.D. 40 year-old female □ 2 years college □ Divorced □ Employed 5 years □ DES score = 91 □ MMSE = 28	□ A.J. 51-year-old female □ 2 years of college □ Divorced, widowed □ Unemployed × 3 years □ DES score = 65 □ MMSE = 30
Diagnostic history			
□ History of substance abuse □ Current meds = prozac, melaril □ Diagnosed during course of substance-abuse treatment circa 1988, but reports symptoms have been present since at least teen years	□ History of eating disorder □ Current meds = clonapin, wellbutrin, synthroid □ Diagnosed 1991, but reports symptoms date back to at least teen years	□ No history of substance abuse □ Meds = clonapin, zoloft, promazine □ History of harm to self, hospitalizations □ In therapy on/off since 1983 □ Reports symptoms present since early childhood	□ Conversion reaction' 3 years ago after seizures, blindness, paralysis □ Subsequent diagnosis of DID □ History of hospitalizations, suicide attempts (earliest at age 6) □ Current meds = effexor, synthroid
Diagnostically relevant symptoms			
□ Reports at least 10 identities □ Daily gaps in memory/ missing time □ Has found unfamiliar clothes in closet □ Has found furniture re-arranged without knowledge □ Describes internal dialogues between identities	□ Reports at least 13 identities □ Daily gaps in memory/missing time □ Describes dramatic changes in abilities (e.g. artistic and computer work) □ Reports age-regression experiences □ Describes internal dialogues between identities	□ Reports at least 4 identities □ Daily gaps in memory/ missing time □ Distant travel without knowing why □ Has been called other names by people who insist they know her □ Has found clothes, jewelry, furniture she did not remember buying □ Describes changes in handwriting during journal entries	□ Reports at least 6 identities, one consistent alter throughout high school, and five now □ Daily gaps in memory/missing time □ Reports internal dialogues between identities □ Reports forgetting how to dress self, brush teeth (when a child identity) □ Reports switching under duress, especially during family conflicts □ Reports memories of ritual abuse
Observations of tested identities			
□ Reports 'co-consciousness' between tested identities	□ Reports *no* 'co-consciousness' between tested identities	□ Reports *no* 'co-consciousness' tested identities □ Experimenter had clinical impression that participant was experiencing depersonalization	□ Reports that only one of the learned words was available to both tested identities (limited 'co-consciousness') □ One identity manifests German accent □ Reports that at least some of her identities are not 'co-conscious'

[a]*Note.* All symptoms reported by the participants. DES, Dissociative Experiences Scale (max = 100, score over 30 suggests possible dissociative disorder); MMSE, Mini Mental Status Examination (max = 30).

that phenomenon, and therefore wished to test two distinct identities that were likely to be unaware of each others' activities. We asked them to think through whether it would be possible to first test one identity, and then ask them to switch to another identity that would be unlikely to know anything that just happened to the first identity. We further stipulated that both identities must be able to read common English words on a computer screen.

Upon arrival to the Psychophysiology Laboratory, participants were shown the control room and equipment, and the adjacent sound-attenuated subject chamber. Participants were then prepared for electroencephalographic (EEG) recording by fitting a stretch lycra cap and seven individual electrodes. After preparation, participants were escorted to the adjacent chamber to begin the memory assessment procedure. Participants were seated in front of a computer monitor, with an experimenter seated just behind and to the left of the participant.

2.2.1. Identity A's tasks

Each participant was first asked to report her identity. Once she confirmed that this was in fact the identity she had previously indicated should be tested, identity A was shown a list of six unrelated words and asked to learn them to a criterion of two perfect serial recitations. This list was selected randomly from among a set of seven lists of six high-frequency words that had been matched for word frequency using the word frequency norms of Carroll et al. (1971). For learning, the words were presented at a rate of one word every 2 s (stimulus duration = 1530 ms) on a computer screen mounted on a movable arm and positioned approximately 85 cm in front of the participant's face. The experimenter recited each word aloud as it appeared. Words subtended approximately 0.34° of visual angle vertically and 1.01–1.48° horizontally. Upon completion of this learning task identity A performed a simple computerized recognition task in which she pressed a button with the thumb of her dominant hand to indicate recognition of words she had learned and pressed a button with her non-dominant thumb for words she did not recognize. Learned words appeared on 1/5 of the trials and unlearned words appeared on 4/5 of the trials. These unlearned words were matched for frequency of occurrence with the seven lists (mentioned above) that participants could learn, and these unlearned words were never seen subsequently in any other tasks.

Identity A was then thanked for her participation and the experimenter requested that she switch to identity B. Each participant had her own idiosyncratic method of switching identities, and in some cases, multiple attempts were required in order for the participant to switch to the intended identity. The switch from identity A to identity B sometimes also occurred via another identity (i.e. identity A to identity X to identity B).

2.2.2. Identity B's tasks

Identity B was asked to recall what had happened during the experiment, and specifically if she remembered learning any words at this point. Participant M02 recalled identity A's words, and participant M05 recalled one of identity A's words. Neither participant M03 nor M04 could recall any words, or even recall having learned any words.

Identity B then learned a list of six unrelated words other than that learned by identity A. This list was randomly selected from among the remaining six of the original seven matched lists. This list was presented in the same manner as the earlier list. Once identity B had learned the list to the criterion of two perfect serial recitations, she was presented with the critical procedure in which words learned by identity A appeared on 1/7 or the trials, words learned by identity B appeared on 1/7 of the trials, and words matched for frequency but not previously seen by either identity appeared on 5/7 of the trials.[4] This arrangement made it likely that the learned lists, appearing relatively infrequently against a background

[4] It was at this point — after instructions for this procedure were administered — that the first potential participant (the one who claimed over 600 personalities and for whom we do not report data) reported that she suddenly developed an incapacitating migraine and needed to discontinue testing.

of unlearned material, would elicit a large P3 if they were recognized. Identity B was presented with five blocks of 42 words, with each block containing a different random ordering of all six words from all seven lists (two learned, and five unlearned). At the beginning of each block (in addition to the 42 words), three additional words appeared that were not included for purposes of signal averaging. Stimulus duration was 306 ms, and onset-to-onset interval was 1999 ms. Monitor distance and visual angles subtended by stimuli were identical to those of the learning phase. The experimenter instructed identity B that she would see a long list of words consisting of words she had learned as well as other words. She was instructed to press the 'YES' button with the thumb of her dominant hand if the word on the screen was one of the words she remembered learning today, and to press the 'NO' button with the non-dominant thumb for all the other words. She was told to respond as quickly and accurately as possible.

