BRIEF REPORT

Improving Pulse Oximetry Pitch Perception with Multisensory Perceptual Training

Joseph J. Schlesinger, MD,*† Ryan A. Stevenson, PhD,‡§‖¶** Matthew S. Shotwell, PhD,*#
and Mark T. Wallace, PhD‡§‖**††

The pulse oximeter is a critical monitor in anesthesia practice designed to improve patient safety. Here, we present an approach to improve the ability of anesthesiologists to monitor arterial oxygen saturation via pulse oximetry through an audiovisual training process. Fifteen residents' abilities to detect auditory changes in pulse oximetry were measured before and after perceptual training. Training resulted in a 9% (95% confidence interval, 4%–14%, $P = 0.0004$, $t_{166} = 3.60$) increase in detection accuracy, and a 72-millisecond (95% confidence interval, 40–103 milliseconds, $P < 0.0001$, $t_{166} = -4.52$) speeding of response times in attentionally demanding and noisy conditions that were designed to simulate an operating room. This study illustrates the benefits of multisensory training and sets the stage for further work to better define the role of perceptual training in clinical anesthesiology. (Anesth Analg 2014;118:1249–53)

An anesthesia provider must critically manage patients in the operating room and respond to any physiologic aberration in an environment that contains a large amount of sensory information. Relevant information must be appropriately assimilated and irrelevant information filtered.[1] Attention to the pulse oximeter, a vital piece of monitoring equipment, needs to be particularly acute during induction and emergence, the period when attention is heavily directed toward the patient and when errors are most likely to occur.[2,3]

Dividing attention and triaging clinical events and tasks are imperative to the anesthesiologist, and these challenges are amplified under circumstances in which the physician's attentional resources are at or beyond capacity.[4] The consequences of divided attention have been the subject of a significant amount of psychological, neuroscience, and human factors research, including applied studies asking how such abilities might be improved.[4–10] Indeed, training in which subjects make use of combined visual and auditory (i.e., multisensory) cues has been shown to be able to improve sensory performance and perception.[10–12] The ability to integrate information across multiple senses leads to improvements in detection,[13,14] accuracy,[15,16] and response times.[17,18]

We assessed whether perceptual training would improve the ability of resident anesthesiologists to detect frequency changes used to report oxygen saturation.

METHODS

This study was approved by the Vanderbilt University Medical Center IRB, and written consent was obtained from all participants. Anesthesia residents were recruited by electronic mail. Forty electronic mail messages were sent to residents in the Department of Anesthesiology who were postgraduate years 1 to 4 and scheduled on a clinical rotation that was amenable to study participation. Fifteen residents replied and participated in the study. All participants had normal hearing and normal or correct-to-normal vision. Before multisensory perceptual training, all 15 resident anesthesiologists completed a session measuring their sensitivity to changes in pulse oximetry frequency.[4] This session tested sensitivity to changes in simulated pulse oximeter frequency over 3 levels of attentional load, as well as in 2 levels of noise (in quiet and with typical operating room background noise (Fig. 1, A–C)). The order of the 6 conditions was randomly permuted but uniformly applied to each participant. Participants then underwent multisensory perceptual training regimens where visual rings and auditory beeps were presented with varying stimulus onset asynchronies (SOAs) (Fig. 2A), after which they completed a posttraining pulse oximetry session identical to that on experiment day 1. The details of the paradigm and training are contained in Figure 1.[4]

STATISTICAL METHODS

Response time and accuracy (dependent variables) were separately examined using a linear mixed-effects model with fixed effects for attentional load (low, medium, or high),

---

From the *Department of Anesthesiology, Division of Critical Care Medicine, and †Department of Anesthesiology, BH Robbins Scholars Program, ‡Department of Hearing and Speech Sciences, Vanderbilt University Medical Center; §Vanderbilt Brain Institute; ‖Vanderbilt Kennedy Center, Nashville, Tennessee; ¶Department of Psychology, University of Toronto, Toronto, Ontario, Canada; #Department of Biostatistics, Vanderbilt University, Nashville, Tennessee; **Department of Psychology, Vanderbilt University, Nashville, Tennessee; and ††Department of Psychiatry, Vanderbilt University, Nashville, Tennessee.

