Subanesthetic Concentrations of Isoflurane Suppress Learning as Defined by the Category-Example Task

Ben S. Chortkoff, M.D.,* H. L. Bennett, Ph.D.,† E. I. Eger II, M.D.‡

Background: Previously, we found unconscious (implicit) learning in subjects given subanesthetic, but not anesthetic, concentrations of isoflurane. Other investigators, using different learning tasks, have reported implicit learning at anesthetic concentrations. We investigated whether one of these tasks might provide a more sensitive test of implicit learning. In addition, to determine whether suppression of explicit or implicit learning is dose-dependent, we studied one of the tasks at three subanesthetic concentrations.

Methods: We applied a category-example task at 0.15, 0.28, and 0.4 minimum alveolar concentration (MAC) of isoflurane, and a behavior task only at 0.4 MAC. After anesthesia, we determined whether volunteers more frequently listed an example of a category (*e.g.*, flute as an example of musical instrument) presented during anesthesia and/or demonstrated a behavior (touching ear, chin, or knee) suggested to them at 0.4 MAC.

Results: Results from the category task indicated implicit learning only at 0.15 MAC, a concentration that also permitted significant explicit learning. Explicit learning was demonstrated at 0.28 but not at 0.4 MAC (ED_{50} of 0.20 MAC and ED_{95} of 0.4 MAC). Results from the behavior task revealed neither implicit nor explicit learning.

Conclusions: The ED_{50} that suppressed explicit learning in our volunteers equaled that previously reported (0.2 MAC) for implicit learning in volunteers measured using a different task. Combined, these results suggest that less than 0.45 MAC

* Assistant Professor, Anesthesia.

† Associate Professor, Anesthesiology.

‡ Professor and Vice Chairman for Research, Anesthesia.

isoflurane suppresses learning in volunteers. (Key words: Anesthesia, depth: awareness; learning. Anesthetics, volatile: isoflurane. Memory, recall: implicit; explicit.)

LEARNING during anesthesia may be explicit (the patient aware of what was learned) or implicit (the patient unaware of what was learned). Anesthetizing concentrations of volatile anesthetics reliably suppress explicit learning. In contrast, findings for implicit learning vary: Bennett and Davis demonstrated that anesthetized patients told to touch and rub their ear during the postoperative interview did so more often than uninstructed control patients, though no patient explicitly recalled the instruction.[1] Goldmann *et al.* replicated these results.[2] Other investigators applying this behavioral test found additional evidence of implicit learning in anesthetized patients.[3] Failure to replicate these results also has been reported in patients[4] and volunteers.[5]

Block *et al.*[6] showed that subanesthetic concentrations of 30% nitrous oxide suppressed explicit but not implicit learning as measured by a test that provided examples of categories to volunteers breathing nitrous oxide, supporting the observation that implicit learning may be more resistant than explicit learning to suppression by anesthetics. Roorda-Hrdlickova *et al.* demonstrated that repeated presentations of examples of a category (*e.g.*, green as an example of color) to anesthetized patients markedly increased the incidence with which such words were chosen after anesthesia.[7] Jelicic *et al.* reproduced these results[8] with a less remarkable effect. In contrast, when applying this task to anesthetized patients, Block *et al.*[3] found no difference between subjects and controls.

The results obtained by Block *et al.*[6] and Roorda-Hrdlickova *et al.*[7] suggest that the category-example task is a sensitive measure of implicit learning, more sensitive than tasks our group applied previously.[5] Yet, not all investigators have found implicit learning with this task. Perhaps the use of higher anesthetic concen-

trations or different anesthetic combinations by other investigators accounts for the lack of consistent findings with this task. In the present study, we sought a task more sensitive than those we used previously. To define the sensitivity of the category-example task, we applied it to a range of subanesthetic concentrations of isoflurane. To maximize the sensitivity, as well as apply the task in its original form, we also tailored the task to each individual. The latter test used examples generated by each volunteer and allowed a larger list of the volunteer's responses to be examined for evidence of learning. Our expectation was to define the dose-response relationship between end-tidal isoflurane concentration and explicit and implicit learning.