2.2.3. Forced-choice recognition tasks

Following this oddball recognition task, identity B was given a forced-choice recognition task. The intent of the forced-choice task was to provide an index of whether participants were systematically avoiding acknowledging words learned by identity A, and whether words seen as distractors by identity A would show evidence of transfer to identity B. Veritable amnesia would not result in systematic avoidance of learned material, but rather would show up as the tendency to respond with recognition 50% of the time, because if one were amnestic for previously learned words, one would choose them 50% of the time by chance in this forced choice task. Two words were presented side-by-side on the computer screen simultaneously, and identity B was instructed that she should indicate which of the two words she had learned. She was further instructed that if she was unsure, she should guess, as we wanted her to make a response for each and every pair of words. This task consisted of three different types of stimuli: (1) Explicit Memory Probes ($n = 6$), which contained words learned by identity A paired with previously unseen items matched for frequency; (2) Incidental Memory Probes ($n = 12$), which contained words that had been previously seen as unlearned distractors by identity A paired with previously unseen items matched for frequency; (3) Control Trials ($n = 6$), which were two previously unseen items matched in frequency to all other words. For the first two types of stimuli, words that had been previously seen appeared on the left side of the screen for half the trials and on the right side of the screen for the other half of the trials. Following this task, identity B was thanked for her time and the experimenter once again requested to meet with identity A. After participants reported a successful switch to identity A, a similar forced choice recognition task was administered to identity A to determine whether identity A was aware of what identity B had learned. In this task, the learned words (Explicit Probes) and the previously seen words (Incidental Probes) were those that were seen only by identity B.

Unfortunately, this forced choice recognition task was implemented after the testing of participant M02, and response apparatus problems rendered data from participant M03 useless. Therefore, only data for participants M04 and M05 are available.

2.3. Psychophysiological recording and analysis

2.3.1. Signal acquisition and analysis

Electroencephalographic signals were recorded from 26 scalp sites using tin electrodes mounted in an electrode cap with seven additional tin electrodes outside the cap. Three of the additional electrodes were placed to monitor electrooculographic (EOG) activity: one at the nasion and one each directly below each eye (20% of the nasion–inion distance below FP1 and FP2). Additionally, an electrode was placed on the left and the right mastoid and two of the electrodes were used with FP1 and FP2 to create two bipolar electromyographic (EMG) channels over the left and right medial frontalis muscle. Recording sites in the electrocap included FP1, FP2, F7, F8, F3, Fz, F4, FTC1, FTC2, C3, Cz, C4, T3, T4, T5, T6, TCP1, TCP2, P3, Pz, P4, PO1, PO2, O1, Oz, and O2. All impedances were below 5 kΩ and all channels were referenced online to the left mastoid and amplified using a Grass model 12 Neurodata system with a band-pass of 0.1–100 Hz and digitized continuously at a sampling rate of 256 Hz. All signals were amplified 20 000 times except for FP1, FP2, the three EOG sites, and the EMG sites, which were all amplified 5000 times. The continuous files were digitally filtered with a 12.5-Hz, 96-dB per octave low-pass filter, and then corrected for ocular artifact using the approach of Semlitsch et al. (1986) using a bipolar channel derived from the nasion and the site below the left eye. The data files were then epoched from 150 ms prior to stimulus onset to 1500 ms after stimulus onset, linear detrended (to remove any prolonged DC shift in the waveforms), baseline corrected to a prestimulus baseline of −150–0 ms, and finally re-referenced off-line to linked mastoids. Averaged ERPs were then created for each condition (two learned and five unlearned lists). For analyses involving the Bayesian prediction, the waveforms were additionally filtered with a 5.75-Hz low-pass filter just as in Allen et al. (1992).

3. Results

Although most of the analyses involve the use of the Bayesian approach detailed in Appendix A, parametric statistics are presented to round out the interpretation of the data. In cases where repeated measures analyses of variance are reported, the Greenhouse–Geisser correction was used. We report the P-values associated with the Epsilon-corrected degrees of freedom, but report the original degrees of freedom.

3.1. Behavioral data during identity B's recognition task

Table 2 shows the number of times (30 maximum) that identity B responded 'YES' to indicate recognition of different word lists. As can be seen from the table, all four DID participants denied recognition on a majority of trials in which words learned by identity A were presented, and very seldom indicated recognition for words that should have been unfamiliar (lists U1–U5). With the exception of participant M04, these response data are virtually identical to the pattern seen in college-student controls (Allen et al., 1992) who

Table 2
Affirmative responses by identity B indicating recognition of items from the various lists[a]

ID	DID Participants Responses as Identity B						
	Identity B	Identity A	U1	U2	U3	U4	U5
M02	21	4	0	1	1	1	1
M03	25	2	1	1	2	0	0
M04	13	3	0	0	1	0	1
M05	26	1	1	0	0	3	0
Mean	21.3	2.5	0.5	0.5	1.0	1.0	0.5
S.D.	5.9	1.3	0.6	0.6	0.8	1.4	0.6

ID	College student controls						
	Nconceal	Conceal	U1	U2	U3	U4	U5
Mean	24.3	1.5	0.0	0.1	0.0	0.0	0.0
S.D.	5.0	1.9	0.1	0.4	0.2	0.0	0.2

[a]*Note.* Maximum = 30. U1 through U5 are five different unfamiliar lists. Identity B refers to words learned by identity B and identity A refers to words learned by identity A. The college student controls are the 60 participants reported in Allen et al. (1992), who were instructed to deny knowledge of a previously learned list (Conceal) and to acknowledge learning a recently learned list (NConceal).

Table 3
Response latency (ms) by identity B[a]

ID	DID participants responses as identity B						
	Identity B	Identity A	U1	U2	U3	U4	U5
M02	707	689	540	578	548	535	579
M03	817[b]	775	639	559	589	545	605
M04	794	660	537	567	516	556	567
M05	742	780	635	678	635	709	696
Mean	765	726	588	596	572	586	612
S.D.	44	61	57	56	52	82	58
ID	College student controls						
	Nconceal	Conceal	U1	U2	U3	U4	U5
Mean	582	546	456	462	463	458	460
S.D.	88	101	79	77	81	77	81

[a]*Note.* Mean response latencies are calculated for 'YES' responses to words learned by identity B, and to 'NO' responses to all other words. The college student controls are the 60 participants reported in Allen et al. (1992), who were instructed to deny knowledge of a previously learned list (Conceal) and to acknowledge learning a recently learned list (NConceal).

