npg

Effect of General Anesthesia in Infancy on Long-Term Recognition Memory in Humans and Rats

Greg Stratmann*[1], Joshua Lee[2], Jeffrey W Sall[1], Bradley H Lee[1], Rehan S Alvi[1], Jennifer Shih[1], Allison M Rowe[1], Tatiana M Ramage[1], Flora L Chang[1], Terri G Alexander[1], David K Lempert[1], Nan Lin[1], Kasey H Siu[1], Sophie A Elphick[1], Alice Wong[1], Caitlin I Schnair[1], Alexander F Vu[1], John T Chan[1], Huizhen Zai[1], Michelle K Wong[1], Amanda M Anthony[1], Kyle C Barbour[1], Dana Ben-Tzur[1], Natalie E Kazarian[1], Joyce YY Lee[1], Jay R Shen[1], Eric Liu[1], Gurbir S Behniwal[1], Cathy R Lammers[3], Zoel Quinones[3], Anuj Aggarwal[1], Elizabeth Cedars[1], Andrew P Yonelinas[4] and Simona Ghetti[5]

[1]Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA, USA; [2]Center for Mind and Brain, University of California, Davis, CA, USA; [3]Department of Anesthesia and Pain Medicine, University of California, Davis, CA, USA; [4]Department of Psychology, University of California, Davis, CA, USA; [5]Department of Psychology, Center for Mind and Brain, University of California, Davis, CA, USA

Anesthesia in infancy impairs performance in recognition memory tasks in mammalian animals, but it is unknown if this occurs in humans. Successful recognition can be based on stimulus familiarity or recollection of event details. Several brain structures involved in recollection are affected by anesthesia-induced neurodegeneration in animals. Therefore, we hypothesized that anesthesia in infancy impairs recollection later in life in humans and rats. Twenty eight children ages 6–11 who had undergone a procedure requiring general anesthesia before age 1 were compared with 28 age- and gender-matched children who had not undergone anesthesia. Recollection and familiarity were assessed in an object recognition memory test using receiver operator characteristic analysis. In addition, IQ and Child Behavior Checklist scores were assessed. In parallel, thirty three 7-day-old rats were randomized to receive anesthesia or sham anesthesia. Over 10 months, recollection and familiarity were assessed using an odor recognition test. We found that anesthetized children had significantly lower recollection scores and were impaired at recollecting associative information compared with controls. Familiarity, IQ, and Child Behavior Checklist scores were not different between groups. In rats, anesthetized subjects had significantly lower recollection scores than controls while familiarity was unaffected. Rats that had undergone tissue injury during anesthesia had similar recollection indices as rats that had been anesthetized without tissue injury. These findings suggest that general anesthesia in infancy impairs recollection later in life in humans and rats. In rats, this effect is independent of underlying disease or tissue injury.

Neuropsychopharmacology (2014) **39**, 2275–2287; doi:10.1038/npp.2014.134; published online 2 July 2014

INTRODUCTION

The behavioral consequences of anesthetic neurotoxicity in the developing brain were first described in rats a decade ago (Jevtovic-Todorovic *et al*, 2003). Exposure to an anesthetic, consisting of nitrous oxide, isoflurane, and midazolam caused widespread neuronal death and impaired learning and memory months later (Jevtovic-Todorovic *et al*, 2003). Almost immediately, scientists, clinicians, and the public began to ask if the same could possibly happen in humans, although this question remains unanswered. Some studies have suggested that there are long-term effects of anesthesia, while others have not (Bartels *et al*, 2009;

Block *et al*, 2012; DiMaggio *et al*, 2009; DiMaggio *et al*, 2011; Hansen *et al*, 2013; Ing *et al*, 2012; Kalkman *et al*, 2009; Satomoto *et al*, 2009; Sprung *et al*, 2012; Sprung *et al*, 2009; Wilder *et al*, 2009). The most important limitation of this body of literature is the inability to dissociate general anesthesia from the underlying disease or surgical procedure that required anesthesia. Also, the degree to which co-existing conditions, known to affect cognition in children, such as low birth weight (Pyhala *et al*, 2011), or prematurity (for review see the study by (Volpe, 2009)), affected the results is unclear. In rats, anesthesia can be administered in the absence of surgical procedures. Further, randomization should distribute rats with preexisting conditions, if any, evenly between anesthesia and control groups. If anesthesia caused a similar cognitive outcome in humans and rats, it could be reasonably assumed that this outcome is indeed due to the anesthetic.

Animal studies have linked anesthesia to impairments on recognition memory tests (Jevtovic-Todorovic *et al*, 2003; Kodama *et al*, 2011; Sanders *et al*, 2009; Satomoto *et al*,

2009; Shih et al, 2012; Stratmann et al, 2009b; Stratmann et al, 2009c; Zhu et al, 2010). Recognition memory judgments can depend on either recollecting specific event details or experiencing familiarity for the event (for review see (Yonelinas et al, (2010))). Recollection is supported by the hippocampus, anterior thalamic nuclei, frontal cortex, retrosplenial/posterior cingulate cortex, and white matter tracts connecting these regions (Yonelinas et al, 2010). Some of these anatomic areas are prime targets of anesthesia-induced neurodegeneration in the developing brain of animals (for review see (Loepke and Soriano, (2008))). Familiarity, by contrast, is supported by anterior medial temporal cortices including the perirhinal cortex (for review see (Eichenbaum et al, (2007))), a region that is commonly not affected by anesthesia-induced neurodegeneration (for review see (Loepke and Soriano, (2008))). We tested the hypothesis that anesthesia in infancy impairs recognition memory, and that these effects would arise because of a reduction in recollection rather than in familiarity. Recollection and familiarity can be tested in humans and recollection- and familiarity-like memory can be tested in rats using similar cognitive assays (Eichenbaum et al, 2007; Fortin et al, 2004; Sauvage et al, 2008; Yonelinas, 1997). Here we show that general anesthesia in infancy impairs recognition memory performance due to impaired recollection but not familiarity in both humans and rats.

MATERIALS AND METHODS

Human Study

With IRB approval, anesthesia billing databases for the year 2004 for the University of California San Francisco and University of California Davis were queried for eligible participants. Information included name, date of birth, gender, American Society of Anesthesiologists (ASA) physical status, body weight, date of surgery, type of surgery, diagnosis, anesthesia start time, time of entry into the operating room, and time of exit from the operating room. Children were eligible if they received a general anesthetic before age 2, were 6–11 years of age at the time of testing, and generally healthy (ASA physical status 1 or 2 without diagnoses or operations that might be associated with cognitive impairment, such as neurosurgical procedures or congenital heart disease). The lower age limit of 6 years was chosen because younger children do not reliably perform the recognition memory tests. The upper age limit of 11 years was chosen because older children were more likely to have received anesthetic agents that are no longer commonly used such as thiopental, halothane, or enflurane.

The chart was reviewed for further inclusion criteria such as anesthetic dose of greater than 120 minimum alveolar concentration (MAC) $\times$ min, induction of general anesthesia with either a volatile agent ($+/-$ nitrous oxide) or propofol, and maintenance with either a volatile agent or combination of a volatile agent and nitrous oxide. MAC, or minimum alveolar concentration, is a measure of anesthetic potency and is the concentration of inhaled anesthetic required to prevent movement in 50% of subjects in response to painful stimulus. We chose 120 MAC $\times$ min, because this anesthetic dose caused robust neurodegeneration in the developing rodent brain, whereas 60 MAC $\times$ min

did not (Stratmann et al, 2009a). Due to clerical error, two patients were enrolled whose anesthetic doses were less than 120 MAC $\times$ min (40 and 59 min). Their data were included because the results and conclusion were not changed.

Parents of eligible participants were contacted to evaluate exclusion criteria and to obtain consent. Exclusion criteria, identified from phone interview and chart review, included ASA physical status 3 or greater, impairments of attention or learning, preexisting conditions possibly associated with neurocognitive impairment (eg, a history of disease or trauma to the central nervous system), cancer, premature delivery (<37 weeks gestation), low birth weight (<5 lb), and known genetic syndromes. Further exclusion criteria included potentially confounding intraoperative physiologic factors persisting for more than 5 min—hypotension (systolic, diastolic, or mean arterial pressure $<30\%$ from baseline), bradycardia (heart rate $<30\%$ from baseline), hypoxemia (oxygen saturation reading $<93\%$), hypercarbia (PaCO2 >60 mm Hg), or dysthermia (temperature <35 °C or >35 °C). Finally, patients were excluded when meaningful participation in the study was unlikely (eg, color blindness or inability to speak English).