Materials and Methods

We prospectively studied ten healthy male volunteers (aged 22–30 yr) with their informed consent and approval from the Committee on Human Research at the University of California, San Francisco. Volunteers received isoflurane in three increasing subanesthetic concentrations. The capacity to learn auditory information (information presented verbally) was tested at each concentration, using the category-example task. This test assessed whether the capacity to learn was dose-dependent. A second task, the behavior task, was administered only at the highest concentration.

The category-example task and the behavior task measure learning[1,7] by eliciting whether intraanesthetic suggestions influence subsequent verbal and physical behaviors, respectively. For the category-example task, examples of specific categories are presented to anesthetized volunteers, who are tested for their memory of these examples after recovering from anesthesia. We assess whether the examples presented during anesthesia are cited more frequently than control (unpresented) examples. For the behavior task, we instruct the anesthetized volunteer to touch his ear, chin, or knee and compare the preanesthetic and postanesthetic incidence of touching these sites.

Anesthetic Management

All volunteers fasted for the 12 h preceding administration of anesthesia, then swallowed a nonparticulate antacid (Alka-Seltzer Gold, Miles, Elkhart, IN, in 30 ml of water). Each volunteer breathed 100% O_2 through a molded mouthpiece connected to a semiclosed anesthetic circuit; the nares were occluded with a soft clamp. Isoflurane was introduced to maintain an end-tidal concentration of 0.15 minimum alveolar concentration (MAC; 0.19% isoflurane), as measured by mass spectroscopy. To prevent contamination of end-tidal samples with inspired gas, deadspace was augmented at the sampling port (between mouth- and Y-piece) with tubing having an internal volume of 95 ml. To ensure equilibration with the brain, isoflurane was maintained at the desired end-tidal concentration for 15 min before applying any test. Subsequently, volunteers were exposed to 0.275 MAC (0.35% isoflurane), then 0.4 MAC (0.51% isoflurane), each with a 15-min equilibration period. The three anesthetic concentrations were administered in increasing order to replicate the conditions of the study to which we will compare our results.[5] Volunteers were allowed to recover from anesthesia for 1 h before learning was tested.

Study Protocol

Category-Example Test. We applied the category example task in two ways at each anesthetic level: (1) using examples of categories identified in a preliminary study (normative data examples) and (2) modified to increase test sensitivity by incorporating examples provided by each individual volunteer (self-generated examples) and by allowing a larger list of the volunteer's responses to be examined for evidence of learning. These idiosyncratic examples were obtained during a preanesthetic interview.

We conducted a preliminary study to identify a sufficient number of categories and examples to permit administration of the test at more than one anesthetic concentration. We pretested 50 categories in 50 students similar in age, sex, and education to our study population and asked the students to list the first 12 examples of each category as they came to mind. We then selected for study ten categories that elicited examples meeting the following criteria: (1) two examples in the category were cited by at least 20% of all students, and (2) fewer than one-third of all students cited these two examples as their first or second choice. If more than two examples in a single category met these criteria, we selected the two most often cited. That is, the selected examples represented common but not the most popular (*i.e.*, neither first nor second choice) examples of a category. Examples chosen by this method comprise "normative data examples." The resulting categories included those used by Roorda-Hrdlickova *et al.* (colors and fruits) and ranks in the

army, ice cream flavors, musical instruments, fast food restaurants, edible fish, farm animals, presidents of United States, and planets.

To obtain self-generated examples (*i.e.,* those obtained from and applied to a given volunteer), we conducted a preanesthetic interview during which the volunteer provided 8–12 examples for each of five categories randomly selected from the above set of ten. The sequence of responses was recorded. From the resulting list for each category, two items were selected for use during anesthesia. For lists with an odd total (*e.g.,* 9 items), the midpoint and immediate subsequent item were selected (*e.g.,* items 5 and 6). For lists with an even total (*e.g.,* 8 items), the items on either side of the midpoint were selected (*e.g.,* items 4 and 5).

