Unconscious learning during surgery with propofol anaesthesia

C. Deeprose1*, J. Andrade1, S. Varma2 and N. Edwards3

1Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TP, UK. 2Sunderland Royal Hospital, Sunderland, UK. 3Northern General Hospital, Sheffield, UK
*Corresponding author. E-mail c.deeprose@shef.ac.uk

Background. Learning during anaesthesia has been demonstrated, but little is known about the circumstances under which it may occur. This study investigated the hypothesis that learning during anaesthesia occurs during, but not before, surgical stimulation.

Methods. Words were played through headphones to 64 day-surgery patients during propofol anaesthesia. Fourteen words were played repeatedly (15 times) for 1 min each either before (n=32) or during (n=32) surgical stimulation. The depth of anaesthesia was estimated using the bispectral index (BIS). Heart rate, ventilatory frequency, mean arterial pressure, end-tidal carbon dioxide concentration, and infusion rate of propofol were recorded at 1 min intervals during word presentation. On recovery, memory was assessed using an auditory word stem completion test and word recognition test.

Results. The mean BIS, arterial pressure, end-tidal carbon dioxide and heart rate during word presentation did not differ between the groups. The infusion rate of propofol and the ventilatory frequency were significantly greater in the during-surgical stimulation group. There was no evidence for explicit recall or recognition, nor of awareness during anaesthesia (median mean-BIS=38 in the before-surgical stimulation group and 42 in the during-surgical stimulation group). Only patients who were played words during surgical stimulation showed significant implicit memory on recovery (mean score=0.08, P<0.02) However, their scores were not significantly higher than those of the before-surgical stimulation group (mean score=0.01).

Conclusions. Learning during anaesthesia seems more likely to occur during rather than before surgical stimulation at comparable anaesthetic depth. We hypothesize that surgical stimulation facilitates learning during anaesthesia, independently of its effects on anaesthetic depth.


Keywords: anaesthesia, depth; memory, implicit; surgery

Accepted for publication: September 13, 2003

In a typical study of learning during anaesthesia, words are played to patients during general anaesthesia and, on recovery, the patients are tested for explicit and implicit memory of those words. Explicit memory refers to conscious recall or recognition of the words. Evidence for explicit memory would suggest conscious learning of the words during surgery (awareness), yet many studies have used insensitive tests of explicit memory (e.g. simply asking patients for their recollections of surgery). Implicit memory refers to a change in behaviour, for example an increased tendency to complete word stems (e.g. TRA-) with studied words (e.g. TRACTOR), unaccompanied by conscious recollection of the studied words.

Reviews of learning during anaesthesia present an array of positive and negative findings, with around half of all studies demonstrating significant results.1–3 Merkle and Daneman4 found statistically significant implicit memory in their meta-analysis of learning during anaesthesia, providing that learning was tested within 36 h of surgery. However, little is known about the circumstances under which learning during anaesthesia may occur. We hypothesized that two factors contribute to learning during anaesthesia: depth of anaesthesia and surgical stimulation. Although the evidence for a contribution of depth of anaesthesia is mixed,5–9 a study by Lubke and colleagues⁶

---

© The Board of Management and Trustees of the British Journal of Anaesthesia 2004
provided quite clear evidence for a relationship between depth and learning by studying trauma patients subject to wide variations in anaesthetic depth. They found a significant correlation between bispectral index™ (BIS™) and learning. However, BIS™ explained a relatively small proportion (12%) of the variance in patients’ memory scores, suggesting that depth of anaesthesia is not the only determinant of learning.

We have hypothesized that surgery facilitates learning during anaesthesia, independently of its effects on depth of anaesthesia.8 10 12 Surgical stimulation triggers a rapid increase in blood and brain concentrations of catecholamines and cortisol, which have been shown to exert influences via the amygdala on learning and memory in both animal and human subjects.11 Tentative support for a role of surgery comes from recent studies showing no learning during light sedation before surgery.10 12 However, Kerssens and colleagues11 did find learning before surgery despite deeper sedation.