The control sample was recruited from a registry of parents who previously expressed interest in having their children participate in research. Children were selected to match age and gender. Potential control participants were excluded if the child received general anesthesia, met any exclusion criteria described for the anesthesia group, or were not of the targeted age range.

Recognition memory was tested using two separate tasks—color task and spatial task. Afterward, the Wechsler Abbreviated Scale of Intelligence was administered that provided verbal, performance, and full IQ scores. While participants were tested, their parents provided demographic and disease-related information on a questionnaire, which was consolidated with information from medical records. Parents also filled out the Child Behavior Checklist, a standardized parent-reported questionnaire on which a child is rated for behavioral adjustment.

Recognition Memory Measures

Two separate tests were used to assess recognition memory as previously described (Ghetti et al, 2010). The examiner was blind to exposure status, and interactions with test subjects were standardized and scripted. The test results provide measures of recollection index, familiarity index, and source memory. Recollection and familiarity were determined using analysis of receiver operator characteristics (ROC) of recognition memory data (Ghetti and Angelini, 2008; Yonelinas, 1994; Yonelinas et al, 2002).

For each task, eighty black ink drawings on a white, square background were presented sequentially on a computer screen. In the color task, the drawing was presented with one of four colored borders (red, blue, yellow, or green). In the spatial task, the drawing was presented in one of four quadrants of the screen. Subjects were instructed to remember the drawing and the border color or spatial location associated with it. The four conditions of color or location were used with equal frequency and in random order, and the sequence of the two tests (color and spatial) was counterbalanced across participants.

Items were chosen from a set of 320 unambiguous line drawings that were validated with child participants for familiarity, visual complexity, and name agreement (Cycowicz et al, 1997).

Five minutes after viewing all 80 drawings, participants were given a self-paced recognition test that included 80 original ('old') drawings and 80 previously unseen ('new') drawings. Images were presented in the center of the screen without a colored border. Participants first determined whether they had seen the drawing before. Next, they rated the confidence in their recognition response (not at all, a little, or very confident). If the drawing was recognized, then participants were further asked to recall the color of the border or the quadrant of the screen in which it had appeared previously.

Recollection and familiarity were determined using ROC curves that were generated by plotting the rate of correct 'old' judgment (hits) against the rate of incorrect 'old' judgments (false alarms) as a function of response confidence, ranging from 1 ('very sure this is new') to 6 ('very sure this is old') (Ghetti and Angelini, 2008). A curve was then fitted to these data points using a least squares model. Based on the dual-process model (Yonelinas, 1994), at each confidence level, hit rates are described by the equation: $P(\text{Ing et al}) = R + (1 - R)\ \Phi\ (d'/2 - ci)$, reflecting the independent contribution of the single threshold process of recollection (R) and the continuous process of familiarity ($\Phi\ (d'/2 - c)$ (ie, the proportion of old item distribution that exceeds the response criterion c), whereas false alarm rates are expected to be described by the equation: $P(\text{false alarm}) = \Phi(-d'/2 - c)$, reflecting only the contribution of familiarity. Given that recollection and familiarity are assumed to remain constant along the ROC curve while only the criterion varies, the set of equations above (ie, two for each confidence level) can be solved to derive R and d'.

In addition, source recollection, or the ability to recollect specific details, was evaluated by measuring the proportion of trials in which individuals correctly remembered the associated color or location of images. Source recollection was calculated by dividing the number of correctly identified sources (color or spatial location) by the number of 'old' judgments.

Rat Study

Anesthesia. With IRB approval, cross-fostered 7-day-old male Sprague–Dawley rats ($n = 42$) were randomized to receive either sevoflurane anesthesia for 4 h at 1 MAC ($n = 25$) as previously described (Shih et al, 2012) or no anesthesia ($n = 17$). As MAC is not stable in immature rodents (Kodama et al, 2011; Stratmann et al, 2009b), roughly half of the sevoflurane-treated animals were tail clamped during anesthesia to titrate the anesthetic concentration to the clinically relevant endpoint of MAC. Tail clamping also produces tissue inflammation and scarring similar to surgery. The anesthesia protocol has been published in detail elsewhere (Shih et al, 2012; Stratmann et al, 2009b), and the dose and duration were chosen based on prior experiments that resulted in neuronal death and an identifiable cognitive deficit.

Odor recognition task training. Odor recognition testing was performed as before by experimenters blind to group assignment (Fortin et al, 2004; Robitsek et al, 2008; Sauvage et al, 2008). Rats were trained to perform a distinct behavioral sequence for the following: (1) an 'old' odor that had been previously presented on the same day, or (2) a 'new' odor that had not been previously experienced. The rats' responses were biased across five conditions yielding discrete false alarm rates. Biasing conditions were based on cup height, food reward for 'old' judgment, and food reward for 'new' judgment as shown in Table 1. Two unscored odors (one old, one new) were presented initially after the delay to allow the rat to experience the biasing conditions, followed by 20 scored odors in random order (10 old, 10 new).

For each bias level, a 'hit' rate and 'false alarm' rate were obtained. Once false alarm rates varied by less than 20% over five consecutive sessions, at least five responses per bias level were scored and averaged. As the duration over which rats learned the rules of the task varied (between 3 and 9 months), only the last 5 weeks of testing was scored. Thus, all animals were of similar ages, between 9–10 months, at the time of data collection. Recollection and familiarity were determined using analysis of ROC of recognition memory data (Fortin et al, 2004; Robitsek et al, 2008; Sauvage et al, 2008).

Odor dilution test. We investigated whether anesthesia affects rodents' sense of smell by assessing odor identification in stepwise dilutions of the original scent (0.5 g of scented sand per 100 g of clean, unscented sand), starting with $1:10^5$ dilution down to a final dilution of $1:10^9$. This allowed us to compare control and anesthetized rats' ability to identify scents at dilution levels lower than those used in the recognition memory testing. If anesthesia impairs the sense of smell, then anesthetized rats would have lower accuracy of odor identification relative to control rats at the same scent dilutions. Following the dilution series, the scent was reintroduced at the highest dilution associated with normal performance ($>80\%$ accuracy). Eight randomly selected rats (five control, three anesthetized) were tested by experimenters blinded to group assignment.

Table I Strategy for Biasing Responses Toward 'old' or 'new' Judgments

Bias level	5	4	3	2	1
Cup height	Short	Short	Regular	Regular	Tall
Old reward	1/2	1/4	1/4	1/4	1/4
New reward	1/2	1/2	1	2	3

The numbers in the bottom two rows represent the cereal ring(s) awarded for correct 'old' or 'new' judgments. Each bias level has a distinct combination of cup height and food reward ratio, resulting in a distinct false alarm and hit rate (ie, one of five data points used to generate the ROC curves). Liberal criteria such as bias level 1 (digging is more difficult, thus biasing toward 'old' judgments) result in high false alarm and high hit rates, whereas conservative criteria, such as bias level 5 (digging is easier, thus biasing towards 'new' judgments) result in low false alarm and hit rates.

Statistical Analysis

Sample size determination. For the human study, it was determined that a total of 52 patients would be required to detect a difference in means in the spatial recollection index of 0.16 with 80% power at a significance level of 0.05, if the s.d. of the recollection index is 0.20. SAS (version 9.2, SAS, Cary, NC) Proc Power for two sample *t*-tests based on mean differences was used to calculate the sample size. This effect size was derived from pilot data using six patients and six matched controls. We planned to enroll 28 patients per group to create a small margin of error. SAS (version 9.2, SAS) Proc Power for two sample *t*-tests based on mean differences was used to calculate the sample size.

In rats, the pilot experiment to determine anesthetic effect size in five animals per group revealed a statistically significant difference in recollection between groups, making sample size calculation redundant. To validate these results, the experiment was repeated twice using similar sample sizes as in the pilot experiment.

Data fulfilling parametric assumptions were expressed as means and 95% confidence intervals of the means. Nonparametric data were expressed as medians and inter-quartile ranges or medians and ranges (IQ data). Group differences in the primary outcome measures—recollection and familiarity indices—were tested using Student's *t*-test in humans and Mann–Whitney *U*-test in rats (data in the sevoflurane group did not fulfill parametric assumptions because half of all animals had recollection index of zero). Analyses were conducted using SPSS (version 20, IBM, Somers, NY) and Prism 5 for Mac OS X (GraphPad, San Diego, CA). A *P*-value of 0.05 was considered statistically significant.