To test learning in each volunteer, two examples from each of six categories (three assigned normative examples and three assigned self-generated examples), randomly chosen, were recorded on an audio-cassette tape tailored to each volunteer. These six data sets were presented to volunteers during anesthesia. The remaining four data sets (two categories using normative and two using self-generated examples) were used as controls. To avoid category selection bias, the categories were randomized and balanced across the study population; that is, we applied ten categories across ten volunteers, providing a 10×10 matrix. Categories were assigned randomly through this matrix to ensure equal representation of each category in each position in the matrix.

The audio-cassette tape was produced by the investigator who conducted the preanesthetic interview (HLB). The tape was divided into three sections, one for each level of anesthesia. Each section provided two examples from a normative category and two from a self-generated category.

At each anesthetic test concentration, one word was spoken every 1.5 s, and each set of four words was repeated 30 times. Examples from each category were alternated (*e.g.,* flute, chicken, drum, sheep).

At each test concentration before and after playing of the tape, consciousness was assessed by the volunteer's response to a command to squeeze the investigator's hand a specified number of times (1–3 times).

The Behavior Test. The behavior task was applied only at 0.4 MAC isoflurane. During the preanesthetic interview, the number of times the volunteer touched his knee, chin, or ear was discreetly noted. Two sites comparable in number and duration of touches were selected; if a volunteer touched all three sites equally,

two were randomly chosen. Three audio-cassette tapes had been recorded previously, each with the instructions to touch one of the three sites during the postanesthetic interview. Only one tape was used for testing, randomly selected from the two tapes (touch sites) appropriate to each volunteer.

To ensure that the category-example and behavioral messages were clearly audible (normal listening tone) and that the appropriate tapes were used with each volunteer, an investigator who did not participate in the postanesthetic interview monitored all tape presentations *via* parallel headphones.

Postanesthetic Interview. One hour after anesthesia, each volunteer was asked to provide 8–12 examples for each of the five self-generated and the five normative data categories. The resulting lists (transcribed by the interviewer) were returned to the volunteer. We instructed the volunteers to examine each list and to circle the two words on each list they believed had been presented during the anesthetic. They were obliged to circle two words on each list, guessing if necessary. Then, they reviewed their choices and stated whether each selection was made because they remembered or recognized having heard the word during the anesthetic or were guessing. We considered learning to be explicit if volunteers either spontaneously recalled or consciously recognized the word as one presented during anesthesia. Implicit learning was defined as follows.

Category-example Data. To score the self-generated examples, each volunteer's postanesthetic list was compared to the preanesthetic list to determine whether the word presented during anesthesia had advanced to a higher position or ranking on the list (a "hit"). A word that maintained position or moved to a lower ranking on the list was considered a "miss." Two examples from each of the two self-generated categories that served as controls (examples not presented during anesthesia) were treated identically to the experimental categories. Hits from the two control categories were summed, divided by two, and subtracted from the sum of the hits for each experimental category at each anesthetic concentration.

For the normative data categories, a hit was defined as the appearance of the normative example within the first three examples in the postanesthetic list. As with the self-generated categories, hits from two control categories were averaged and subtracted from the sum of hits for each experimental category at each anesthetic concentration.

The Behavior Test. During the postanesthetic interview, an observer discreetly counted the number and duration of chin, ear, and knee touches. The number and duration of touches to each of the two possible touch sites identified for each volunteer during the preanesthetic interview was compared to the postanesthetic interview tally. To evaluate whether learning occurred, we determined whether the change in the number and duration of touches to the directed site differed from the change at the control site. This portion of the study was double-blinded for seven of the ten volunteers, and data from only these volunteers were analyzed.