[b]Reaction time estimated as the 'YES' button failed to record latencies for this particular participant. This estimate was derived by taking the ratio of A/B response latencies for the other three participants, and using this scaling factor to estimate this participant's response latency for this condition. Re-running the analysis without this participant yielded virtually identical results. It is worth noting that this ratio of reaction times (A/B) for the DID participants was 0.95, and for the college student controls it was 0.94.

were instructed to deny knowledge of a previous learned list. To compare the DID participants and undergraduate participants, a 7 (List) by 2 (Group: DID, Control) repeated measures analysis of variance was conducted. Data were first recoded to express the number of 'incorrect' responses, that is the number of times each participant did not respond in accord with the expectation of how someone who was amnestic for identity A's words would respond. This meant that 'NO' responses for identity B's words and the 'YES' responses for identity A's words were analyzed. The rationale for analyzing the data in this fashion was that this transformation put all lists on a similar metric, rather than having one list with a disproportionately large value. Moreover, the data expressed in this fashion serve as the input data for the Bayesian combination (see below) of behavioral indicators. The anticipated main effect for List ($F_{6,372} = 24.8$, $P < 0.001$) indicated that participants made more incorrect responses to the learned lists than to the unlearned lists ($P < 0.05$ by Student Newman Kuhls). Additionally participants made significantly more errors ($P < 0.05$ by SNK) to the recently learned list (identity B's words) than they did to the previously learned list (identity A's list). A main effect for Group ($F_{1,62} = 5.3$, $P < 0.05$) indicated that DID participants made more errors overall, but the absence of a List by Group interaction ($F < 1$) indicated that their tendency to commit more errors was not isolated to any particular list.

Table 3 shows the mean response latency for 'correct' responses; i.e. indicating 'YES' for identity B's items, 'NO' for identity A's items, and 'NO' for all unfamiliar items. The rationale for examining response latencies for only such trials was based on: (1) the fact that — in most cases — there were very few incorrect responses on which to calculate a reliable response latency; (2) the fact that this method was used in analyzing the data from the college student control subjects (Allen et al., 1992); and (3) the fact that the ERPs were analyzed using only these trials. As shown in

Table 3, the DID participants' pattern of response latencies were very similar to that demonstrated by the college-student control subjects. Because the DID participants were slower overall ($F_{1,62} = 13.0$, $P < 0.001$), we examined within subject Z-scores to determine whether the pattern of response latencies was the same for the two groups, after removing the overall difference in reaction time. The main effect of list ($F_{6,372} = 94.6$, $P < 0.001$) and the absence of a List by Group interaction ($F_{6,372} = 1.2$, ns) indicated that all subjects had longer response latencies to the learned words than to the unlearned words ($P < 0.05$ by SNK). Thus even though DID participants responded 'NO' to words learned to identity A and to unlearned words, they were significantly slower to responds to words learned by identity A than to unlearned words. These response data can also inform whether — at an individual subject level — a list of words had been learned. In the validity studies of Allen et al. (1992), the Bayesian combination of response latency and incorrect responses accurately identified 96% of learned items and 98% of unlearned items. As can be seen in Table 4, these behavioral indicators perfectly identify word lists learned by both identities as having been learned, and very seldom (2 times among 20 chances) identify unlearned words as having been learned. On the basis of slowed reaction time and an increased propensity to make incorrect responses (i.e. indicating 'NO' to identity B's words, and indicting 'YES' to identity A's words or to unlearned words), these behavioral measures suggest that identity A's words are influencing identity B's behavior in much the same way identity B's words do. Moreover, these behavioral measures suggest that identity A's words are influencing identity B's behavior in much the same way as previously learned words affect the college student controls — who deliberately concealed their recognition of previously-learned items.

3.2. Electrophysiological data during identity B's recognition task

The upper region of Fig. 1 presents the ERP waveforms for each of the four DID participants and four selected control subjects, and the lower panel presents the grand average waveforms across all four DID participants and all 60 college student control subjects. Note the inter-subject variability in waveform morphology. Descriptively, while all DID participants exhibit a larger P3 amplitude (maximum amplitude from 350 to 850 ms) to words learned by identity B than to unlearned words, only participants M02 and M05 exhibit a larger P3 amplitude to words learned by identity A than to unlearned words. Participants M03 and M04 , on the other hand, do not exhibit

Table 4
Bayesian posterior probability that a list was recognized as learned for DID Participants, calculated on the basis of behavioral responses[a]

Preditor		Sensitivity	(C Sample)	Specificity	(C Sample)	Z Cutscore	
Response latency		1.000	(0.950)	1.000	(1.000)	0.383	
Incorrect responses		0.875	(0.750)	0.900	(1.000)	−0.153	
ID	Identity B	Identity A	U1	U2	U3	U4	U5
M02	1.00	1.00	0.01	0.01	0.01	0.01	0.01
M03	1.00	1.00	0.01	0.01	1.00	0.01	0.01
M04	1.00	1.00	0.01	0.01	0.01	0.01	0.01
M05	1.00	1.00	0.01	0.01	0.01	1.00	0.01

[a]*Note.* Sensitivities and specificities of the indicators for the DID sample are presented, along with the values from the comparison sample (C Sample). Z cutscore = the cutscore established in the study of Allen et al. (1992) and applied here. Values in the lower panel of the table indicate the probability that a list of words was recognized as learned by identity B, based on response latency and the number of 'incorrect' responses (see text). For example, for M02, there is an extremely high probability (rounds to 1.00) that words learned by both identity A and identity B appear recognized, and a very low probability (rounds to 0.01) that any of the five unlearned lists appear recognized.