The present study tested the hypothesis that surgery facilitates learning by investigating learning of words presented after induction of propofol anaesthesia, but either before or during surgical stimulation. We manipulated the presence of surgical stimulation during word presentation while holding depth of anaesthesia constant. As we were interested in the effects of arousal, we checked comparability of anxiety levels between our two study groups, using a pre-operative anxiety questionnaire, the State Trait Anxiety Inventory.14 We measured depth of anaesthesia using BIS™, and pilot tested our memory tests to ensure they were sensitive tests of explicit and implicit memory.

Methods

Approval for the study was obtained from North Sheffield Medical Ethics Committee. Only adults (ASA I) undergoing day surgery and who spoke English as their first language were considered. Patients with known visual or hearing impairments, language difficulties, neurological disorders, contra-indications to the proposed anaesthetic technique, or who were taking medication known to affect the central nervous system were excluded. Written informed consent was obtained from 65 patients. Patients were assigned randomly to one of two experimental groups, one group receiving stimulus words post-induction but before surgical stimulation and the other receiving the words during surgical stimulation.

Construction of memory test

Practically all memory tests can be performed using a combination of implicit and explicit memory. We wanted tests that were relatively pure, that is an implicit memory test on which performance was minimally influenced by explicit memory and vice versa. We also needed sensitive tests because previous research suggests only small amounts of learning during anaesthesia.4 We therefore developed our memory tests as follows.

In a laboratory study with an undergraduate sample, we compared the purity and sensitivity of three common tests of implicit memory (word stem completion, word identification, and a modified version of the process dissociation procedure using word stem completion) and three tests of explicit memory (yes/no recognition, forced choice recognition, and the modified process dissociation procedure) using auditory presentation and test. Participants (n=48) were asked to try and remember one list of words (full attention condition) and were played another list of words while completing a demanding visual task (divided attention condition). Condition order, and presentation and test lists were counterbalanced. The rationale for using an attentional manipulation is that implicit memory tends to be insensitive to such manipulations, in contrast to explicit memory.15 Relatively pure measures of implicit memory will therefore be less affected by the manipulation than measures that are more contaminated by explicit memory, whereas relatively pure measures of explicit memory will be affected more than less pure measures. We used the following operational definitions: sensitivity was defined as effect size in the divided attention condition, that is, the extent of the difference between responses with or to studied and distractor words. For implicit memory, a pure test was one that showed minimal effect of dividing attention (compared with full attention). For explicit memory, a pure test was one that showed a large effect of dividing attention. On the basis of this pilot work, we chose word stem completion as the implicit memory test (effect size in divided attention condition=0.44; the effect of the attentional manipulation was not statistically significant) and yes/no recognition as the explicit memory test for the current study (effect size in divided attention condition=0.95; the effect of the attentional manipulation was statistically significant).

From the pilot study, 28 words with low unstudied stem completion rates were selected for the learning during anaesthesia study. Words were recorded onto a Macintosh Powerbook (1400cs/133, Apple Computer Inc., CA) at a sample rate of 44.1 kHz and 16 bit sample size using a microphone (Apple Plaintalk) and SoundEdit software (16 version 2, Macromedia, US). In a copy of the word file, SoundEdit was used to remove the tail of each word, leaving an auditory stem approximately three phonemes long.

Four word lists were created, matched for proportion of correct spontaneous completion rates in the pilot study (mean=0.32) and frequency (mean=50.1), familiarity (mean=525.5), imageability (mean=495.3), and concreteness (mean=463.5).

Two word lists were presented an equal number of times to patients in each study group (before-surgical stimulation and during-surgical stimulation). The words in the study lists were presented in a different random order to each patient. For the explicit, yes/no recognition test, patients heard a randomized series of word stems taken from one of
the studied word lists and one of the distractor word lists and decided which words they recognized hearing during anaesthesia. For the implicit, word stem completion test, patients heard a series of word stems from the remaining study and distractor word lists and responded with the first completion that came to mind. Each word list appeared equally often as target and distractor stimuli on both the implicit and explicit tests. Test order (explicit then implicit and vice versa) was fully counterbalanced to maximize the probability of detecting explicit as well as implicit memory.

**Anaesthetic technique and experimental procedures**

Within 2 h before the surgery, patients completed the State Trait Anxiety Inventory as a measure of pre-operative anxiety.