Data Analysis

To calculate the percentage of explicit learning at each anesthetic test concentration, we divided the total number of examples spontaneously recalled or consciously recognized by the total number of examples presented, and multiplied by 100. In addition, we fit a sigmoid equation using BMDPAR–derivative-free nonlinear regression§ to calculate the ED_{50} for suppression of learning, *i.e.,* the concentration of isoflurane that permitted (or suppressed) conscious recall of 50% of the examples presented.

The category-example task was assessed as follows. Data for each method (self-generated *vs.* normative) of applying the category-example task were analyzed independently and compared using a paired *t*-test. As there was no significant difference between the results ($P = 0.22$), the data were combined and linear regression analysis was applied to determine the correlation between concentration of isoflurane and implicit learning. In addition, we compared the number of hits for the presented examples at each anesthetic test concentration with the number of hits for the control examples, using a two-tailed Wilcoxon signed-rank test. $P < 0.05$ was used to identify significant differences.

As an additional assessment of implicit learning, we focused on the volunteer's ability to guess which examples had been presented. We eliminated the examples that the volunteer remembered or recognized. For the remaining examples (the group identified by the volunteer as guessed), we counted the number correctly identified at each anesthetic test concentration (*i.e.,* the example had been presented during anes-
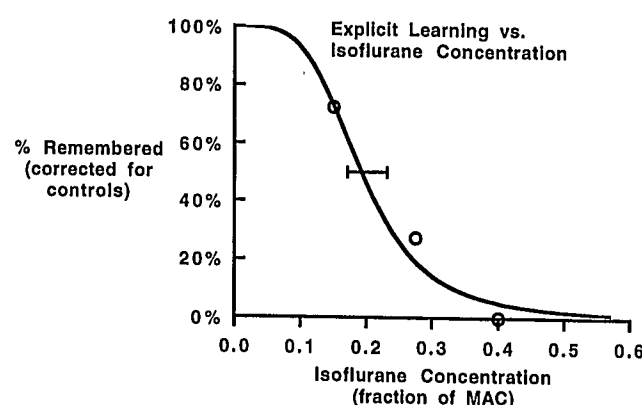
thesia), divided by the total number of guesses at that anesthetic concentration, and multiplied by 100. We compared the results at each anesthetic test concentration by analysis of variance, accepting $P < 0.05$ to identify significant differences.

To evaluate learning in response to the behavioral message, we used chi-square analysis to determine whether change in the number and duration of touches to the directed site differed from change at the control site. $P < 0.05$ was used to identify significant differences.

Results

The highest isoflurane concentration was slightly less than the ED_{50} for response to command, *i.e.,* at 0.4 MAC, we obtained 12 of 20 responses to command. At 0.28 MAC, all volunteers responded properly to the command.

Increasing concentrations of isoflurane decreased the percentage of explicitly remembered words (fig. 1). Seventy-two percent of all examples presented during anesthesia (both self-generated and normative data) were remembered and identified correctly at 0.15 MAC, 28% at 0.28 MAC, and 0% at 0.4 MAC. The resultant ED_{50} for explicit learning calculated by nonlinear regression was 0.20 MAC isoflurane (95% confidence limits 0.17–0.23). The ED_{95}, the concentration of isoflurane that suppressed 95% of explicit learning, was

Fig. 1. The three points represent the percentage of presented words spontaneously recalled or consciously recognized by the ten volunteers at each isoflurane test concentration (measured in minimum alveolar concentration fractions). The curve is derived by fitting the data to a sigmoid equation (see text) to allow interpolation of the ED_{50} value. The error bars represent the 95% confidence interval for ED_{50}.

§ Ralston M: Derivative-Free Nonlinear Regression: BMDP Statistical Software Manual. Edited by Dixon WJ. Berkeley, University of California, 1990, pp 396–422.

≤0.4 MAC (*i.e.,* the amount of explicit learning at 0.4 MAC was indistinguishable from control).