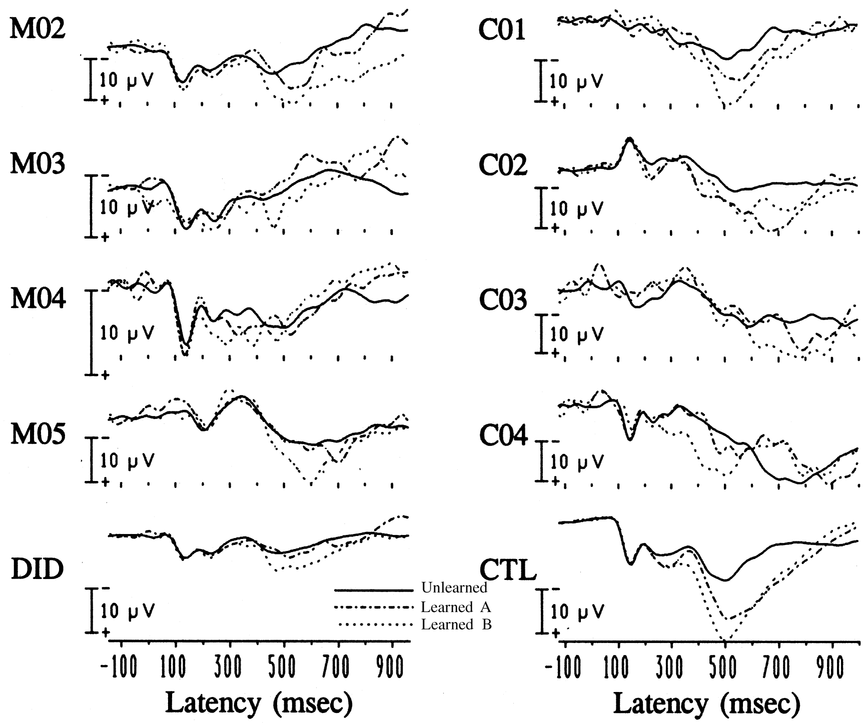

Fig. 1. Left column shows identity B's ERPs at site Pz for words learned by identity A and identity B, and for unlearned words. Note that participants M02 and M05 reported some degree of mutual awareness between personalities, while M03 and M04 did not. Right column shows comparable data from four college student controls (labeled C01–C04, selected from among the 60). At the bottom of each column are grand average ERP waveforms at site Pz for DID participants and college-student controls. DID participants were tested as identity B for recognition of words learned by identity A, identity B, and unlearned words. College student controls were tested for recognition of words for which they denied and concealed knowledge (labeled 'Learned A'), for which they acknowledged learning (labeled 'Learned B'), and for words they had not learned. Only trials where participants responded with recognition of 'Learned B' items and denied recognition of 'Learned A' items and unlearned items are included in the ERP averages. Top panel plots each subject on a different microvolt scale to maximize viewing; Bottom panel plots DID and controls on the same scale. The Bayesian procedure identified both learned lists as recognized for all eight subjects depicted in the figure, despite considerable inter-subject variability in the appearance of P3.

a clearly identifiably P3 peak, and the amplitude in the P3 range to words learned by identity A that is in the range of amplitudes seen for unlearned words. These descriptive statements are based not only on visual inspection of Fig. 1, but also upon a comparison of the present data to the normative data of the college student controls. In the Allen et al. (1992) study, P3 amplitudes were first converted to within subject Z-scores in order to observe the pattern of amplitudes within a subject, independent of between subjects' variations in amplitudes. A cutscore of a $Z = 0.114$ was then established to maximally separate ERPs to learned words from ERPs to unlearned words. (This cutscore was derived on a sample of 20 individuals and cross-validated on two independent samples of 20 individuals each.) Using this cutscore with the DID participants' ERP data, P3 amplitude (maximum amplitude in the 350–850-ms range) for identity B's words exceeded the cutscore for all four participants,[5] but P3 amplitude for identity A's words exceeded this cutscore for only 50% of the DID participants, specifically

[5]Subjects M03 and M04 fail to show a clearly identifiable P3 peak for words learned by identity B as well, but the amplitude in this region did exceed the cutoff for identity B's words.

Table 5
Bayesian posterior probability that a list was recognized as learned for DID Participants, calculated on the basis of P3 amplitude and four communality (H^2)measures[a]

Predictor		Sensitivity	(C Sample)	Specificity	(C Sample)	Z Cutscore	
P3 Amplitude		0.750	(0.925)	0.750	(0.920)	0.114	
H^2 original		0.875	(0.950)	0.800	(0.920)	−0.139	
H^2 1st derivative		1.000	(0.875)	0.850	(0.810)	0.028	
H^2 2nd derivative		0.875	(0.750)	0.700	(0.740)	0.039	
H^2 deviation		0.875	(0.925)	0.750	(0.920)	−0.108	
ID	Identity B	Identity A	U1	U2	U3	U4	U5
M02	0.98	1.00	0.00	0.87	0.19	0.00	0.00
M03	1.00	0.98	0.00	0.00	0.56	0.01	0.00
M04	1.00	0.98	0.00	0.13	0.00	0.00	0.00
M05	1.00	0.98	0.00	0.00	0.98	0.00	0.00

[a]*Note.* Sensitivities and specificities of the indicators for the DID sample are presented, along with the values from the comparison sample (C Sample). Z cutscore = the cutscore established in the study of Allen et al. (1992) and applied here. Values in the lower panel of the table indicate the probability that a list of words was recognized as learned by identity B, based on the five ERP indicators. Only epochs where identity B acknowledged learning identity B's words, and denied recognizing identity A's words or unlearned words were included in this analysis.

those reporting some degree of co-consciousness (M02 and M05). For comparison, among the 60 college control subjects, P3 amplitude to the concealed list exceeded this cutscore for 88.3% of the subjects.

Although P3 amplitude was not uniformly sensitive to the words learned by identity A, other features of the ERP waveform were, as indicated by the results of the Bayesian classification procedure (see Table 5). Based solely on features of the ERP waveform that have proven sensitive to recognized items in previous work, there is a greater than 98% chance that ERPs to words learned by either identity (A or B) are indicative of recognized items. It is further important to reiterate that these ERP data were based only on trials where identity B responded 'YES' to indicate recognition of identity B's items, and responded 'NO' to indicate non-recognition of identity A's items.

3.3. Forced-choice task data

As shown in the top panel of Table 6, it appears that participant M04 — when tested as identity B — systematically avoided all of the words learned by identity A (binomial probability $< 2\%$). Participant M05, by contrast, acknowledged 2 of the 6 words learned by identity A, a result not significantly different from chance (i.e. 50%). In terms of evidence of transfer of incidentally learned information, neither M04 nor M05 — when tested as identity B — showed any compelling evidence of incidental learning, acknowledging 6 of 12 and 5 of 12 items, respectively (chance = 50%).

Following this first forced choice task, participants were asked to switch again to identity A, and where then tested for words learned by identity B. While participant M04 did not perform

Table 6
Percentage of times identity B and identity A responded indicating recognition of items seen by the other identity[a]

	Explicit probes	Incidental probes
As identity B, responding to probes for identity A		
M04	0%	50%
M05	33%	42%
As identity A, responding to probes for identity B		
M04	33%	75%
M05	83%	42%

[a]*Note.* Explicit probes contain words learned intentionally by the other identity, while incidental probes contain words that were seen as distractors by the other personality. Chance in each case is 50%.

significantly different from chance performance, M05 recognized 5 of 6 words learned by identity B (binomial probability $< 10\%$), suggesting that her identity A may have been aware of some of the words learned by identity B. For neither participant was there clear evidence of transfer of incidental learning.

4. Discussion

Amnesia is a central descriptive and diagnostic feature of Dissociative Identity Disorder, and one that is typically assessed on the basis of clients' reports of memory disturbance. The present study was designed to provide an objective and independent evaluation of memory in DID, using event-related potentials and indirect behavioral measures.