Accepted for publication January 29, 2014.

Funding: Joseph Schlesinger: Vanderbilt Institute for Clinical and Translational Research VR2236; Ryan Stevenson: Banting Postdoctoral Fellowship, NIH 1F32 DC011993; Mark Wallace: Vanderbilt Brain Institute.

The authors declare no conflicts of interest.

This report was previously presented, in part, at the American Society of Anesthesiologists, Society of Critical Care Anesthesiologists, Society for Neuroscience, American Medical Association, Association of University Anesthesiologists, Society of Critical Care Medicine, and Society for Technology in Anesthesia.

Reprints will not be available from the authors.

Address correspondence to Joseph J. Schlesinger, MD, 1211 21st Avenue South, Medical Arts Building, Suite 526, Nashville, TN, 37212. Address e-mail to joseph.j.schlesinger@vanderbilt.edu.

DOI: 10.1213/ANE.0000000000000222

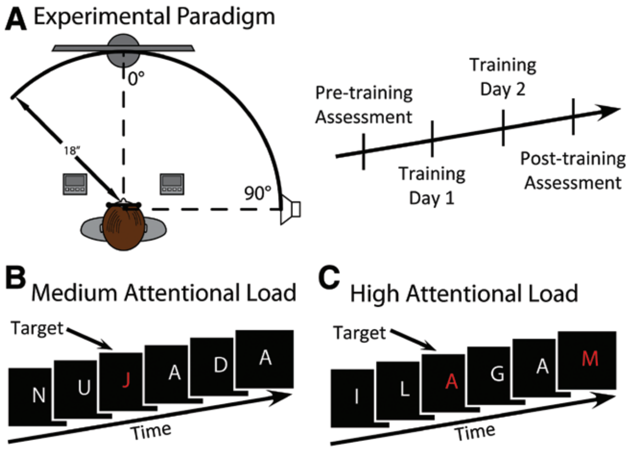

Figure 1. A depicts the experimental setup. Individuals performed the visual tasks shown in B and C in the central field of vision while simultaneously performing a pulse oximetry task that was presented at a 90° angle from the participant's head on their right side, similar to a standard pulse oximetry setup. This task consisted of detecting a reduction in pitch associated with a decrease in saturation from 99% (648 Hz) to 98% (630 Hz) delivered at a rate of 75 pulses/min and at a volume of 80 dB sound pressure level corresponding to the default QRS volume setting of 2 on the Philips patient monitor. Background noise, consisting of prerecorded operating room noise during off-pump cardiac surgery without pulse oximetry beeps at an average of 67 dB sound pressure level, was presented on half of the conditions, completing a 3 × 2 (attentional load × noise) experimental design. Participants responded to a decrease in frequency via key press. B and C depict the medium- and high-attentional load visual tasks. Low-attentional load is not depicted because it merely consisted of visual fixation. Visual stimuli consisted of individual letters presented in the central visual field in rapid serial visual presentation (RSVP) at a rate of 10 Hz. Visual stimuli included 25 capital letters (excluding "Y") presented in either red or white. For the medium-attentional load condition, participants' task was to detect any red letter. For the high-attentional load condition, participants' task was to detect red vowels. Targets were always presented on 4% of trials. Participants were presented with these tasks in blocks of 6000 letter presentations for each condition, with 240 targets. Participants simultaneously completed the auditory task described above, with 750 pulses presented, 10% of which (75) were targets. These 6 conditions were completed both before and after training, with accuracies and response times measured for each participant and each condition. For the auditory task, participants responded to a decrease in frequency. For the visual tasks, participants responded to a red letter in the medium-attentional load condition and a red vowel in the high-attentional load condition. All responses were made via key press.

background noise (noise or no noise), and training (before or after). All possible interactions were considered. A random intercept for each participant was used to account for the correlation among repeated measurements and to reflect the random sampling nature of participant enrollment.