The dose-response relationship for implicit learning was less clear. Analyzing the self-generated or normative data sets independently provided no evidence of implicit learning at any anesthetic test concentration. That is, compared with control responses, examples from either of these two data sets did not advance significantly in rank during the postanesthetic interview. However, analysis of the combined data indicated that implicit learning occurred at 0.15 MAC ($P < 0.05$) but not 0.28 MAC (fig. 2). Linear regression analysis of the combined data revealed a negative slope (slope = -4.8, $r^2 = 0.15$, $P = 0.033$).

Examination of the volunteer's ability to guess which examples had been presented did not demonstrate implicit learning. The incidence of correct guesses (correctly identified examples that were not remembered) did not differ from chance at any anesthetic concentration.

The behavior task did not show implicit learning. There was no difference in the incidence of the suggested behavior before *versus* after anesthesia. Among the seven volunteers for whom data were collected,
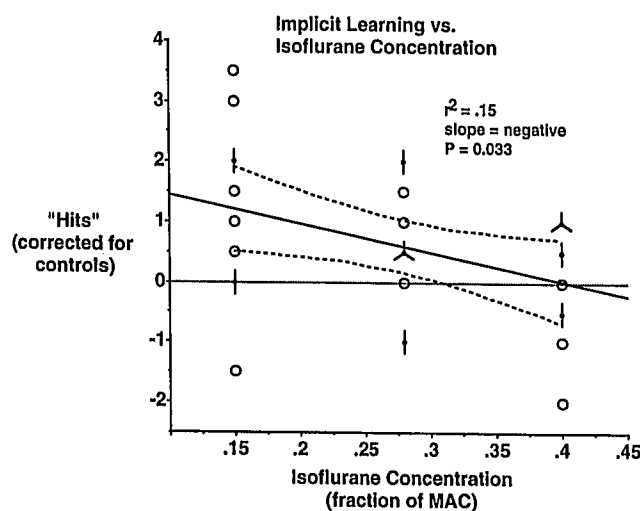

Fig. 2. Hits (see text) minus controls, a measure of implicit learning, for each of the ten volunteers at each anesthetic test concentration. The data shown are obtained by combining the scores of both modifications of the category task. Implicit learning is significantly different from zero only at 0.15 minimum alveolar concentration of isoflurane ($P < 0.05$, two-tailed Wilcoxon signed-rank test). Linear regression analysis applied to the combined results yields a negative slope consistent with a significant dose-dependent suppression of learning ($r^2 = 0.152$, $P = 0.033$). The dotted line represents the 95% confidence intervals. † represents two overlying data points. ⅄ represents three overlying data points.

two demonstrated the suggested behavior more often after than before anesthesia, three less often, and two equivalently. Assessing duration of touch yielded similar results.

Discussion

Our finding of 12 of 20 responses to command at 0.4 MAC supports the previously reported ED_{50} for response to command (MAC-awake) of 0.38 MAC (± 0.03).[5]

Our ED_{50} of 0.20 MAC isoflurane for explicit learning does not differ from the ED_{50} for implicit learning reported previously by Dwyer *et al.*,[5] who used a different memory testing tool, one that possibly combined elements of explicit and implicit learning. Our present task also may have combined elements of explicit and implicit learning by defining recognition as explicit learning. If implicit learning makes some contribution to recognition judgments,[9] then we may have overestimated the extent of explicit learning.

Our finding of explicit rather than implicit learning reported by Dwyer *et al.* may be explained by the time of testing after anesthesia. Dwyer *et al.* tested volunteers 24 h after anesthesia, whereas we tested our subjects at 1 h. Had we waited 24 h, the explicit learning in our volunteers might have decayed or been forgotten, thereby allowing implicit learning to be measured (if the presence of explicit learning hinders the measurement of implicit learning and if implicit learning decays at a slower rate than explicit learning). The similarity in the ED_{50} values obtained by Dwyer *et al.* and our present study implies that implicit learning is suppressed at concentrations nearly identical to those that suppress explicit learning. Both studies found complete suppression of learning (implicit and explicit) in volunteers at less than 0.45 MAC isoflurane.