The effect of demographic variables (age at testing, gender, socioeconomic status) and the effects of clinically relevant variables (anesthetic duration, age at first exposure, anesthetic agent, cumulative dose expressed as MAC × hours) were explored using multivariate analysis of covariance (ANCOVA) and tested using Bonferroni correction to test for the interactive effects below. The effect of test type (spatial *vs* color) on source recollection was assessed with a 2-way repeated measures analysis of variance (ANOVA). Further exploratory analyses of clinically relevant variables were conducted, which included both linear regression and Spearman's rank order coefficients correlation analyses. In addition, Spearman's correlation analyses were conducted to evaluate whether familiarity or recollection outcomes in either task were influenced by the total dose (MAC min) of anesthetic received. Analyses were conducted using SPSS (version 20, IBM). A *P*-value of 0.05 was considered statistically significant.

RESULTS

Human Data

Patient enrollment is summarized in Figure 1. Eight-hundred thirty four records were screened for anesthetic exposure before age 2, and 350 potential participants met enrollment criteria. Of these, 58 families had valid contact information and were contacted. Thirty of these families agreed to participate (52%). Over a period of 20 months (May 2011–January 2013), a total of 21 boys and 9 girls were enrolled and tested. Two enrolled patients were excluded from analysis (inability to comply with test instructions). By chance, all 28 included patients received their first anesthetic before age one. Cases ($n = 28$) and controls ($n = 28$) were well-matched for age and gender, IQ, and CBCL Total Problems scores. One exception is that patient families earned significantly more income than control families (Table 2). The mean age at first exposure was 6 months (95% CI: 5.3 to 7.6 months, range: 2.2 to 11.2 months) with boys being exposed earlier than girls (5.7 *vs* 8.3 months respectively, 95%CI of difference between means: 0.5 to 4.8 months, $P = 0.02$)

A summary of the patient characteristics, including details of the anesthetic exposures and surgeries performed, are listed in Table 3. The median anesthetic dose was 203 MAC × min (interquartile range 155–325). The mean
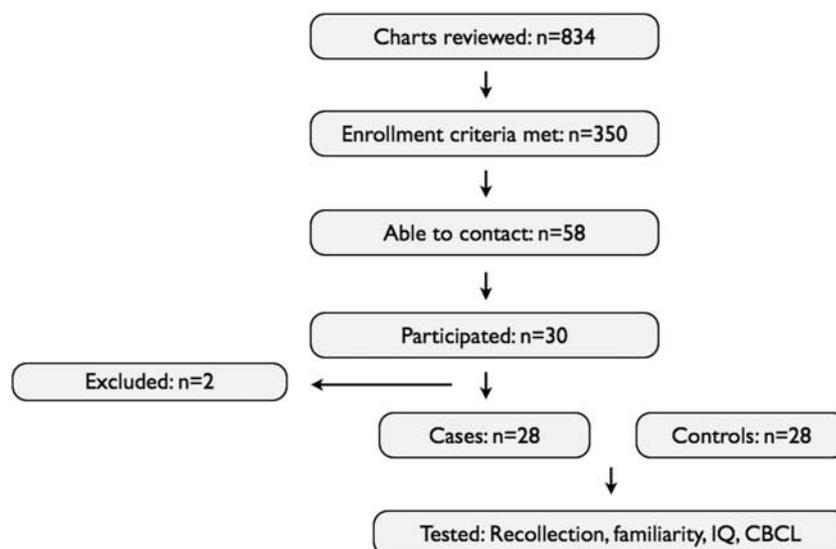
Figure 1 Flow chart summarizing participant enrollment. CBCL, Child Behavior Checklist.

Table 2 Demographic Characteristics of Sample

Characterisitic	Anestheisa (N = 8)	Control (N = 28)	Total (N = 56)	P-value
Female participants, No. (%)	8 (30)	11 (39)	19 (34)	0.4
Age at testing, mean (SD), year	9.4 (1.6)	9.13 (1.6)	9.27 (1.6)	0.53
Family income, mean (SD), $1000	85.2 (25.0)	69.6 (27.3)	77.4 (27.1)	0.03
Full-scale IQ, mean score (SD)	111.7 (16.4)	110.2 (16.2)	111.0 (16.1)	0.73
Verbal IQ, mean score (SD)	112.8 (17.6)	109.6 (13.9)	111.2 (15.8)	0.45
Performance IQ, mean score (SD)	108.4 (15.4)	112.5 (16.2)	110.4 (15.8)	0.34
CBCL total problems (SD)	46.5 (8.0)	46.1 (8.0)	46.3 (7.9)	0.87

One-way ANOVA was used to test for differences between anesthesia and control groups on each demographic variable. Only family income differed between groups $F(1,54) = 4.94$, $P = 0.03$. The families of anesthetized children earned significantly more annual income than those of the control group.

anesthetic duration, defined by the time of entry into the operating room until time of exit from the operating room, was 148 min (95% CI: 119–178). Eighteen patients received a single anesthetic of a median dose of 238 MAC × min (interquartile range: 179–355). Ten patients received multiple anesthetics before testing with an initial median anesthetic dose of 169 MAC × min (interquartile range: 148–203). There was no difference in IQ scores between anesthetized children (median 112, range: 66–146) and controls (median 110, range: 66–133, rank sum difference 4, $P = 0.98$, Mann–Whitney U-test).

The anesthetics consisted of a mixture of nitrous oxide and a volatile anesthetic (all patients were exposed to nitrous oxide). Propofol was used in three patients but never as the sole agent. The most frequently used volatile anesthetic was sevoflurane ($n = 26$) followed by isoflurane ($n = 13$) and halothane ($n = 7$). Eleven patients were exposed to only sevoflurane and one patient was exposed to only halothane. Most patients were exposed to a combination of volatile agents, including sevoflurane and isoflurane ($n = 10$), sevoflurane and halothane ($n = 3$), halothane and isoflurane ($n = 1$), or all three agents ($n = 2$).

Recognition memory

Item recognition memory, ROC curves were obtained for participants, and estimates of recollection index (R, the y intercept of the ROC curve) and familiarity index (d', degree to which the ROC curve bows upwards) were determined for anesthetized and non-anesthetized children (Figure 2a and b). The recollection index of anesthetized children was significantly lower than that of controls in both the color task (anesthesia mean 0.34 vs control mean 0.47, $P = 0.02$, t-test) and spatial task (anesthesia mean 0.38 vs control mean 0.49, $P = 0.04$, t-test). This difference retained statistical significance when family income, the only variable that distinguished the groups, was included in a covariate 2 (patients vs control) × 2 (color recollection index vs spatial recollection index) mixed measures ANCOVA (F(1,53) = 8.37, $P = 0.006$). The familiarity index, however, was not different between groups in either the color task (anesthesia mean 0.75 vs control mean 0.77, $P = 0.88$, t-test) or spatial task (anesthesia mean 0.84 vs control mean 0.78, $P = 0.77$, t-test).

To further evaluate the nature of memory impairments, separate analyses were conducted on the hit rates (eg, the probability of correctly recognizing an old item as old) and false alarm rates (eg, the probability of incorrectly recognizing a new item as old) (Figure 3). Hit rates alone, although, cannot inform the nature of the deficits because they reflect the contribution of both recollection and familiarity processes. For the color task, the hit rate was significantly lower in the anesthesia group (anesthesia median 54.4% vs control median 71.3%, $P = 0.002$, Mann–Whitney U-test, Figure 3a). The hit rate in the spatial task was not significantly different between groups (anesthesia median 63.8% vs control median 73.8%, $P = 0.15$, Mann–Whitney U-test, Figure 3b). False alarm rates were not different between groups in either task (Figure 3c, d).

Source recollection in the color task was significantly impaired in the anesthesia group (correct source recollection in color task: anesthesia mean 18.2% vs control mean 22.7%, $P = 0.03$, t-test, Figure 3e). In the spatial task, the source recollection reduction in the anesthetized group approached, but did not reach, statistical significance (correct source recollection in spatial task: anesthesia mean 31% vs control mean 39.2%, $P = 0.08$, t-test, Figure 3f). A direct comparison indicated that the location memory deficit was not significantly smaller than the color memory deficit (treatment $F_{(1,54)} = 5.96$, $P = 0.02$, test $F_{(1,54)} = 35.84$, $P < 0.0001$, interaction $F_{(1,54)} = 0.54$, $P = 0.47$, two-way repeated measures ANOVA). Overall, these results show that anesthetized children exhibit reduced recognition memory performance.