A likelihood ratio (chunk) test was used to assess the interaction and overall effects of each factor. Contrasts for the effect of training in each combination of attentional load and noise level were summarized using a point estimate and Wald-type 95% confidence interval (CI) test. Quantile-quantile (QQ) plots were used to assess residual normality. Levene test was used to assess homogeneity of variance among experimental stratum, with a significance threshold of $P < 0.01$. The symbols η_p^2, χ_k^2, and t_k represent the partial η^2, χ^2, and Student t statistic with k degrees of freedom.

Residual diagnostics were evaluated in the pulse oximetry accuracy and response time data. Levene test resulted in P values 0.71 and 0.044 for response time and accuracy, respectively. The QQ plot for response time residuals was unremarkable. The QQ plot for accuracy indicated that residuals were sufficiently heavy tailed at the lower end to warrant additional scrutiny. We elected to use a conservative significance threshold of 0.01 for P values associated with the accuracy end point. A conventional threshold of 0.05 was used for tests associated with response time. CIs that had failed to include an appropriate null value were also considered statistically significant.

We conducted a bootstrap-based simulation study to assess the post hoc power and type-I error associated with the likelihood ratio test for an overall training effect, by resampling the observed sample data at the participant level. To estimate type-I error, the null hypothesis was imposed on the sample data by subtracting the estimated main and interaction effects due to training. The estimated type-I error was within 1% point of the nominal value for both response time (estimated type-I error: 0.048) and accuracy (estimated type-I error: 0.015). Given the magnitude of the effect of training in these data, the power to detect such an effect under the current design and test was >0.99 for response time, and >0.76 for accuracy.

RESULTS

Environmental noise significantly impaired residents' accuracy and response time to detect changes in tones designed to simulate decreasing oxygen saturation by pulse oximetry. Attentional load also significantly impaired response time. Multisensory training significantly improved both accuracy and response time (Table 1). A significant interaction was observed between attentional load and training for response time ($P = 0.0058$, $\eta_p^2 = 0.08$, $\chi_4^2 = 14.52$). No other 2- or 4-way interactions were significant.

Figures 3 and 4 illustrate the effects of training for each combination of attentional load and noise on accuracy and response time, respectively. In quiet, there were significant improvements in response time for the high-attentional load condition (95% CI, 9–71 milliseconds, $P = 0.011$, $t_{166} = -2.58$) but not for the medium (95% CI, −6 to 58 milliseconds, $P = 0.11$, $t_{166} = -1.62$) or low (95% CI, −32 to 30 milliseconds, $P = 0.95$, $t_{166} = 0.06$) attentional load conditions. The corresponding effects on accuracy were not significant for the high (95% CI, 0%–10%, $P = 0.048$, $t_{166} = 1.99$), medium (95% CI, −2% to 8%, $P = 0.30$, $t_{166} = 1.04$), or low (95% CI, −5% to 5%, $P = 0.91$, $t_{166} = -0.11$) attentional load conditions. In noise, there was a significant improvement in accuracy and response time for the high-attentional load condition (95% CI, 4%–14%, $P = 0.0004$, $t_{166} = 3.60$ and 95% CI, 40–103 milliseconds, $P < 0.0001$, $t_{166} = -4.52$, respectively) but not for the medium (95% CI, −3% to 7%, $P = 0.45$, $t_{166} = 0.75$ and 95% CI, 0–62 milliseconds, $P = 0.052$, $t_{166} = -1.96$) or low load condition (95% CI, −5% to 4%, $P = 0.83$, $t_{166} = -0.21$, and 95% CI, −36 to 27 milliseconds, $P = 0.76$, $t_{166} = 0.31$).