One goal of our study was to replicate earlier findings[7,8] strongly supporting implicit learning during isoflurane/nitrous oxide anesthesia in surgical patients, using a study design that also defined the dose-response relationship for learning. A second goal was to enhance the sensitivity of the test by tailoring it to each volunteer. Accordingly, we applied the category-example task at three subanesthetic concentrations of isoflurane in two ways: (1) replicating (as closely as possible) the task described by Roorda-Hrdlickova *et al.* and (2) using examples idiosyncratic to each volunteer and allowing a larger list of the volunteer's responses (from three possibilities to four or five possibilities) to be examined for evidence of learning.

Neither goal was achieved. The modification tailoring the test to the individual did not enhance sensitivity. Although we found both implicit and explicit learning at 0.15 MAC isoflurane, we found only explicit learning at 0.28 MAC, and 0.4 MAC suppressed all learning (explicit and implicit). Linear regression analysis of implicit learning (as measured by the category-example task) yielded a negative slope consistent with a dose-dependent suppression of implicit learning, but the concentration of isoflurane causing suppression was far less than that found in investigations revealing learning at anesthetizing concentrations of isoflurane.

The anesthetic protocol for the present study was identical to that used by Dwyer et al.[5] (i.e., delivery of subanesthetic concentrations of agent administered in ascending order). Holding the order of administration constant could have affected our findings if fatigue, practice, or interval between exposure to the information and testing for learning (less than 45 min between the first and final concentrations) alter learning as demonstrated by our tasks. However, Dwyer et al. tested for and found no difference in learning at a defined anesthetic concentration, regardless whether it occurred at the beginning or at the end of a series of other concentrations.

We did not use a placebo control group because we believe that such a group could not be blinded to the absence of treatment (anesthesia). However, such a control group might provide information on the impact of practice and fatigue on the experimental results. Future investigators also might consider that pretreatment testing (baseline testing) would have served different functions to those provided by a control group and might have added to the strength of our results, particularly as a study of the cognitive effects of anesthesia per se.[10] We believe this to be a minor limitation because the results obtained at the lowest anesthetic concentration applied could not have been appreciably improved. Thus, only if a pretreatment testing would have worsened our results could it have modified our conclusions.

Differences between our findings and those of other investigators may arise from differences in anesthetic regimen, study population, or cultural/linguistic sensitivity. First, the more dramatic results showing learning with the category-example task came from studies that included nitrous oxide as a part of the anesthetic technique[7,8]; we administered isoflurane alone. Similarly, other studies using tasks different from the category-example successfully demonstrated learning

during anesthesia that included nitrous oxide.[1,11–13] Exogenous catecholamines improve learning in anesthetized rats.[14] Possibly, the sympathetic stimulating properties of nitrous oxide antagonize the capacity of other components of anesthesia to suppress learning. Consistent with this possibility, Dwyer et al. found that nitrous oxide is less potent than isoflurane (in MAC equivalents) in suppressing learning.[5] Contradicting this possibility, Khilstrom et al. found evidence of implicit learning in patients administered isoflurane without nitrous oxide.[15]

Second, we conducted our study in volunteers. Patients may have increased catecholamines consequent to preoperative anxiety, surgical stimulation, and/or intraoperative stress, and perhaps such increases in catecholamines predispose to learning.

Third, subanesthetic concentrations of isoflurane may be inadequate to reveal implicit learning (i.e., a greater depression may be needed). However, we found a negative dose-response relationship, and our highest subanesthetic concentration (0.4 MAC) inhibited all measurable learning (explicit and implicit). Consequently, we find it difficult to accept that greater concentrations of anesthetic might improve learning. Furthermore, Block et al.,[3] in a study applying the category-example task to fully anesthetized patients, did not demonstrate learning.