Exploratory Analyses

We collected data of clinical relevance that may help guide future investigations. To examine the effects of current age and sex, we conducted two separate 2 (participant group: patient vs control) × 2 (sex: male vs female) ANCOVA, entering age at testing as a covariate with recollection in the color and spatial tasks. In the color task, there was a significant main effect of participant group, $F(1,51) = 4.92$, $P = 0.03$, while the main or interactive effects of sex and age did not achieve statistical significance, $Fs(1,51) < 3.79$, $Ps > 0.06$. The spatial task, meanwhile, showed a significant interaction between sex and participant group, $F(1,51) = 4.60$, $P = 0.04$. Males in the patient group performed significantly worse than males in the control group (mean adjusted for covariate: patients − 0.34, 95% CI: 0.26–0.43 vs control − 0.54, 95% CI: 0.44–0.63); however, this was not

Table 3 Patient Characteristics Including Details about Anesthetic Exposures and Surgeries as well as Primary Outcomes for each Patient

Diagnosis	Type of surgery	Age	Sex	Age at 1st exposure	Kg at 1st exposure	Anesthetics before age 2	Exposed after age 2	Anesthetic duration	N20	Halo-thane	Iso-flurane	Sevo-flurane	Sum	Propofol (mg)	Fentanyl (mcg)	IQ	Color R	Color F	Spatial R	Spatial F
									MAC × min					IV agents, total			Primary outcomes			
Hip dysplasias, chronic skin ulcer	Bilateral open reduction	6.09	F	0.48	7	3	Yes	572	191.00	481.00	218.00	186.50	1,077	0	20	98.00	0.49	1.97	0.70	0.75
Ureterocele	Cystoscopy/circumcision	7.29	M	0.27	6	2	No	257	57.50	0.00	66.30	168.00	292	0	0	125.00	0.01	1.52	0.35	0.54
Cleft lip	Cleft lip repair	9.88	F	0.25	6	1	No	120	35.00	0.00	63.60	37.50	136	0	15	107.00	0.12	0.51	0.62	1.87
Ureterocele	Ureteral endopuncture	6.81	M	0.19	6	2	No	199	99.50	0.00	150.00	105.00	355	0	10	101.00	0.03	0.33	0.12	0.00
Hypospadias	Hypospadias repair	10.06	M	0.58	10	1	No	133	81.00	157.00	0.00	0.00	238	20	10	66.00	0.00	0.00	0.46	0.02
Right eyelid hemangioma	Right eyelid hemangioma injection	10.00	F	0.93	8	1	No	15	10.00	14.00	0.00	16.00	40	0	0	108.00	0.32	0.74	0.73	1.69
Pulmonary sequestration	Thoracoscopy, thoracotomy, resection of pulmonary sequestration	6.90	M	0.73	9	1	No	195	32.40	0.00	0.00	297.00	329	0	0	117.00	0.38	0.00	0.14	0.31
Polydactyly	Excision of extra member	10.22	M	0.38	7	1	No	194	100.00	0.00	0.00	159.00	259	0	35	109.00	0.38	0.43	0.27	0.66
Hypospadias	Hypospadias repair	10.49	M	0.55	7	2	No	235	110.00	130.00	0.00	115.00	355	0	8	136.00	0.41	0.60	0.35	1.74
Tethered cord	Tethered cord release	6.90	F	0.61	9	5	Yes	613	115.00	0.00	58.30	329.00	502	50	50	120.00	0.27	0.35	0.23	1.59
Sebaceous cyst	Excision	11.53	F	0.73	8	1	No	97	40.00	56.00	76.00	0.00	172	0	0	100.00	0.44	0.81	0.51	0.70
Sternal notch mass	Excision	9.29	M	0.53	8	1	No	45	18.00	0.00	71.00	48.00	137	0	0	110.00	0.45	0.00	0.57	1.13
Anal fistula	Anal fistulectomy	9.88	M	0.24	5	2	No	77	28.60	0.00	40.80	100.00	169	0	5	116.00	0.27	1.19	0.19	1.59
Testicular torsion	Left orchidopexy	10.05	M	0.18	6	1	No	117	12.00	0.00	0.00	47.00	59	0	0	107.00	0.49	0.44	0.12	0.00
Umbilical hernia and periauricular lesion	Repair and excision	10.46	M	0.27	6	1	No	87	56.00	0.00	56.00	38.00	150	0	5	124.00	0.49	1.58	0.38	0.86
Left complex bifid thumb (duplication)	Amputation	9.66	F	0.54	7	1	No	113	38.00	0.00	39.00	123.00	200	0	0	116.00	0.06	0.00	0.50	0.12
Vesicourethral reflux	Bilateral urethral reimplantation	11.81	F	0.93	10	1	No	166	26.00	0.00	0.00	207.00	233	0	5	96.00	0.84	0.00	0.31	1.33
Hypospadias	Hypospadia repair	10.75	M	0.90	12	1	No	128	58.00	0.00	0.00	90.00	148	0	0	113.00	0.23	0.56	0.34	0.85
Hypospadias	Hypospadia repair	9.78	M	0.47	6	1	No	110	50.50	0.00	0.00	128.00	179	0	5	118.00	0.75	1.89	0.42	0.88
Bilateral inguinal hernias	Bilateral inguinal hernia repair	6.41	M	0.33	8	1	No	156	90.00	0.00	0.00	115.00	205	0	10	86.00	0.16	0.56	0.00	0.00
Vascular malformation left hand	Left hand sclerosis	7.75	M	0.91	10	2	Yes	385	13.00	0.00	0.00	462.60	476	40	25	146.00	0.23	0.92	0.69	1.19
Hypospadia	Hypospadias repair	9.67	M	0.61	9	1	No	125	57.00	0.00	0.00	122.00	179	0	0	100.00	0.21	0.33	0.27	0.64
Vesicourethral reflux	Vesicoureteral reimplantation, cystogram	10.31	F	0.86	10	1	No	179	101.00	0.00	0.00	97.00	198	0	0	96.00	0.33	0.49	0.21	0.00
Giant hairy nevus	Resect and skin graft from buttocks	11.25	F	0.48	7	1	No	316	151.00	0.00	206.00	24.00	381	0	70	115.00	0.43	0.94	0.48	1.56
Clubfoot	Achilles tenotomy; tendon lengthening; leg cast	10.77	M	0.31	7	4	No	245	123.00	79.00	24.00	84.00	310	0	0	128.00	0.60	2.23	0.53	1.32
Hypospadias	Complex hypospadias repair	9.39	M	0.60	10	1	No	138	63.00	114.00	0.00	24.00	201	0	15	137.00	0.33	0.92	0.27	0.15
Bilateral inguinal hernias	Bilateral inguinal hernia repair	11.23	M	0.67	10	1	No	123	80.00	0.00	49.00	18.00	147	0	25	108.00	0.49	0.97	0.52	1.02
Hypospadias	Complex hypospadias repair	8.77	M	0.48	8	1	No	138	68.00	0.00	0.00	139.00	207	0	5	125.00	0.20	0.73	0.20	0.91

Abbreviations: F, familiarity index; kg, weight in kilograms; MAC, minimum alveolar concentration; N20, nitrous oxide; R, recollection index.
Age refers to the age at testing, calculated as years + (days/365) and rounded to 1 decimal point. Age at 1st exposure is likewise expressed as (a fraction of) years and rounded to 2 decimal points.
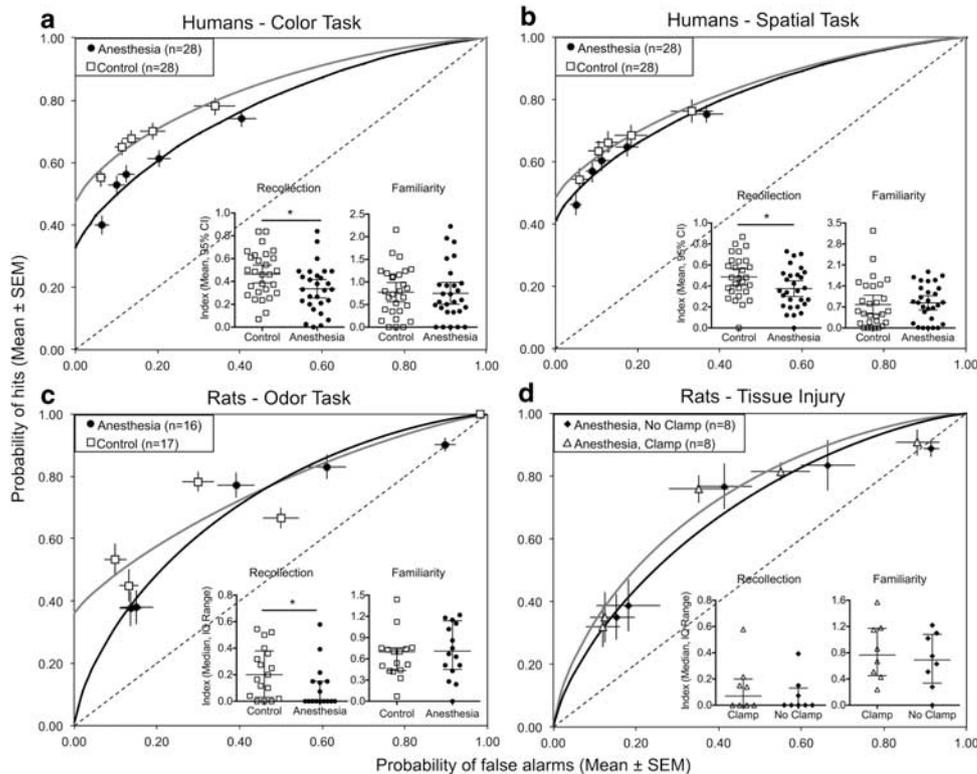
Figure 2 ROC curves and derived estimates of recollection and familiarity of a color recognition memory task in humans (a), a spatial recognition memory task in humans (b) and an odor recognition memory task in rats (c). The composite ROC curve for each human group is fit to the six (human) and five (rat) data points representing the means of all 'hit' and 'false alarm' rates for each of the confidence (humans) and bias (rats) categories. The *y* intercept of this curve is the recollection index *R*, the degree to which the curve bows upwards (asymmetry) is the familiarity index *F*. ROC curve error bars are SE. The diagonal straight line represents chance performance. Recollection and familiarity data are medians and interquartile ranges. Group ROC curves and recollection/familiarity estimates of anesthesia with and without tissue injury caused by tail clamping during anesthesia (d), showing that tissue injury does not change the effect size of anesthesia-induced impairment in recollection. Tissue injury was caused by tail clamping 50% of rats under anesthesia to determine anesthetic depth. SEM, standard error of the mean.