Virtually every measure provided evidence consistent with the hypothesis that DID participants did not have amnesia between tested identities. In one case (M02), the participant clearly acknowledged this by reporting 'co-consciousness'; nonetheless, her identity B responded during the ERP task as if she did not recognize the items learned by identity A. For participants M03 and M04, neither participant reported that there was any sharing of memory between personalities. For participant M05, she denied 'co-consciousness', but did — as identity B — recall one of the words learned by identity A. But similar to participant M02, she did not acknowledge this word consistently during the ERP task.

4.1. ERP results

While the appearance of a large P3 seems to covary with the subjective report of 'co-consciousness' (as only M02 and M05 had P3 amplitudes to identity A's words that were outside the range of P3 amplitudes to unlearned words), other features of the ERP waveforms (empirically derived features detailed in Appendix A) suggest that the ERPs were elicited by recognized items.

It is provocative that P3 amplitude elicited by identity A's words was reduced in those participants reporting no co-consciousness. P3 amplitude, however, was smaller overall in DID patients (see Fig. 1), suggesting that P3 amplitude alone may not be the most reliable index in these patients. Moreover, the two subjects reporting no co-consciousness (M03 and M04) failed to produce clearly definable P3 peaks to any items. The smaller P3 amplitude among DID patients compared to controls, and the lack of a clearly defined P3 peak in two of the patients, could be due to the fact that the DID patients are older or that they were on medication (see Table 1), or could reflect information processing deficits associated with their psychopathology.

One of the two subjects (M04) with a small P3 to identity A's words showed a clear pattern of malingering on the test designed to assess systematic avoidance of identity A's words. It is unlikely therefore that the small P3 to identity A's words could indicate amnesia in this subject. It is worth noting that P3 amplitude alone has a reasonably high false negative rate, with 12% of the college-student subjects with intact memory failing to show a large P3 amplitude to learned items. These caveats, considered in the context of the entirety of the findings, suggest that the small P3 amplitude in response to identity A's words among those without reported co-consciousness is intriguing, but far from definitive.

The results of the Bayesian classification, which uses features of the ERP that have been empirically derived to be sensitive to previous learning, provide strong evidence that in all cases identity B was familiar with material learned by both identity A and identity B. The procedure has a built in control so that one can determine whether it is vulnerable to false positive or false negative outcomes. The former would be indicated by the proportion of times that the procedure mistakenly identifies unlearned items as learned (4% in the college student sample, and 10% in the DID sample). The latter would be indicated by the number of times the procedure mistakenly identifies the overtly acknowledged list (the recently learned list) as unlearned (3% in the college student sample and 0% in the DID sample).

4.2. Behavioral response data

Behavioral responses also suggest in all cases that the words learned by identity A were familiar to identity B. Just like the college student comparison sample, identity B showed slower 'NO' responses to words learned by identity A relative to unlearned words. In the college student controls we have conceptualized this as a response conflict associated with the requirement to say 'YES' to recently learned words, but 'NO' to previously learned words (Allen and Iacono, 1997). Previously-learned words involve priming of both response channels: 'YES' by virtue of being learned, and 'NO' by virtue of the instruction to deny and conceal knowledge. Note that this response conflict would only arise if the previously learned words were in fact recognized. The magnitude of the slowing is virtually identical in the DID participants and the college student controls (see note in Table 3). It would therefore appear that identity B has knowledge of words learned by identity A. The response data in Table 2 further corroborate this interpretation, as identity B does classify identity A's words as learned more often than she classifies unlearned words as learned. Again, the pattern of results is remarkably similar to the college-student comparison sample (see Table 2), and is consistent with the hypothesis that DID participants — as identity B — were aware of, but denied knowledge of, words learned by identity A.

4.3. Implicit memory vs. intentional production of amnesia

Whereas previous authors have concluded that transfer of information across personalities may be indicative of implicit memory (e.g. Nissen et al., 1988), the present study suggests that this is not always the case. For the two participants for whom the forced-choice task was operative, one displayed a rather clear malingering profile while the other did not. Yet in both cases, the ERP and behavioral data are strongly suggestive of interpersonality recognition. In the report of Nissen et al., 1988, the MPD participant did not show evidence of malingering, as she responded at or above chance in a forced-choice face recognition task to items seen by a different identity, such as faces. So while the Nissen et al. (1988) data are consistent with, but do not unequivocally support, the interpretation of, implicit memory, the present data do not support the interpretation that there is implicit memory across identities.

Of additional interest are the two cases (M02 and M05) where identity B, upon inquiry, recalled some of identity A's words prior to learning identity B's words. One would expect, then, that identity B would acknowledge identity A's words in the recognition task. But as the data in Table 2 reveal, this was clearly not the case. Rather, for these participants identity B responded as if she were amnestic for the items learned by identity A. It is possible that in the short interval between the recall inquiry and the recognition task that identity B became amnestic for identity A's words, but the preponderance of evidence in this report suggests that this possibility is unlikely. The alternative is that identity B interpreted our instruction — to respond yes to any words on the screen that she remembered learning today — to indicate that she should respond yes to only words identity B learned, even if she recognized the words learned by identity A. Upon debriefing, this is in fact what M02 reported. This observation is supportive of the sociocognitive perspective (Spanos, 1994) of DID that holds that the symptoms of DID are context bounded, goal-directed, social behavior produced in response to demand characteristics. It is worth cautioning, however, that simply because the symptoms of DID can be produced in response to demand characteristics, it in no way implies they must always be. On the other hand, the scientific community can view Spanos' perspective as an hypothesis with considerable support, but one that has the potential to be falsifiable if one could find cases that had unequivocal and profound inter-identity amnesia, and that emerged in the documented absence of social demand. While the assessment paradigm detailed in this report could assist in documenting amnesia, the latter stipulation (absence of social demand) is admittedly difficult to conclusively establish.

4.4. Limitations to generalizability

At present, it would be premature to generalize to all diagnosed cases of DID. There may exist some cases that in fact show a dramatically different pattern than that evidenced by the four participants in this study. On the other hand, the demographic profile of the four DID participants in the present study is not that different than many of the case reports in the literature (Spanos, 1994). It is also worth noting that the present study did not, and likely could not, test all possible pairs of identities for any given participant. It may be the case that, even among the four participants tested in this report, some pairs of identities may demonstrate a mutual amnesia that would produce dramatically different results than those seen here.