BIS™ monitoring commenced in the anaesthetic room using an Aspect-1000 monitor (software version 2.51; Aspect Medical Systems, Framingham, MA, USA) with bifrontal montage (F7, F8, reference Fp2). BIS™ was monitored continuously throughout the anaesthetic period. Although the anaesthetist was not blind to the BIS™ score, only standard clinical signs were used to maintain anaesthesia.

The patients did not receive preoperative medication. At induction of anaesthesia patients received fentanyl 1.5 µg kg⁻¹ followed by a ‘sleep’ dose of propofol. Anaesthesia was maintained with an infusion of propofol at 5–7 mg ml⁻¹ h⁻¹ depending on patient response. Patients breathed nitrous oxide 66% and oxygen 33% spontaneously through a laryngeal mask. Postoperative analgesia was given as oxycodone 66% and oxygen 33% spontaneously through a laryngeal mask. Postoperative analgesia was given as expressed as a proportion of the total number of studied items on the test (i.e. 7).

In the word stem completion test, patients were presented with auditory stems of seven target and seven distractor words. Each stem was presented twice, whereupon patients were asked to complete the stem with the ‘first complete word, which comes to mind’. Responses on the word stem completion test were scored as ‘hits’ if they were words from the study or distractor list. Implicit memory scores were calculated by subtracting the number of hits on the distractor stems from the number of hits on the target stems, and presented as proportions.

**Statistical methods**

Mean-BIS™ scores recorded during word presentation were compared across groups using the Mann–Whitney test for independent samples. The heart rate, mean arterial pressure, end-tidal carbon dioxide concentration, ventilatory frequency, and infusion rate of propofol were compared between groups using two tailed t-tests for independent samples. One-tailed, one sample t-tests were used to test whether the memory scores in the before-surgery and during-surgery study groups exceeded zero. Pearson’s correlations were used to test the relationship between memory scores and anaesthetic variables. Spearman’s rank correlations were used to test the relationship between mean-BIS™ and implicit memory score on recovery. For these analyses, we report rho values corrected for ties. Statistical significance was assessed with alpha=0.05 unless otherwise stated. All analyses were performed using StatView 5.0 (SAS Institute Inc., Cary, NC).

**Results**

We consented a total of 65 patients. One patient in the before-surgical stimulation group refused to complete the memory test on recovery, and was replaced, to give 32 patients in each group. Age, weight, anxiety levels (trait or state), duration of surgery, time to test, and types of surgery
in each group are shown in Table 1. The mean time between the end of surgery and the onset of testing was 69 min.

Anaesthetic variables are shown in Table 2. Only ventilatory frequency and the mean amount of propofol given during word presentation varied significantly between the two groups. There were no significant differences in anaesthetic depth as measured by BIS™. Seven patients in the before-surgical stimulation group and four patients in the during-surgical stimulation group received between 5 and 10 mg morphine postoperatively. Preliminary analyses showed no effect of test order (implicit then explicit or vice versa) on explicit or implicit memory performance. No patient had spontaneous recall for intra-operative events and there was no evidence for explicit memory using the structured postoperative interview. Explicit memory scores on the yes/no recognition test did not exceed zero in either group, and did not differ across groups (Table 3).

Implicit memory (Table 3) exceeded zero for patients who received the stimulus words during surgical stimulation ($P<0.02$) but not for patients who received the words before-surgical stimulation ($P>0.15$). However, implicit memory scores did not differ significantly between the two groups.

In the during-surgical stimulation group, implicit memory did not correlate significantly with time to testing ($r=0.03$), length of surgery ($r=0.28$), dose of propofol ($r=0.11$), state anxiety ($r=-0.09$), or trait anxiety ($r=0.19$). Mean-BIS™ showed a stronger, but still non-significant correlation with implicit memory ($r=0.30$, $P=0.10$). However, closer inspection of the data (Fig. 1) suggested that this apparent correlation might be spurious, because of a significant decrease in hits to distractor items as mean-BIS™ increased ($r=0.38$) rather than an increase in hits to target items ($r=-0.06$). Because distractor hits might decrease because patients more often fail to think of any word that completes a stem, we also tested the relationship between mean-BIS™ and number of omitted responses to distractor items. This correlation was also non-significant ($r=0.11$).