the case in females (mean adjusted for covariate: patients − 0.45, 95% CI: 0.31–0.59 *vs* controls − 0.41, 95% CI: 0.30–0.53). No other significant main or interactive effects emerged, Fs(1,51) < 2.17, Ps > 0.15.

To explore the effects of age at first exposure and anesthetic duration, we created two median split variables to identify: (1) patients above or below the median age (6.4 months) at first exposure and (2) patients whose exposure was above or below the median duration (132 minutes). We then conducted a 2 (sex: male *vs* female) × 2 (duration: short *vs* long) × 2 (exposure timing: early *vs* late) ANCOVA, entering age at testing as a covariate and recollection in each task as repeated measures. This analysis confirmed the main effect of sex on recollection and also revealed a significant interaction between sex and age at first exposure, F(1,20) = 7.63, *P* = 0.01. In females, those with early exposures exhibited better overall recollection than those with late exposures (mean adjusted for covariate: early—0.78, 95% CI: 0.51–0.97 *vs* late—0.37, 95% CI: 0.27–0.47), and females had higher recollection than males regardless of when they received anesthesia. There was also a significant interaction between age at first exposure and duration of exposure, F(1,20) = 8.86, *P* = 0.01. Children with early, short exposures had higher recollection than either those with early, long exposures (mean adjusted for covariate: short—0.66, 95% CI: 0.46–0.85 *vs* long—0.39,

95% CI: 0.27–0.50) or those with late exposures regardless of duration (mean adjusted for covariate: short—0.31, 95% CI: 0.20–0.41; long—0.37, 95% CI: 0.28–0.46).

We also computed nonparametric Spearman's rank order correlation coefficients between anesthetic status, age at testing, CBCL total problems, gender, full-scale IQ, family income, and recollection and familiarity estimates for spatial and color tasks (Table 4). Notably, there was a significant association between family income (higher in patients) and IQ but not with recognition memory measures. Thus, although income positively correlated with IQ, it did not extend to memory performance and did not alter the relationship between anesthetic exposure and recollection. In addition, separate Spearman's correlation analyses were conducted but did not identify a significant correlation between anesthetic dose (MAC min) and outcomes in the memory tasks: color task recollection (*r* = − 0.098, *P* = 0.62), color task familiarity (*r* = 0.040, *P* = 0.84), spatial task recollection (*r* = − 0.095, *P* = 0.63), or spatial task familiarity (*r* = − 0.022, *P* = 0.91).

Given that the only difference between patients and controls was family income, we further conducted four separate linear regressions in which anesthesia status, age at testing, CBCL score, gender, IQ, and family income were used to predict color recollection, spatial recollection, color familiarity, or spatial familiarity. Results of these analyses
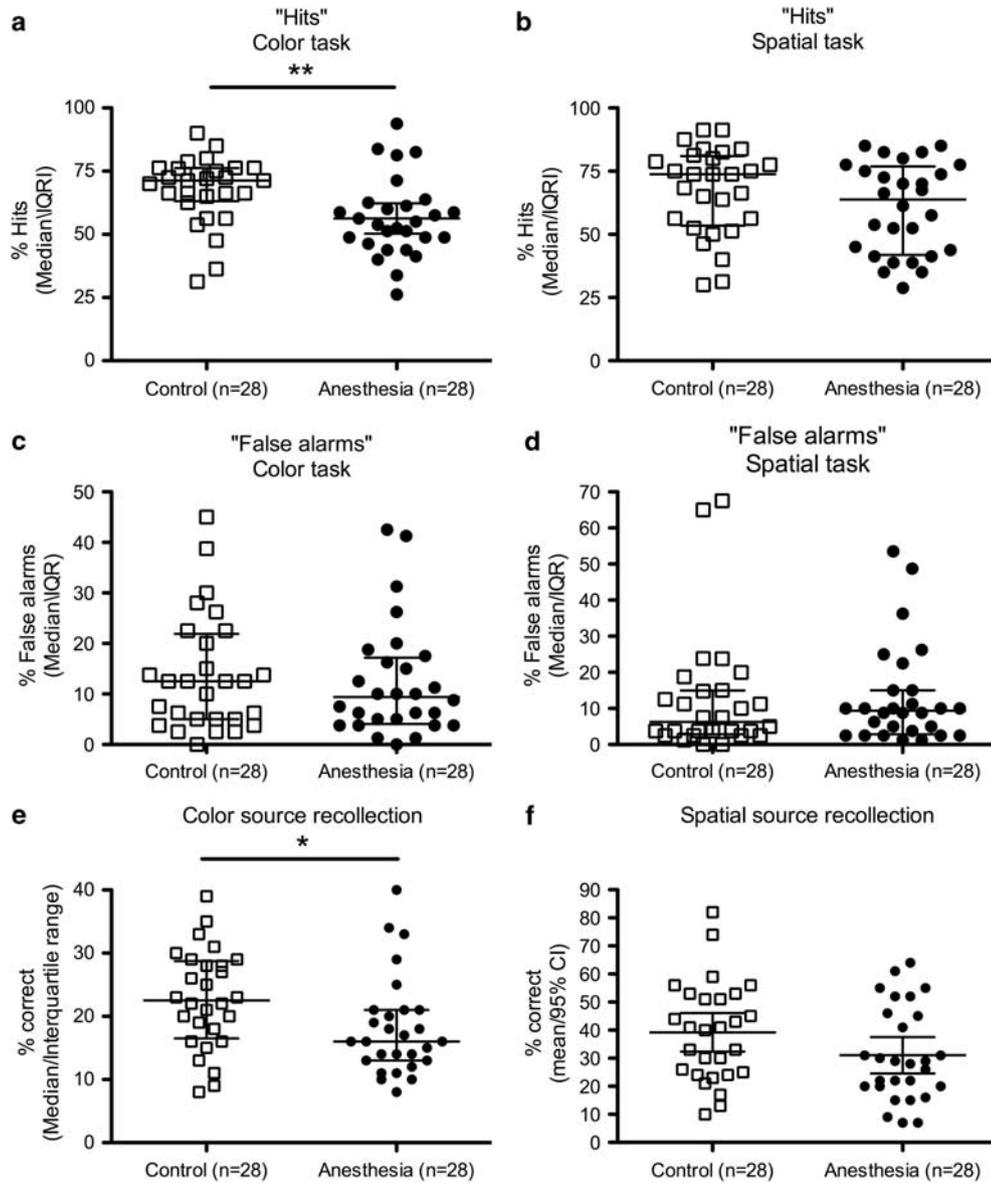
Figure 3 Human memory testing results. (a–d) The hit rates (eg, the probability of correctly recognizing an old item as old) and false alarm rates (eg, the probability of incorrectly recognizing a new item as old) are shown for each task. For the color task, the hit rate was significantly lower in the anesthesia group. False alarm rates were not different between groups in either task. (e and f) Source recollection performance in the color task was significantly impaired in the anesthesia group.

are presented in Table 5. Only participant status and age predicted recollection indices, whereas the other variables did not significantly predict either recollection or familiarity indices. These exploratory results are consistent with our primary analyses in which history of anesthetic exposure was associated with lower performance on recollection but not familiarity parameters.