4.5. Concluding remarks

The present study details an objective method of assessing amnesia in those who meet diagnostic criteria for DID. These results suggest the need to be cautious in making the diagnosis of DID based on self-report, even with the use of a structured clinical interview. The present findings, electrophysiological as well as behavioral, indicate that individuals who appear — by self-report and diagnostic interview — to meet diagnostic criteria for DID may not be assumed to show unequivocal evidence of inter-identity amnesia.

The behavioral data from the present study suggest the utility of screening additional DID cases using only behavioral response (and not ERP) data. Two behavioral indicators — slowed responses to learned items and increased misclassifications of learned items — can be used to derive statistically supported classifications of items as recognized or not recognized. While not many investigators with access to DID patients may have access to an ERP laboratory, most such researchers could certainly implement the present protocol using the behavioral measures. (The present task, as programmed, can be run on any PC computer, and is available from the first author.) If, through such screening, a DID participant or participants were identified that appeared to show objective evidence of mutual amnesia between identities, then a follow-up study with such individuals in an ERP protocol could prove worthwhile.

Acknowledgements

This research was supported, in part, by a grant from the McDonnell-Pew Program in Cognitive Neuroscience. The authors wish to thank the participants for volunteering of their time to participate in the study, and to thank James A. Coan, John Polich, and Cyma Van Pe for helpful comments on an earlier draft of this manuscript. Portions of this data set were presented at the annual meeting of the Society for Psychophysiological Research, Vancouver, BC, Canada, October 1996.

Appendix A. Using Bayesian procedures to arrive at statistically-supported inferences for each individual

Bayesian analysis can be used to compute the probability that items from each of several lists are familiar to a subject. In the procedure detailed in the manuscript, two sets of items are those that subjects have learned, while substantially more sets are those that have not been learned. ERPs are derived for each set of items. Our analytical approach uses several ERP indicators (e.g. P3 amplitude), each of which may imperfectly distinguish previously-seen from new material. While any given indicator may distinguish such items very well, it would seldom do so perfectly. These particular ERP indicators included P3 amplitude and four measures derived from within-subject factor analyses of different ERP waveforms for each subject (see below). These latter factor-analytic measures were devised because they take into account the entire waveform rather than a single feature such as P3 amplitude. Moreover, even among unselected college students, P3 can be rather variable on in individual subject level of analysis (Allen and Iacono, 1997) and it imperfectly discriminates learned from unlearned material (Allen et al., 1992).

In order to provide statistically informed decisions concerning whether a single subject's ERPs are likely to be in response to items that the subject recognizes, we have taken a statistical approach that involves combining several features of the ERP waveform that tend to distinguish ERPs to recognized items from ERPs to unfamiliar items. The essential features of this approach (detailed fully in Allen et al., 1992), involve: (1) using each participant as his or her own control; (2) examining patterns within each participant after removing mean differences between subjects; (3) using measures that are not dependent on any single component of the ERP but that capture aspects of the entire waveform's morphology; and (4) combining information in Bayesian fashion to arrive at the posterior probability for any given individual that an ERP is indicative of recognition. This approach uses ERP data from site Pz, as adding information from additional sites did not significantly improve the overall accuracy of classifying recognized and unfamiliar material (Allen et al., 1992). This may be due, at least in part, to the fact that data from a single site provided highly accurate classifications (94% true positive rate and 96% true negative rate).

A.1. Each participant as own control

First, each participant serves as his or her own control in that the participant sees words that have been previously learned, as well as new words. By comparing ERPs elicited by these two classes of stimuli, one can assess how ERPs to learned items differ from those to unlearned items *for a given participant* rather than with respect to normative data. Additionally, the ERPs elicited by the unlearned items provide an opportunity to assess the extent to which the procedure produces false positive outcomes, that is erroneously concluding that ERPs to unlearned words were likely to have been elicited by learned stimuli.

A.2. Patterns within participants

Second, the pattern across ERPs to different conditions can be examined for a given subject by first computing an ERP to each list: that learned previously, that learned recently, and each of the five unlearned lists. Each of the resultant ERPs is then in response to a single list of words, and comprised of roughly the same number of trials. Any given index, such as P3 amplitude, is calculated for each of these seven ERP waveforms, and is then transformed to a Z-score *for each participant*. Between-subject differences in P3 amplitude are then removed, leaving scores that capture within-subject variations between the seven ERP waveforms.

A.3. Measures that capture the morphology the of the entire waveform

Third, the approach involves deriving measures of the ERP waveform that are not dependent on any single component (e.g. P3 amplitude), but that rely instead on a comparison of the entire waveforms between conditions. This is accomplished by a conducting a factor analysis *for each subject*, using as variables the ERP waveforms elicited by each of the seven lists, and using as observations time points from -100 ms to $+1050$ ms. The resultant 7×7 variance-covariance summarizes the similarities and differences among the waveforms, and serves as the input matrix for the factor analysis. In our previous work (Allen et al., 1992) we noted that ERPs elicited by learned items tended to have larger variance and covariance, and would often load highly on the first or second extracted principal components, although sometimes on the third principal component. The communality (sum of squared loadings) of a variable (i.e. an ERP waveform) across the first three principal components is a convenient way to summarize the results of the factor analysis and indexes the extent to which that ERP waveform has important variance. (Incidentally, the communality across these three extracted components is unchanged by rotation.) In our previous work, we conducted such a factor analysis not only on the original ERP waveforms, but also on three transformations of those waveforms that highlight different aspects of the morphology. The first derivative of the waveform highlights the slope of the waveform at any point, the second derivative highlights the change in slope, and the 'deviation' waveform (calculated as a given ERP minus the grand mean ERP for that subject; cf. Farwell and Donchin, 1991) highlights how an ERP deviates from the average for that subject.

A.4. Bayesian combination

Finally, each of these five indictors — P3 amplitude (maximum amplitude from 350 to 850 ms), and the four communalities from the factor analyses — are combined in Bayesian fashion. For a group of subjects, each indicator will have an associated sensitivity (i.e. the proportion of all ERPs to learned items that show this particular characteristic, such as large P3 amplitude) and specificity (i.e. the proportion of all ERPs to unlearned items that do *not* show this particular characteristic). The sensitivity and the specificity, along with the base rate of ERPs elicited by learned items (i.e. 2/7 of the ERPs) will determine the probability that a given ERP was elicited by learned items given the presence of the indicator. This is determined by:

$$\frac{Ll}{Ll+Uu}$$

where:

- L is the base rate of ERPs elicited by learned items (2/7);
- l is the sensitivity of the indicator;
- U is the rate of occurrence of the ERPs elicited by unlearned items ($1-L$); and
- u is the complement of the specificity ($1-$ specificity) of the indicator, which is simply the rate at which unlearned items produce ERPs with the indicator.