### Discussion

We compared learning of words presented to anaesthetized patients before or during surgical stimulation. No patient had explicit recall of the anaesthetic period, nor explicit recognition of the stimulus words, despite using counter-balancing of the order of the explicit and implicit memory tests to maximize the accuracy of our assessment of explicit as well as implicit memory. Patients who were played words before surgical stimulation did not show implicit memory for the words, as tested by an auditory word stem completion task. Patients who were played words during surgical stimulation did show implicit memory for the words; their implicit memory scores were significantly greater than zero even if alpha is adjusted to 0.025 to allow for multiple comparisons.

### Table 1

| Mean patient characteristics, surgery duration and time from end of surgery to memory testing (range and (SD)) |
|---|---|---|---|---|---|---|---|---|
| **Age (yr)** | **Weight (kg)** | **Sex ratio male:female** | **Trait anxiety** | **State anxiety** | **Types of surgery** | **Surgery duration (min)** | **Time to test (min)** |
| Before surgical stimulation (n=32) | | | | | | | |
| 41 (20–75; 13.8) | 75.5 (50–108; 14.9) | 17:15 | 36.2 (21–55; 8.1) | 37.5 (22–62; 11.2) | Hernia repair: 11; varicose vein removal: 9; hernia repair: 3; misc.: 9 | 31 (9–55; 10.8) | 70 (30–155; 30.5) |
| During surgical stimulation (n=32) | | | | | | | |
| 44 (17–68; 13.7) | 76.4 (55–113; 14.8) | 18:14 | 35.1 (20–52; 7.3) | 39.3 (21–73; 13.4) | Hernia repair: 8; hernia repair: 9; misc.: 8 | 33 (14–98; 20.8) | 68 (25–140; 32.7) |

### Table 2

Median mean-BIS™ (interquartile range), and heart rate (HR), mean arterial pressure (MAP), end-tidal carbon dioxide ($E\phi CO_2$), ventilatory frequency, and propofol infusion rate (mean (SD)). (Mean-BIS™ is the mean of the 28 bispectral index™ values recorded during word presentation for each patient). *$P<0.001$; **$P<0.01$.

<table>
<thead>
<tr>
<th><strong>Mean-BIS™</strong></th>
<th><strong>HR (min⁻¹)</strong></th>
<th><strong>MAP (mmHg)</strong></th>
<th><strong>$E\phi CO_2$ (kPa)</strong></th>
<th><strong>Ventilatory frequency (min⁻¹)</strong></th>
<th><strong>Propofol infusion rate (mg ml⁻¹)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Before surgical stimulation</td>
<td>37.9 (10.6)</td>
<td>66 (9.3)</td>
<td>69 (7.4)</td>
<td>6.8 (0.7)</td>
<td>9 (3.6)</td>
</tr>
<tr>
<td>During surgical stimulation</td>
<td>43.6 (16.4)</td>
<td>72 (16.6)</td>
<td>75 (22.9)</td>
<td>6.7 (0.8)</td>
<td>15 (4.9)</td>
</tr>
</tbody>
</table>

### Table 3

Explicit and implicit memory for patients receiving word presentation before or during surgical stimulation.

<table>
<thead>
<tr>
<th><strong>Mean implicit memory (with SD) 95% CI</strong></th>
<th><strong>Mean explicit memory (with SD) 95% CI</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Before surgical stimulation</td>
<td>0.02 (0.19) -0.05 to 0.09</td>
</tr>
<tr>
<td>During surgical stimulation</td>
<td>0.08 (0.19) 0.01 to 0.15</td>
</tr>
</tbody>
</table>

Deeprose et al.
for repeated \( t \) testing. This is the first replication, within a single study, of the trend in the literature to find learning during surgery\(^4\) but no learning at similar or lighter anaesthetic depths in the absence of surgery\(^{10, 12}\) (but see \(^{13}\)).

However, the apparent effect of surgery is weak. The difference between the implicit memory scores for the before-surgical stimulation and the during-surgical stimulation groups was not statistically significant. Because of resource constraints (including the need for an additional period of anaesthesia in the before-surgery group), we chose our sample size to give good power (0.80) to detect learning in either patient group, based on the effect size of 0.44 for the word stem completion test in the divided attention condition of the pilot study. The actual effect size for the during-surgical stimulation group in this experiment was 0.40. This is a relatively good effect size in this field. Compared with the studies included in the meta-analysis by Merikle and Daneman,\(^4\) this is lower than the maximum of 0.59 demonstrated by Roorda-Hrdlickova and colleagues\(^ {16}\) but considerably higher than the mean effect size of 0.23 for studies testing recall within 12 h, as one might expect given our short retention interval and pilot testing of the stimulus materials.