Rat Data

Recognition memory was tested in rats using odor as a memory trace. Hits and false alarms were recorded for each of five response bias levels, reoccurring in random order once a week. One sevoflurane-treated rat had to be excluded for inability to learn the rules of the task. Once stable,

performance at each bias level did not decrease over time. Data averaged over the last 5 weeks of testing were included in the analysis.

The ROC curves, as well as recollection and familiarity estimates, for rats anesthetized with sevoflurane at a dose of 240 MAC × min (4 h at 1 MAC) on day 7 of life ($n = 16$) and non-anesthetized controls ($n = 17$) are shown in Figure 2c. Anesthetized rats had significantly reduced recollection relative to controls. The group recollection index of controls was 0.36, whereas the group recollection index of anesthetized rats was zero. This impression was confirmed by analysis of data derived from the individual ROC curves: 8 of 16 rats in the sevoflurane group had a recollection index of zero, whereas only 2 of 17 controls had a recollection index of zero. The median recollection index of controls was 0.2 compared with 8.9×10^{-5} for sevo-

Table 4 Spearman Rank Order Correlation Coefficients with Characteristics and Outcome Variables

Item	1	2	3	4	5	6	7	8	9	10
Anesthesia status (0 = patient; 1 = control)	1									
Age at testing	− 0.11	1								
Total problems CBCL	− 0.04	0.1	1.99							
Gender	0.11	− 0.07	− 0.01	1						
Full-scale IQ	− 0.02	− 0.14	− 0.27[a]	− 0.07	1					
Family income $1000	− 0.31[a]	− 0.09	− 0.12	− 0.1	0.5[b]	1				
Color recollection	0.32[a]	0.25	− 0.12	0.15	− 0.06	− 0.03	1			
Spatial recollection	0.27[a]	0.19	− 0.03	− 0.01	− 0.19	− 0.22	0.2	1		
Color familiarity	0.06	0.23	− 0.02	− 0.12	0.05	− 0.14	0.21	0.22	1	
Spatial familiarity	− 0.1	0.24	− 0.23	− 0.12	0.2	0.02	0.06	0.32[a]	0.26	1

Abbreviation: CBCL, child behavior checklist.
[a]Correlation is significant at the 0.05 level (two-tailed).
[b]Correlation is significant at the 0.01 (two-tailed).

Table 5 Regression Betas of Demographic Characteristics on Recollection and Familiarity Estimates

Characteristic	Recollection				Familiarity			
	Color task		Spatial task		Color task		Spatial task	
	β	SE	β	SE	β	SE	β	SE
Anesthesia status (0 = Patient; 1 = Control)	0.35[a]	0.06	0.29[a]	0.06	0	0.17	− 0.06	0.19
Age at testing	0.27[a]	0.02	0.21	0.02	0.14	0.05	0.24	0.06
Total Problems CBCL	− 0.17	0	− 0.1	0	− 0.02	0.01	− 0.21	0.01
Gender (0 = Male; 1 = Female)	0.14	0.06	− 0.04	0.06	− 0.12	0.17	− 0.1	0.2
Full-Scale IQ	− 0.11	0	− 0.16	0	0.17	0.01	0.22	0.01
Family Income $1000	0.15	0	0	0	− 0.17	0	− 0.2	0
Constant	—	0.35	—	0.35	—	1.04	—	1.2

Beta values are reported from four separate linear regressions in which anesthesia status, age at testing, total problems CBCL, gender, total IQ, and family income predict color recollection, spatial recollection, color familiarity, or spatial familiarity, respectively. SE for each coefficient is reported in parentheses below each standardized beta coefficient. Results reveal ΔR^2 (root mean squared error) for color recollection, 0.21 (0.20); spatial recollection, 0.14 (0.20); color familiarity 0.06 (0.60); and spatial familiarity 0.16 (0.68), with $N = 56$, in each case.
[a]Significance $P \leqslant 0.05$.

flurane ($P = 0.027$, Mann–Whitney U-test). Familiarity indices, on the other hand, were not different between groups (control median: 0.73 vs anesthesia median: 0.67; $P = 0.69$, Mann–Whitney U-test).

An exploratory analysis revealed that group ROC curves, as well as individual recollection and familiarity indices, for rats who had been tail clamped during anesthesia ($n = 8$) were not different from those of rats that had received anesthesia without tail clamping (Figure 2d).

Odor dilution test

We performed separate experiments to investigate whether the rodent recognition memory testing may have been affected by the sense of smell. A random sample of anesthetized rats ($n = 3$) and control rats ($n = 5$) underwent odor recognition testing with stepwise dilution of odors.

Even at a dilution of 10^{-9} of the original scented sand—a significantly weaker concentration than that used in the recognition memory testing—both control and anesthetized rats were able to detect the odor with greater than 80% accuracy. Only when new (unscented) playground sand was used did performance drop to chance in both groups as expected (Supplementary Figure 1). Performance returned to > 80% in both groups with reintroduction of the odors at a dilution of 10^{-9}. It remains possible that anesthetized rats have an altered sense of smell; however, the threshold for such a deficit is below a dilution of $1:10^9$ relative to the concentration used in the study.

DISCUSSION

The main finding of this study is that general anesthesia in infancy impairs recollection but not familiarity in humans,

although neither IQ nor CBCL scores were adversely affected. A comparable deficit in recollection-like memory was identified in recognition memory experiments in rats. The fact that a single episode of anesthesia in infancy impairs certain aspects of brain function in animals is now widely accepted (for review see (Stratmann, (2011))). However, the relevance to humans and the particular cognitive effects that might occur following anesthesia have been unclear (DiMaggio et al, 2009; Kalkman et al, 2009; Wilder et al, 2009). The present study provides valuable insight into these questions.

Episodic memory is the memory of past experiences and, as it was first described, involves the conscious recollection of an event (Tulving, 1983). It cannot be confirmed whether episodic memory occurs in nonhuman species, as it requires introspection and consciously re-experiencing the past. However, it has been shown that animals may have episodic-like memory that can be demonstrated through tests involving memory for 'what,' 'where,' and 'when' details of an event. This was first described in birds (Clayton and Dickinson, 1998) and more recently in rodents (Dere et al, 2005; Eacott et al, 2005; Eacott and Norman, 2004; Fortin et al, 2004; Kart-Teke et al, 2006).

In rodents, as in humans, recognition memory may be further explained by a dual-process model where recognition comprises of familiarity and recollection (although not uniformly accepted (Eichenbaum et al, 2008; Wixted and Squire, 2008)). Animal models have since been developed to support the idea of recollection-like memory in rodents, and that it can be distinguished from the process of familiarity (Eacott et al, 2005; Easton and Eacott, 2010; Eichenbaum et al, 2010; Sauvage, 2010). These findings suggest that rodents, like humans, rely on recollection and familiarity to perform recognition tasks. Despite reduced recollection, rats can accurately perform recognition memory tasks by compensating with an increased reliance on familiarity (Sauvage et al, 2008). Nevertheless, a recollection deficit remains important, since under certain circumstances, for instance, if the delay between memory encoding and memory retrieval is increased, recognition memory relies to a greater extent on recollection and could thus be impaired (Fortin et al, 2004).

The finding that anesthesia impairs recollection in children and recollection-like memory in rats has important implications. While the human data alone cannot rule out the contribution of the surgical procedure and/or the underlying disease to memory deficits, because rats have comparable deficits, it suggests that anesthesia in infancy is responsible for impairing recollection later in life. We studied a cohort of children who were likely free of confounding medical conditions. It is reasonable to assume that rats, too, were either free of underlying diseases or that medical conditions were randomly distributed between anesthesia and control groups. We then examined a defined cognitive outcome—recognition memory—and its underlying processes of recollection and familiarity. We found that recognition memory accuracy was impaired in the anesthesia group, which may be attributed to reduced recollection, and this occurred in spite of higher socio-economic backgrounds of anesthetized children. Consistent with the literature (Noble et al, 2007), there was a positive correlation between family income and IQ in our subjects.

Also, although not significant, there was a trend toward improved recognition memory performance with higher income. These findings would have biased against the hypothesis and results.

Rats show a similar decrease in recollection but not familiarity in the absence of concurrent tissue damage, which supports the notion that the anesthetic, rather than the surgical procedure or underlying condition, causes the recollection deficit. Consistent with this, we previously found that both sevoflurane and isoflurane impair short-term memory in a water maze task in rats, regardless of tissue injury (Shih et al, 2012; Stratmann et al, 2009a). The type of tissue injury used in half of the rats to assess anesthetic depth (tail clamping) is at least as noxious as surgical tissue injury (Eger et al, 1965). The fact that a similar test detects an anesthetic effect establishes a translational model of anesthetic neurotoxicity in infantile humans and rats. This should allow a more focused therapeutic approach to this problem by first assessing the effect of interventions in rats before designing clinical trials.