Finally, specific cultural/linguistic sensitivities may predispose to learning with the category-example task. That is, either the test population, the categories and the examples selected for testing, or a combination of these factors, may affect sensitivity to learning. For example, there may exist a specific cultural/linguistic sensitivity in Dutch patients to the two categories and/ or the examples chosen for testing by Roorda-Hrdlickova et al. To evaluate whether their categories of color and fruit were intrinsically more sensitive, we compared our results for these categories with the data from each of our other eight categories, finding no difference in learning across categories. The examples for the categories color and fruit were dictated by our preliminary control results and were not the examples used by Roorda-Hrdlickova et al. (i.e., we used black and orange, rather than yellow and green, and banana and grapes rather than banana and pear). Our findings might differ if the examples applied by Roorda-Hrdlickova et al. provided greater sensitivity.

Despite reports demonstrating the effectiveness of the behavior task (including the report authored by one of us (HLB) that introduced the behavior task), this task

also failed to reveal implicit learning in our volunteers. Bennett *et al.* found that anesthetized patients told to touch and rub their ear during the postoperative interview did so more often than uninstructed control patients, though no patient explicitly recalled the instruction.[1] In contrast, we found no effect on the suggested behavior. This difference may be a function of the study design factors already discussed. Our finding of a lack of effect on subsequent behavior in unconscious volunteers agrees with the findings of Dwyer *et al.* in patients.

References

1. Bennett H, Davis H: Nonverbal response to intraoperative conversation. Br J Anaesth 57:174–179, 1985

2. Goldmann L, Shah M, Hebden M: Memory of cardiac anaesthesia: Psychological sequelae in cardiac patients of intra-operative suggestions and operating room conversation. Anaesthesia 47:596–603, 1987

3. Block R, Ghoneim M, Ping SS, Ali M: Human learning during general anaesthesia and surgery. Br J Anaesth 66:170–178, 1991

4. Jansen CK, Bonke B, Klein J, van Dasselaar N, Hop WC: Failure to demonstrate unconscious perception during balanced anaesthesia by postoperative motor response. Acta Anaesthesiol Scand 35:407–410, 1991

5. Dwyer R, Bennett H, Eger E, Heilbron D: Effects of isoflurane and nitrous oxide in subanesthetic concentrations on memory and responsiveness in volunteers. ANESTHESIOLOGY 77:888–898, 1992

6. Block R, Ghoneim M, Pathak D, Kumar V, Hinrichs J: Effects of a subanesthetic concentration of nitrous oxide on overt and covert assessments of memory and associative process. Psychopharmacology 96:324–331, 1988

7. Roorda-Hrdlickova V, Wolters G, Bonke B, Phaf R: Unconscious perception during general anaesthesia demonstrated by an implicit memory task, Awareness and Memory during Anaesthesia. Edited by Bonke B, Fitch W, Millar K. Amsterdam, Swets, 1990, pp 150–155

8. Jelicic M, Bonke B, Appelboom D: Indirect memory for words presented during anaesthesia (letter). Lancet 336:249, 1990

9. Ghoneim M, Block R: Learning and consciousness during general anesthesia. ANESTHESIOLOGY 76:279–305, 1992

10. Ghoneim M, Ali M, Block R: Appraisal of the quality of assessment of memory in anesthesia and psychopharmacology literature. ANESTHESIOLOGY 73:815–820, 1990

11. Levinson B: States of awareness during general anaesthesia. Br J Anaesth 37:544–546, 1965

12. McLintock T, Aitken H, Downie C, Kenny G: Postoperative analgesic requirements in patients exposed to positive intraoperative suggestions. Br Med J 301:788–790, 1990

13. Evans C, Richardson P: Improved recovery and reduced postoperative stay after therapeutic suggestions during general anaesthesia. Lancet 2:491–493, 1988

14. Weinberger N, Gold P, Sternberg D: Epinephrine enables Pavlovian fear conditioning under anesthesia. Science 223:605–607, 1984

15. Kihlstrom JF, Schacter DL, Cork RC, Hurt CA, Behr SE: Implicit and explicit memory following surgical anesthesia. Psychol Sci 1:303–306, 1990