To ensure increases in pulse oximetry performance were not due to decreased attention to the visual tasks after training, an identical mixed-effects regression analysis was conducted on visual task accuracies and response times with training. The only significant overall effect on accuracy or response time was due to attentional load ($P < 0.0001$, $\eta_p^2 = 0.70$, $\chi_4^2 = 116.99$, and $P < 0.0001$, $\eta_p^2 = 0.87$, $\chi_4^2 = 212.45$, respectively). Relative to low-attentional load, high load

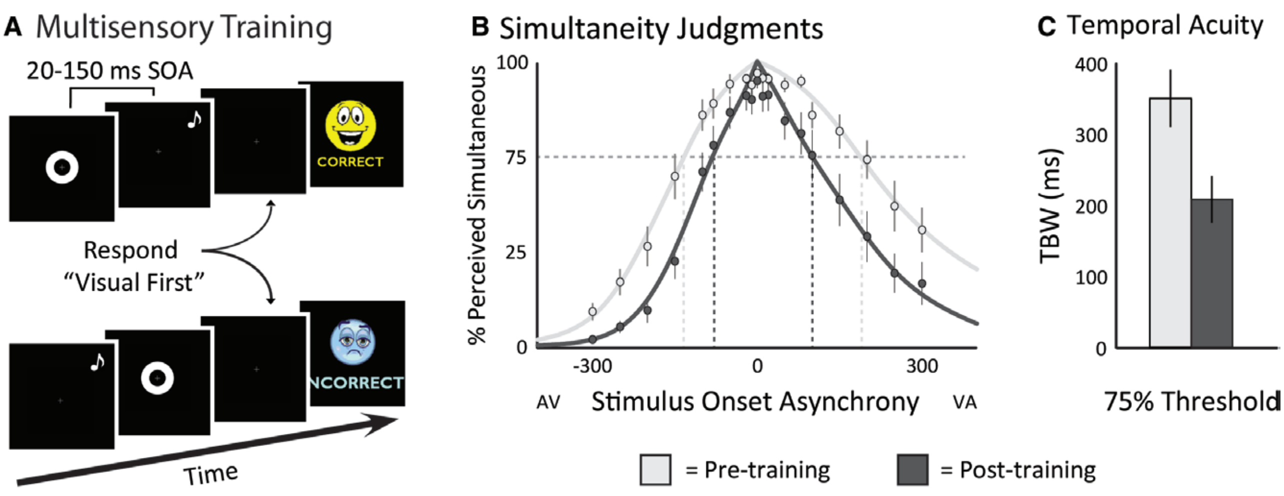

Figure 2. A depicts the multisensory training paradigm. Visual rings and auditory beeps were presented with varying stimulus onset asynchronies (SOAs). Participants responded via key press indicating if they perceived the presentation as "auditory-first" or "visual-first." Each run consisted of 25 audiovisual presentations at each SOA, with SOAs including ±150, ±100, ±50, ±20 for a total of 200 trials (8 conditions × 25 trials each). After participants' responses, feedback indicating the accuracy of the response was provided. B shows rates of perceived simultaneity with each SOA were calculated for each individual. Two psychometric sigmoid functions were then fit by the glmfit function in MATLAB to the rates of perceived simultaneity across SOAs, one to the audio-first presentations and a second to the visual-first presentations that will be referred to as the left (AV) and right (VA) windows, respectively. The 0-millisecond SOA condition was used in both fits. Each participant's individual temporal binding windows (TBWs) were then estimated as the SOA at which the best-fit sigmoid's y value equaled a 75% rate of perceived simultaneity. Five to 10 mean group TBWs were then calculated. To ensure the efficacy of the training procedure, paired-samples *t* tests were used to compare, pre- and posttraining measures of multisensory temporal processing, depicted in C. The TBW was significantly narrowed with training from 350 milliseconds pretraining to 208-millisecond posttraining ($P = 0.0044$, $t_{14} = 3.39$, Cohen's $d = 1.03$). A reduction was observed in 14 of 15 residents ($P = 0.0010$), where only half are expected under the null hypothesis (i.e., training is ineffective). This is strong evidence that the mean effect was not driven by a small subset of individuals.