Among 28 anesthetized subjects, nine had a familiarity index of zero, and two of these also had a recollection index of zero. The zero-value indices, however, occurred in different tasks—the color task for one individual and the spatial task for the other. The higher frequency of zero-value familiarity indices does not indicate that anesthetized subjects were primarily affected in familiarity. Eight individuals in the control group also had a familiarity index of zero, and only the recollection indices for the anesthetized group were significantly lower than the control group. A zero value for familiarity may be in part due to the constraints of the model and fitting an individual's performance to those parameters and more likely reflects the variability between individuals in their use of familiarity and recollection to recognize items, as one may compensate for the other to achieve successful recognition (Sauvage et al, 2008; Fortin et al, 2004).

Clinical significance of the observed primary outcome

What problems might an impairment in recollection-based memory present in day-to-day life? Recollection has demonstrated roles in autobiographical memory, prospection, classroom learning, reading comprehension, etc. (Ghetti and Bauer, 2012). Thus, even subtle recollection deficits may have immediate consequences and reduce the child's potential to learn over time, which future studies should examine more closely. However, other factors such as motivation, attention, and intelligence also determine a child's ability to learn and succeed. It has been suggested that anesthesia that impairs motivation in primates years after exposure (Paule et al, 2011) is associated with attention deficit disorder in humans (Sprung et al, 2012), and with disabilities in language and cognition (Ing et al, 2012) in humans. The contributions of these potential anesthetic effects to the cognitive state in general, and their significance to recognition memory in particular, will have to be unraveled in future studies. Also, the cumulative lifetime impact of 20–25% impairment in children's recollection memory, as observed in this study, may be substantially greater than what might be apparent at 6–11 years of age and also warrants further study.

Demographic and clinically relevant variables

There were gender differences with respect to age at first anesthetic exposure, anesthetic duration, and vulnerability to anesthesia-induced impairment in recollection. Boys were overrepresented in the anesthetic cohort, which is expected, given the male preponderance among infants requiring surgery (Bartels *et al*, 2009; Block *et al*, 2012; DiMaggio *et al*, 2009; Kalkman *et al*, 2009; Sprung *et al*, 2012; Wilder *et al*, 2009). Boys were exposed roughly 2.5 months earlier than girls, and boys appeared to be more vulnerable to the effect of anesthesia on recollection than girls, which is consistent with recently published results from an animal study (Lee *et al*, 2014). Although recollection deficits were observed in both sexes, only the males exhibited a significant impairment in the spatial task. In addition, there was some indication that females who received anesthesia before 6 months of age performed better than those exposed later in life and that longer durations of anesthesia were associated with worse memory outcomes. The small sample size of the current study, however, limits the conclusiveness of these potential differences, and future studies evaluating gender, anesthetic duration, and age of exposure will be important.

Study limitations and areas of further study

There is a potential source of bias inherent in missing data from inability to contact or recruit participants. We do not know how this might have affected the results of this study. However, the fact that rodent and human data show substantial agreement seemingly validates the anesthetic effects. Further, absence of an effect of tissue injury in rats suggests that anesthesia, and not the surgical procedure, causes the recollection impairment. We cannot rule out the possibility, though, that the underlying condition or surgical procedure caused the deficit in humans, while anesthetic exposure alone led to the same outcome in rodents. This possibility, although unlikely, must be considered in future studies.

A causal link between anesthesia during infancy and long-term impairment of recollection would be further supported if the memory deficit in humans was observed in the absence of surgery (for instance, in children who only underwent magnetic resonance imaging scan under anesthesia). Another possible approach for testing the strength of the association between anesthesia and impaired recollection would be to compare a purely regional anesthetic *vs* general anesthesia for a given underlying condition or surgical procedure. Alternatively, monozygotic twins discordant for anesthetic exposure might be compared using cognitive endpoints.

It has been previously suggested that single anesthetic exposures are not associated with adverse cognitive outcomes, unlike multiple exposures (Sprung *et al*, 2012; Wilder *et al*, 2009). We found no difference between patients receiving single ($n = 18$) *vs* multiple ($n = 10$) anesthetics in any of the measured outcomes. This is consistent with a recent report suggesting that a single anesthetic exposure is associated with deficits in language and cognition (Ing *et al*, 2012) and evidence against the notion that a single anesthetic exposure in infancy is innocuous.

Finally, the anesthetic administered in the human arm of this study always included nitrous oxide and nearly always included sevoflurane and a narcotic. Exploratory analyses did not reveal an effect of the anesthetic agent used, which may be related to the small sample sizes of subgroups receiving a particular anesthetic agent. Given that individual agents act on unique receptors, understanding their differential contributions to the impairment in recollection is an important area of further investigation.

In conclusion, anesthesia in infancy impairs recognition memory performance. Recollection, specifically, accounts for this deficit and is impaired in both humans and rats anesthetized during infancy, while familiarity is unaffected.

AUTHOR CONTRIBUTIONS

GS designed both the human and the rat arms of the study, established the methods for rat recognition memory testing in the lab, organized and oversaw data collection of both trials, performed the data analysis, and wrote the manuscript. JL co-designed the human trial, tested all of the human subjects, co-analyzed the data, and co-authored the manuscript. JS co-designed both arms of the study, co-analyzed the human and rat data, and co-authored the manuscript. BL co-designed the overall rat experiment, co-analyzed the rat data and made Figure 2. RA co-designed the overall study, collection of the rat pilot data, and co-authored the manuscript. JS helped establish the methods for rat recognition memory testing in our lab, co-designed and managed data collection of the rat pilot trial, co-analyzed the data, and co-authored the manuscript. AR co-designed data collection of the rat pilot trial, performed recognition memory testing in rats, co-analyzed the data, and co-authored the manuscript. TR and FC co-designed the second and third rat trials, organized and managed data collection, performed recognition memory testing in rats,

and co-authored the manuscript. TA, DL, KS, SE, JC, co-designed the second rat trial, performed recognition memory testing in rats and co-authored the manuscript. NL co-designed and oversaw the third rat trial, performed recognition memory testing in rats, co-analyzed the data, and co-authored the manuscript. AW, CS, AV, HZ, MW, AA, NK, JL, JS, EL, GB, co-designed the third rat trial, performed recognition memory testing in rats and co-authored the manuscript. KB, DB-T, co-designed the rat pilot trial, performed recognition memory testing in rats, and co-authored the manuscript. CL co-designed the human trial, oversaw data collection at the Davis site, and co-authored the manuscript. ZQ co-designed the human trial, co-authored the manuscript, assisted with data collection, data analysis, and manuscript submission. AA, EC co-designed the human trial, assisted with data collection at the San Francisco site and co-authored the manuscript. AY co-designed the human trial, established the method of recognition memory testing using ROC curve analysis in humans, co-analyzed the data, and co-authored the manuscript. SG co-designed the human trial, established and maintained the registry of control participants, oversaw and managed data collection at the Davis site, co-analyzed the data, and co-authored the manuscript.

REFERENCES

Bartels M, Althoff RR, Boomsma DI (2009). Anesthesia and cognitive performance in children: no evidence for a causal relationship. *Twin Res Hum Genet* **12**: 246–253.

Block RI, Thomas JJ, Bayman EO, Choi JY, Kimble KK, Todd MM (2012). Are anesthesia and surgery during infancy associated with altered academic performance during childhood? *Anesthesiology* **117**: 494–503.

Clayton NS, Dickinson A (1998). Episodic-like memory during cache recovery by scrub jays. *Nature* **395**: 272–274.

Cycowicz YM, Friedman D, Rothstein M, Snodgrass JG (1997). Picture naming by young children: norms for name agreement, familiarity, and visual complexity. *J Exp Child Psychol* **65**: 171–237.

Dere E, Huston JP, De Souza Silva MA (2005). Episodic-like memory in mice: simultaneous assessment of object, place and temporal order memory. *Brain Res Protoc* **16**: 10–19.

DiMaggio C, Sun LS, Kakavouli A, Byrne MW, Li G (2009). A retrospective cohort study of the association of anesthesia and hernia repair surgery with behavioral and developmental disorders in young children. *J Neurosurg Anesthesiol* **21**: 286–291.

DiMaggio C, Sun LS, Li G (2011). Early childhood exposure to anesthesia and risk of developmental and behavioral disorders in a sibling birth cohort. *Anesth Analg* **113**: 1143–1151.

Eacott MJ, Easton A, Zinkivskay A (2005). Recollection in an episodic-like memory task in the rat. *Learn Mem* **12**: 221–223.