The numerator indexes the proportion of all ERPs that are in response to learned items *and* show the presence of the indicator (e.g. large P3). The denominator indexes the proportion of *all* ERPs (learned as well as unlearned) that show the presence of the indicator.

For example, take the single indicator P3 amplitude. In the validation studies (Allen et al., 1992), 92.5% of ERPs elicited by learned items had a large P3 (sensitivity = 0.925) and 92% of the ERPs elicited by unlearned items did *not* have a large P3 (specificity = 0.92). In these studies — like the current study — learned words comprised 2/7 (0.2857) of the trials, and unlearned words comprised 5/7 (0.7143) of the trials. Thus if an ERP has a large P3, the probability that it was elicited by learned items is:

$$\frac{(0.2857)(0.925)}{(0.2857)(0.925)+(0.7143)(0.08)}=0.82$$

With the single indicator of P3 amplitude, which has reasonably high sensitivity and specificity, there is still only an 82% chance that an ERP with such an amplitude was in fact elicited by learned items. For this reason, it is often helpful to combine multiple indicators, which can increase the statistical confidence with which one can identify ERPs elicited by learned items. Extending the approach to multiple indicators, the probability that any particular list is learned, given a particular combination of n indicators, is then determined by the formula:

$$\text{Prob(Learned/combination of } n \text{ indicators)} = \frac{(L_{l_1 l_2 l_3 \ldots l_n})}{(L_{l_1 l_2 l_3 \ldots l_n})+(U_{u_1 u_2 u_3 \ldots u_n})}$$

where

- L = proportion of ERPs elicited by learned items (i.e. the 'base rate' of 2/7);
- U = proportion of ERPs elicited by unlearned items ($1-L$);
- l_i = sensitivity of indicator i if indicator i indicates the ERP was in response to learned items; or
- complement of sensitivity (1 − sensitivity) for indicator i if indicator i indicates the ERP was in response to unlearned items;

- u_i = complement of specificity (1 − specificity) for indicator i if indicator i indicates the ERP was in response to learned items; or
- specificity for indicator i if indicator i indicates the ERP was in response to unlearned items.

Conceptually, this probability ratio is equal to the *proportion of ERPs elicited by learned items with a given combination of indicators divided by the proportion of all ERPs showing this combination of indicators*. If all indicators suggested the ERP was in response to learned items, the probability would be high, but disagreement between indicators would lower the probability.

A final word concerning the indicators is in order. In the example, the magnitude of the P3 indicators was simply discussed in terms of large or not large. One needs to determine a cutpoint such that P3 amplitude greater than some criterion is considered large. While this could be established based on the raw microvolt values of P3, this approach is less than optimal because individual differences between persons in their mean P3 amplitude — regardless of which list elicited it — are likely to be substantial. An alternative to this normative approach would be to use an intra-individual approach to determining whether an indicator is 'large'. By using within-subject Z-scores as mentioned above, between-subject variation in overall level are removed, and the resultant scores reflect the pattern across the ERPs elicited by the learned and unlearned lists. In the validation studies (Allen et al., 1992), a Z-score cutpoint was chosen for each indicator that optimally separated ERPs to learned from ERPs to unlearned material. This cutpoint was selected based on data from the first of these three validation studies, and cross-validated in the two additional studies (Allen et al., 1992). These are the cutpoints that are used in the present report.

References

Allen, J.J., 1992. The development and validation of an event-related potential memory- assessment procedure: application to hypnotically-induced amnesia. Diss. Abstr. Int. 53, 05B.

Allen, J.J.B., Iacono, W.G., 1997. A comparison of methods for the analysis of event-related potentials in deception-detection. Psychophysiology 34, 234–240.

Allen, J.J., Iacono, W.G., Danielson, K.D., 1992. The development and validation of an event-related-potential memory assessment procedure: a methodology for prediction in the face of individual differences. Psychophysiology 29, 504–522.

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. American Psychiatric Association, Washington.

Bahnson, C.B., Smith, K., 1975. Autonomic changes in a multiple personality. Psychos. Med. 37, 85–86.

Bentin, S., Moscovitch, M., 1990. Psychophysiological indices of implicit memory performance. Bull. Psychonom. Soc. 28 (4), 346–352.

Bentin, S., Moscovitch, M., Heth, I., 1992. Memory with and without awareness: performance and electrophysiological evidence of savings. J. Exp. Psychol.: Learn. Mem. Cogn. 18 (6), 1270–1283.

Bernstein, E.M., Putnam, F.W., 1986. Development, reliability, and validity of a dissociation scale. J. Nerv. Ment. Dis. 178, 727–735.

Bliss, E.L., 1980. Multiple personalities: a report of 14 cases with implications for schizophrenia. Arch. Gen. Psychiatry 37, 1388–1397.

Boaz, T.L., Perry, N.W., Raney, G., Fischler, I.S., Shuman, D., 1991. Detection of guilty knowledge with event-related potentials. J. Appl. Psychol. 76, 788–795.

Braun, B.G., 1983. Neurophysiologic changes in multiple personality due to integration: a preliminary report. Am. J. Clin. Hypn. 26, 84–92.

Brende, J.O., 1984. The psychophysiologic manifestations of dissociation: electrodermal responses in a multiple personality patient. Psychiatry Clin. North Am. 7, 41–50.

Brown, P., 1994. Toward a psychobiological model of dissociation and post-traumatic stress disorder. In: Lynn, S.J., Rhue, J.W. (Eds.), Dissociation: Clinical and theoretical perspectives. The Guilford Press, New York.

Cacioppo, J.T., Tassinary, L.G., 1990. Psychophysiology and psychophysiological inference. In: Cacioppo, J.T., Tassinary, L.G. (Eds.), Principles of Psychophysiology: Physical, Social, and Inferential Elements. Cambridge University Press, New York, pp. 3–33.

Carroll, J.B., Davies, P., Richman, B., 1971. The American Heritage Word Frequency Book. Houghton Mifflin, Boston, MA.

Coons, P.M., 1984. The differential diagnosis of multiple personality: a comprehensive review. Psychiatry Clin. North Am. 7, 51–64.

Coons, P.M., 1988. Psychophysiological aspects of multiple personality disorder: a review. Dissociation 1, 47–53.

Coons, P.M., Milstein, V., Marley, C., 1982. EEG studies of two multiple personalities and a control. Arch. Gen. Psychiatry 39, 823–825.

Eich, E., Macaulay, D., Loewenstein, R.J., Dihle, P.H., 1997. Implicit memory, interpersonality amnesia, and Dissociative Identity Disorder: comparing patients with simulators. In: Read, J.D., Lindsay, D.S. (Eds.), Recollections of Trauma: Scientific Evidence and Clinical Practice. Plenum Press, New York, pp. 469–474.