BIS\(^{TM}\) readings showed no significant difference in anaesthetic depth during the two word presentation periods. Thus, the effect of surgical stimulation on learning is not attributable to changes in depth of anaesthesia, which were counteracted by giving more propofol. Together with clinical observations, the BIS\(^{TM}\) readings indicated clinically adequate depths of anaesthesia. Patients received no neuromuscular blocking drugs and therefore were free to move if they became conscious. We therefore conclude that the learning observed during surgical stimulation was unconscious, and not the product of undetected periods of awareness. This conclusion is supported by the correlational analysis of mean-BIS\(^{TM}\) during word presentation and implicit memory. Although there was a marginally significant correlation between these variables, it appeared to be because of a baseline performance getting worse as the anaesthetic got lighter rather than to target performance improving. Poorer baseline performance could arise because patients more often fail to complete distractor word stems, reducing their chance of a ‘hit’. Patients who were more lightly anaesthetized might have omitted more responses because they were experiencing more pain or receiving more analgesia on recovery than patients who were more deeply anaesthetized. This does not seem to have been the case: omission rates did not correlate with mean-BIS\(^{TM}\) scores. Lubke and colleagues\(^6\) found no correlation between BIS\(^{TM}\) and baseline performance on a word stem completion task (note that they tested patients much later, 49 h after surgery compared with 69 min in our study). We therefore assume that the correlation in our study is spurious. Thus, our data provide little support for the hypothesis of a relationship between probability of learning and depth of anaesthesia. This is unsurprising, because depth of anaesthesia in our study was held relatively constant. Lubke and colleagues’ study of trauma patients,\(^6\) with wide variation in BIS\(^{TM}\), provides a stronger test of the relationship between depth of anaesthesia and learning. However, their finding of a significant but relatively weak correlation between BIS\(^{TM}\) and memory also indicates that depth of anaesthesia alone cannot predict whether learning will occur.

Our speculative explanation of the facilitatory effect of surgery on learning during anaesthesia focuses on the fact that the amygdala is implicated in memory formation and is activated by the stress hormones that are released during surgery. The amygdala plays an important role in fear conditioning (see review by LeDoux\(^ {17}\)) and this mechanism may also facilitate enhanced perception and encoding of small amounts of verbal implicit memory. Modality specific neocortical regions form the locus of the perceptual priming that our word stem completion test is detecting (see review by Gabrieli\(^ {18}\)), and also form part of the implicit fear conditioning circuitry.\(^ {17}\) However, whether or not perceptual priming may be enhanced by concurrent arousal in the same way as implicit fear conditioning has yet to be tested.
Nonetheless, this hypothesis provides a tentative explanation for the tendency of surgery to facilitate learning during anaesthesia, and for the relatively weak relationship between depth of anaesthesia and learning in Lubke and colleagues’ study.6 The trauma patients tested by Lubke varied not only in depth of anaesthesia but also in the severity of trauma experienced and the consequent extent of surgery required. Given that both trauma and surgery would trigger catecholamine release, we assume that there would have been considerable variation in catecholamine concentrations across the sample.

An alternative explanation, which also implicates the amygdala, is that learning during surgery taps the memory-enhancing mechanism that ensures we remember emotional events particularly well. Emotional events trigger release of catecholamines that activate amygdala-based enhancement of memory consolidation processes occurring in other brain regions such as the hippocampus.19 Administering epinephrine to animals and humans mimics the effect of emotion and thus improves memory for neutral stimuli.20 21 Thus, in human subjects, surgery might enhance memory consolidation even for neutral word lists. However, this evidence comes from studies of explicit memory and it is not known whether the same mechanism enhances encoding of implicit memories such as those measured in the present study.