Table 1. Likelihood Ratio Test Results for the Overall Effects of Training, Attentional Load, and Noise Level

Dependent measure	Factor	P	DF	χ^2	η_p^2
Accuracy	Training	0.0072	6	17.65	0.10
	Noise level	<0.0001	6	99.26	0.46
	Attn. load	0.0368	8	16.41	0.10
Response Time	Training	<0.0001	6	30.61	0.17
	Noise level	<0.0001	6	66.45	0.33
	Attn. load	<0.0001	8	140.12	0.58

Attn. = attentional; DF = degrees of freedom.

was associated with lower accuracy (95% CI, 25%–40%) and increased response time (95% CI, 239–318 milliseconds). Importantly, there was no evidence for an effect of training on either accuracy or response time ($P = 0.35$, $\eta_p^2 = 0.04$, $\chi_4^2 = 4.40$, and $P = 0.33$, $\eta_p^2 = 0.04$, $\chi_4^2 = 4.59$), ensuring that the training effects seen with auditory pulse oximetry performance were not the result of residents learning to attend to the auditory stimulus at the expense of the primary visual task.

These results show that multisensory training successfully improved individuals' multisensory temporal acuity (Fig. 2). The mean temporal binding window was significantly reduced from 350 milliseconds pretraining to 208 milliseconds posttraining ($P = 0.0044$, $t_{14} = 3.39$, Cohen's $d = 1.03$). A reduction was observed in 14 of 15 residents ($P = 0.0010$), where only half are expected under the null hypothesis (i.e., training is ineffective).

DISCUSSION

In the current study, we attempted to improve performance on a fundamental operating room task, monitoring changes in pulse oximetry, by training resident anesthesiologists to better integrate visual and auditory cues. This multisensory perceptual training of anesthesia residents resulted in statistically significant increases in their ability to detect changes in arterial oxygen saturation via auditory pitch changes and increased the speed at which they could do so in an environment most similar to that seen in the operating room, one that was highly attentionally demanding and noisy. Adverse outcomes are typically the result of a number of minor, compounding events, such as failure to detect a change in arterial oxygen saturation.[19] Lack of monitor vigilance and inattention are 2 of the most commonly cited factors associated with critical incidents in anesthesia.[20,21] While a lab setting has the benefit of isolating independent variables, follow-up studies are needed to test how these perceptual improvements transfer to an operating room setting or in a simulator.[22–24] While not tested in the current study, previous work has shown preservation of training effects over the course of months.[7,9,10] In this study, we have improved anesthesiologists' performance through a perceptual training paradigm that is novel to anesthesiology, but that has been used successfully in basic science research settings in the field of sensory perception.[7,8,10] We believe that our findings are generalizable to anesthesia training programs nationwide. While future work should seek to extend these results to more realistic clinical settings, these findings are a first step in a line of research that has the capacity to save lives

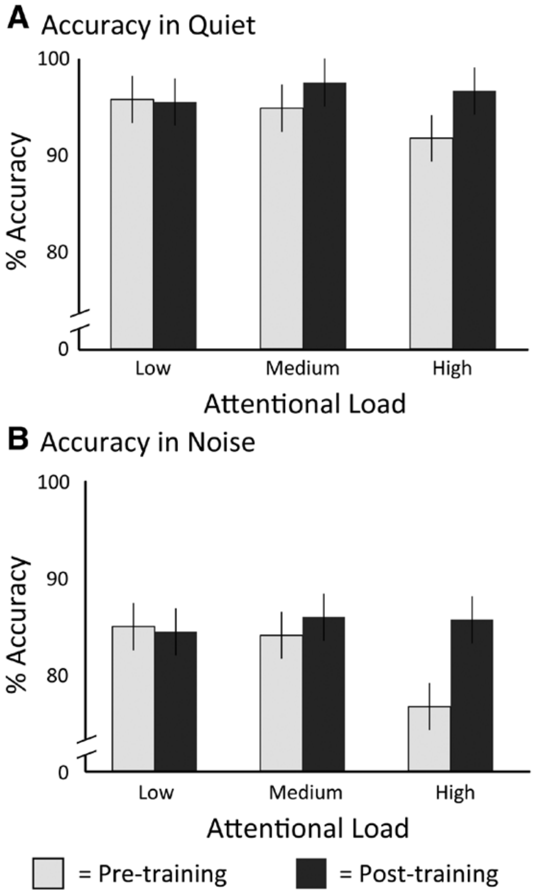

Figure 3. A and B depicts accuracy levels (higher scores are better) for the auditory pulse oximetry pitch detection task in quiet and noise and for each attentional load level. Error bars represent standard error.