Eacott MJ, Norman G (2004). Integrated memory for object, place, and context in rats: a possible model of episodic-like memory? *J Neurosci* **24**: 1948–1953.

Easton A, Eacott MJ (2010). Recollection of episodic memory within the medial temporal lobe: behavioural dissociations from other types of memory. *Behav Brain Res* **215**: 310–317.

Eger EI 2nd, Saidman LJ, Brandstater B (1965). Minimum alveolar anesthetic concentration: a standard of anesthetic potency. *Anesthesiology* **26**: 756–763.

Eichenbaum H, Fortin N, Sauvage M, Robitsek RJ, Farovik A (2010). An animal model of amnesia that uses Receiver Operating Characteristics (ROC) analysis to distinguish recollec-tion from familiarity deficits in recognition memory. *Neuropsychologia* **48**: 2281–2289.

Eichenbaum H, Sauvage MM, Fortin NJ, Yonelinas AP (2008). ROCs in rats? Response to Wixted and Squire. *Learn Mem* **15**: 691–693.

Eichenbaum H, Yonelinas AP, Ranganath C (2007). The medial temporal lobe and recognition memory. *Annu Rev Neurosci* **30**: 123–152.

Fortin NJ, Wright SP, Eichenbaum H (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature* **431**: 188–191.

Ghetti S, Angelini L (2008). The development of recollection and familiarity in childhood and adolescence: evidence from the dual-process signal detection model. *Child Dev* **79**: 339–358.

Ghetti S, Bauer PJ (2012). *Origins And Development Of Recollection: Perspectives From Psychology And Neuroscience*. Oxford University Press: USA.

Ghetti S, Lee JK, Sims CE, Demaster DM, Glaser NS (2010). Diabetic ketoacidosis and memory dysfunction in children with type 1 diabetes. *J Pediatr* **156**: 109–114.

Hansen TG, Pedersen JK, Henneberg SW, Morton NS, Christensen K (2013). Educational outcome in adolescence following pyloric stenosis repair before 3 months of age: a nationwide cohort study. *Paediatr Anaesthesia* **23**: 883–890.

Ing C, DiMaggio C, Whitehouse A, Hegarty MK, Brady J, von Ungern-Sternberg BS *et al* (2012). Long-term differences in language and cognitive function after childhood exposure to anesthesia. *Pediatrics* **130**: e476–e485.

Jevtovic-Todorovic V, Hartman RE, Izumi Y, Benshoff ND, Dikranian K, Zorumski CF *et al* (2003). Early exposure to common anesthetic agents causes widespread neurodegeneration in the developing rat brain and persistent learning deficits. *J Neurosci* **23**: 876–882.

Kalkman CJ, Peelen L, Moons KG, Veenhuizen M, Bruens M, Sinnema G *et al* (2009). Behavior and development in children and age at the time of first anesthetic exposure. *Anesthesiology* **110**: 805–812.

Kart-Teke E, De Souza Silva MA, Huston JP, Dere E (2006). Wistar rats show episodic-like memory for unique experiences. *Neurobiol Learn Mem* **85**: 173–182.

Kodama M, Satoh Y, Otsubo Y, Araki Y, Yonamine R, Masui K *et al* (2011). Neonatal desflurane exposure induces more robust neuroapoptosis than do isoflurane and sevoflurane and impairs working memory. *Anesthesiology* **115**: 979–991.

Lee BH, Chan JT, Kraeva E, Peterson K, Sall J (2014). Isoflurane exposure in newborn rats induces long-term cognitive dysfunction in males but not females. *Neuropharmacology* **83**: 9–17.

Loepke AW, Soriano SG (2008). An assessment of the effects of general anesthetics on developing brain structure and neurocognitive function. *Anesth Analg* **106**: 1681–1707.

Noble KG, McCandliss BD, Farah MJ (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Dev Sci* **10**: 464–480.

Paule MG, Li M, Allen RR, Liu F, Zou X, Hotchkiss C *et al* (2011). Ketamine anesthesia during the first week of life can cause long-lasting cognitive deficits in rhesus monkeys. *Neurotoxicol Teratol* **33**: 220–230.

Pyhala R, Lahti J, Heinonen K, Pesonen AK, Strang-Karlsson S, Hovi P *et al* (2011). Neurocognitive abilities in young adults with very low birth weight. *Neurology* **77**: 2052–2060.

Robitsek RJ, Fortin NJ, Koh MT, Gallagher M, Eichenbaum H (2008). Cognitive aging: a common decline of episodic recollection and spatial memory in rats. *J Neurosci* **28**: 8945–8954.

Sanders RD, Xu J, Shu Y, Januszewski A, Halder S, Fidalgo A *et al* (2009). Dexmedetomidine attenuates isoflurane-induced neurocognitive impairment in neonatal rats. *Anesthesiology* **110**: 1077–1085.

Satomoto M, Satoh Y, Terui K, Miyao H, Takishima K, Ito M *et al* (2009). Neonatal exposure to sevoflurane induces abnormal social behaviors and deficits in fear conditioning in mice. *Anesthesiology* **110**: 628–637.

Sauvage MM (2010). ROC in animals: uncovering the neural substrates of recollection and familiarity in episodic recognition memory. *Conscious Cogn* **19**: 816–828.

Sauvage MM, Fortin NJ, Owens CB, Yonelinas AP, Eichenbaum H (2008). Recognition memory: opposite effects of hippocampal damage on recollection and familiarity. *Nat Neurosci* **11**: 16–18.

Shih J, May LD, Gonzalez HE, Lee EW, Alvi RS, Sall JW *et al* (2012). Delayed environmental enrichment reverses sevoflurane-induced memory impairment in rats. *Anesthesiology* **116**: 586–602.

Sprung J, Flick RP, Katusic SK, Colligan RC, Barbaresi WJ, Bojanic K *et al* (2012). Attention-deficit/hyperactivity disorder after early exposure to procedures requiring general anesthesia. *Mayo Clin Proc* **87**: 120–129.

Sprung J, Flick RP, Wilder RT, Katusic SK, Pike TL, Dingli M *et al* (2009). Anesthesia for cesarean delivery and learning disabilities in a population-based birth cohort. *Anesthesiology* **111**: 302–310.

Stratmann G (2011). Review article: neurotoxicity of anesthetic drugs in the developing brain. *Anesth Analg* **113**: 1170–1179.

Stratmann G, May LD, Sall JW, Alvi RS, Bell JS, Ormerod BK *et al* (2009a). Effect of hypercarbia and isoflurane on brain cell death and neurocognitive dysfunction in 7-day-old rats. *Anesthesiology* **110**: 849–861.

Stratmann G, Sall JW, Eger EI 2nd, Laster MJ, Bell JS, May LD *et al* (2009b). Increasing the duration of isoflurane anesthesia decreases the minimum alveolar anesthetic concentration in 7-day-old but not in 60-day-old rats. *Anesth Analg* **109**: 801–806.

Stratmann G, Sall JW, May LD, Bell JS, Magnusson KR, Rau V *et al* (2009c). Isoflurane differentially affects neurogenesis and long-term neurocognitive function in 60-day-old and 7-day-old rats. *Anesthesiology* **110**: 834–848.

Tulving E (1983). *Elements of Episodic Memory*. Oxford University Press: New York, xi, 351 pp.

Volpe JJ (2009). Cerebellum of the premature infant: rapidly developing, vulnerable, clinically important. *J Child Neurol* **24**: 1085–1104.

Wilder RT, Flick RP, Sprung J, Katusic SK, Barbaresi WJ, Mickelson C *et al* (2009). Early exposure to anesthesia and learning disabilities in a population-based birth cohort. *Anesthesiology* **110**: 796–804.

Wixted JT, Squire LR (2008). Constructing receiver operating characteristics (ROCs) with experimental animals: cautionary notes. *Learn Mem* **15**: 687–690.

Yonelinas AP (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J Exp Psychol Learn Mem Cogn* **20**: 1341–1354.

Yonelinas AP (1997). Recognition memory ROCs for item and associative information: the contribution of recollection and familiarity. *Mem Cogn* **25**: 747–763.

Yonelinas AP, Aly M, Wang WC, Koen JD (2010). Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus* **20**: 1178–1194.

Yonelinas AP, Kroll NE, Quamme JR, Lazzara MM, Sauve MJ, Widaman KF *et al* (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nat Neurosci* **5**: 1236–1241.

Zhu C, Gao J, Karlsson N, Li Q, Zhang Y, Huang Z *et al* (2010). Isoflurane anesthesia induced persistent, progressive memory impairment, caused a loss of neural stem cells, and reduced neurogenesis in young, but not adult, rodents. *J Cereb Blood Flow Metab* **30**: 1017–1030.