Note that we have referred to the formation of implicit memories as ‘learning’. ‘Learning’ implies the encoding of new information or the formation of new associations between existing representations in memory. Forming an explicit memory of a list of familiar words involves forming new associations between those words and their experimental context. Implicit memory formation probably does not involve forming such associations22 and is often referred to not as learning but as ‘priming’, meaning the temporary activation of existing mental representations. However, we have used the term ‘learning during surgery’ for continuity with the existing anaesthesia literature.

A confound in our study is the use of a small dose of fentanyl at induction. Fentanyl has a short half-life23 hence the absence of learning in the before-surgical stimulation group could be attributable to the effect of fentanyl, which would have been waning by the time the words were played to the during-surgical stimulation group. We are aware of no evidence to suggest that fentanyl impairs implicit memory formation. It has almost no effect on explicit recall in sedated volunteers24 and neither does the related drug alfentanil.25 As explicit memory tends to be more sensitive to drug effects than implicit memory,26 it seems unlikely that fentanyl prevented learning in the before-surgery group. Fentanyl may well reduce learning during anaesthesia, but we hypothesize that any effects it has are a result of its suppression of the stress response to surgery27 and not direct effects on learning. Another confound is that patients in the before-surgical stimulation group were anaesthetized for longer than the during-surgical stimulation group and would have potentially required a longer recovery time. However, residual anaesthetic effects in the recovery period are unlikely to explain the absence of implicit memory in this group. Although cognitive functions such as working memory are impaired in the recovery period following propofol anaesthesia, implicit memory remains intact.28

Two other studies have investigated the effects of surgical stimulation on learning. Bethune and colleagues presented words and related phrases (‘Tar—tar makes a mark’) to patients during cardiac surgery and in the immediate postoperative period, or only in the postoperative period.29 Consistent with our findings, patients showed evidence of implicit memory for the stimuli only if they had been presented during surgery as well as postoperatively. MacRae, Thorp, and Millar presented category examples (e.g. ‘peach, grape, melon’) during different phases of the anaesthetic procedure, including periods associated with arousal, for example immediately after the first incision.30 Postoperative memory testing revealed no learning during any phase of the anaesthetic procedure. They did not measure depth of anaesthesia, so it is not possible to judge whether the discrepancy between their results and ours is because of deeper anaesthesia in their study. They used a category generation task to test memory, which requires some degree of semantic processing and therefore may be less sensitive than our word stem completion task. Also, patients in their study received morphine intra-operatively. Morphine reduces the stress response to surgery, decreasing norepinephrine release, thus according to our hypothesis it should have removed the facilitating effect of surgery on memory formation.

One of the difficulties of researching this topic is that the amount of learning during anaesthesia, when it occurs, is rather small, hence there is little leeway for demonstrating effects of experimental manipulations. Although weak, our results combine with the pattern of findings in the literature to provide converging evidence for a role of surgical stimulation in learning during anaesthesia. Future research might maximize the chance of demonstrating learning by maximizing the stress response to surgery and minimizing the use of drugs that suppress this response, such as opioids or local anaesthetics. However, clinical needs limit the extent to which such manipulations can be made.

In conclusion, learning during anaesthesia seems more likely to occur during rather than before surgical stimulation at comparable anaesthetic depth. We speculate that surgical stimulation facilitates learning during anaesthesia, by stimulating amygdala-based memory processes.

Acknowledgements
This work was supported by a Medical Research Council studentship to Catherine Deeprose (née Hanna), Department of Psychology, University of Sheffield, UK.
References

1 Ghoneim MM, Block RL. Learning and consciousness during general anaesthesia. Anesthesiology 1992; 76: 279–305
3 Ghoneim MM, Block RL. Learning and memory during general anesthesia—an update. Anesthesiology 1997; 87: 387–410
4 Merikle PM, Daneman M. Memory for unconsciously perceived events: evidence from anaesthetized patients. Conscious Cogn 1996; 5: 525–41
22 Gooding PA, Mayes AR, Meudell PR. Indirect memory performance is not sensitive to a shift of local context. Eur J Cognit Psychol 1997; 9: 289–312
24 Veselis RA, Reinsel RA, Feshchenko VA, et al. The comparative amnestic effects of midazolam, propofol, thiopental, and fentanyl at equisedative concentrations Anesthesiology 1997; 87: 749–64