Farwell, L.A., Donchin, E., 1991. The truth will out: interrogative polygraphy ('lie detection') with event-related potentials. Psychophysiology 28, 531–547.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. 'Mini-Mental States'. A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Greaves, G.B., 1980. Multiple personality 165 years after Mary Reynolds. J. Nerv. Ment. Dis. 168, 577–596.

Kihlstrom, J.F., Tataryn, D.J., Hoyt, I.P., 1993. Dissociative disorders. In: Sutker, P.B., Adams, H.E. (Eds.), Comprehensive Handbook of Psychopathology, 2nd ed.. Plenum, New York.

Larmore, K., Ludwig, A.M., Cain, R.L., 1977. Multiple personality: an objective case study. Br. J. Psychiatry 131, 35–40.

Ludwig, A.M., Brandsma, J.M., Wilbur, C.B., Bendfeldt, F., Jameson, D.H. (1972). The objective study of a multiple personality: or, are four heads better than one? Arch. Gen. Psychiatry, 26, 298–310.

Lykken, D.T., 1959. The GSR in the detection of guilt. J. Appl. Psychol. 43, 385–388.

Mathew, R.J., Jack, R.A., West, W.S., 1985. Regional cerebral blood flow in a patient with multiple personality. Am. J. Psychiatry 142, 504–505.

Miller, S.D., Triggiano, P.J., 1992. The psychophysiological investigation of multiple personality disorder: review and update. Am. J. Clin. Hypn. 35, 47–61.

Miller, S.D., 1989. Optical differences in cases of multiple personality disorder. J. Nerv. Ment. Dis. 177, 480–486.

Miller, S.D., Blackburn, T., Scholes, G., White, G., Mamalis, N., 1991. Optical differences in cases of multiple personality disorder: a second look. J. Nerv. Ment. Dis. 179, 132–135.

Nissen, M.J., Ross, J.L., Willingham, D.B., Mackenzie, T.B., Schachter, D.L., 1988. Memory and awareness in a patient with multiple personality disorder. Brain Cogn. 8, 117–134.

Neville, H.J., Kutas, M., Chesney, G., Schmidt, A.L., 1986. Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. J. Mem. Lang. 25, 75–92.

Paller, K.A., Kutas, M., 1992. Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. J. Cognit. Neurosci. 4 (4), 375–391.

Perlini, A.H., Spanos, N.P., 1991. EEG alpha methodologies and hynotizability: a critical review. Psychophysiology 28, 511–530.

Plotkin, W.B., 1979. The alpha experience revisited: biofeedback in the transformation of psychological state. Psychol. Bull. 86, 1132–1148.

Putnam, F.W., 1984. The psychophysiologic investigation of multiple personality disorder: a review. Psychiatr. Clin. North Am. 7, 31–39.

Putnam, F.W., Zahn, T.P, Post, R.M., 1990. Differential autonomic nervous system activity in multiple personality disorder. Psychiatry Res. 31, 251–260.

Rosenfeld, J.P., Angell, A., Johnson, M., Qian, J.H., 1991. An ERP-based, control-question lie detector analog: algorithms for discriminating effects within individuals' average waveforms. Psychophysiology 28, 319–335.

Rosenfeld, J.P., Cantwell, B., Nasman, V.T., Wojdac, V., Ivanov, S., Mazzeri, L., 1988. A modified event-related potential-based guilty-knowledge test. Int. J. Neurosci. 42, 157–161.

Rosenfeld, J.P., Reinhart, A.M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., Sweet, J., 1998. P300 correlates of simulated amnesia on a matching-to-sample task: topographic analyses of deception vs. truth-telling responses. Int. J. Psychophysiol. 28, 233–248.

Ross, C.A., Heber, S., Norton, G.R, Anderson, G., 1989a. Differences between multiple personality disorder and other diagnostic groups on structured interview. J. Nerv. Ment. Dis. 179, 487–491.

Ross, C.A., Norton, G.R., Fraser, G.A., 1989b. Evidence against the iatrogenesis of multiple personality disorder. Dissociation 2, 61–65.

Rugg, M.D., 1995. Event-related potential studies of human memory. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. MIT Press, Cambridge MA, pp. 789–802.

Rugg, M.D., Doyle, M.C., 1992. Event-related potentials and recognition memory for low- and high-frequency words. J. Cogn. Neurosci. 4, 69–79.

Rugg, M.D., Doyle, M.C., 1994. Event-related potentials and stimulus repetition in direct and indirect tests of memory. In: Heinze, H., Munte, T., Mangun, G.R. (Eds.), Cognitive Electrophysiology. Birkhauser, Boston, pp. 124–148.

Rugg, M.D., Nagy, M.E., 1989. Event-related potentials and recognition memory for words. Electroencephalogr. Clin. Neurophysiol. 72, 395–406.

Sanquist, T.F., Rohrbaugh, J.W., Sydulko, K., Lindsley, D.B., 1980. Electrocortical levels of processing: perceptual analysis and recognition memory. Psychophysiology 17, 568–576.

Semlitsch, H.V., Anderer, P., Schuster, P., Presslich, O., 1986. A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP. Psychophysiology 23, 695–703.

Sheikh, A.A., Kunzendorf, R.G., 1984. Imagery, physiology, and psychosomatic illness. In: Sheikh, A.A. (Ed.), International Review of Mental Imagery, 1. Human Sciences Press, New York, pp. 95–138.

Silberman, E.K., Putnam, F.W., Weingartner, H., Braun, B.G., Post, R.M., 1985. Dissociative states in multiple personality disorder: a quantitative study. Psychiatry Res. 15, 253–260.

Smith, M.E., 1993. Neurophysiological manifestations of recollective experience during recognition memory judgements. J. Cogn. Neurosci. 5, 1–13.

Smith, M.E., Halgren, E., 1989. Dissociation of recognition memory components following temporal lobe lesions. J. Exp. Psychol.: Learn. Mem. Cogn. 15, 50–60.

Spanos, N.P., 1994. Multiple identity enactments and multiple personality disorder: a sociocognitive perspective. Psychol. Bull. 116, 143–165.

Spanos, N.P., 1996. Multiple identities and false memories: a sociocognitive perspective. American Psychological Association, Washington.

Steinberg, M., 1993. Structured Clinical Interview for DSM-IV Dissociative Disorders (SCID-D). American Psychiatric Press, Washington, DC.

Thigpen, C.H., Cleckley, H.M., 1957. The Three Faces of Eve. Popular Library, New York.