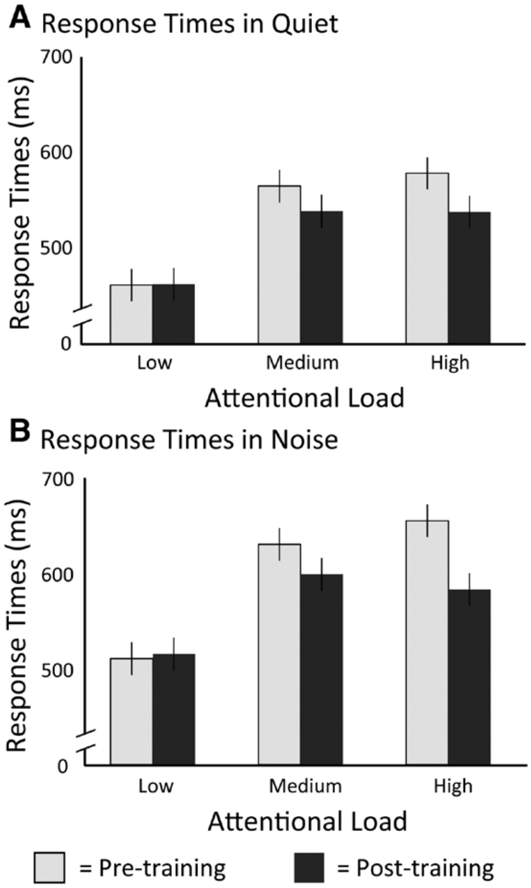

Figure 4. A and B depicts response times (lower values are better) for the auditory pulse oximetry pitch detection task in quiet and noise and for each attentional load level. Error bars represent standard error.

and minimize errors in the operating room by using effective and time-efficient multisensory training to improve an anesthesia provider's subtle unisensory pulse oximetry pitch perception in a dynamic and attentionally demanding perioperative environment.

DISCLOSURES

Name: Joseph J. Schlesinger, MD.
Contribution: This author helped design and conduct the study, analyze the data, and write the manuscript.
Attestation: Joseph J. Schlesinger has seen the original study data, reviewed the analysis of the data, approved the final manuscript, and is the author responsible for archiving the study files.
Name: Ryan A. Stevenson, PhD.
Contribution: This author helped design and conduct the study, analyze the data, and write the manuscript.
Attestation: Ryan Stevenson has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.
Name: Matthew S. Shotwell, PhD.
Contribution: This author helped analyze the data, and write the manuscript.
Attestation: Matthew Shotwell has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.
Name: Mark T. Wallace, PhD.
Contribution: This author helped design the study, analyze the data, and write the manuscript.
Attestation: Mark Wallace has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.
This manuscript was handled by: Maxime Cannesson, MD, PhD.

REFERENCES

1. Mukherjee S. A precarious exchange. N Engl J Med 2004;351:1822–4
2. Gaba DM, Lee T. Measuring the workload of the anesthesiologist. Anesth Analg 1990;71:354–61
3. Loeb RG. Monitor surveillance and vigilance of anesthesia residents. Anesthesiology 1994;80:527–33
4. Stevenson RA, Schlesinger JJ, Wallace MT. Effects of divided attention and operating room noise on perception of pulse oximeter pitch changes: a laboratory study. Anesthesiology 2013;118:376–81
5. Stevenson RA, Zemtsov RK, Wallace MT. Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. J Exp Psychol Hum Percept Perform 2012;38:1517–29
6. Hairston WD, Burdette JH, Flowers DL, Wood FB, Wallace MT. Altered temporal profile of visual-auditory multisensory interactions in dyslexia. Exp Brain Res 2005;166:474–80
7. Stevenson RA, Wilson MM, Powers AR, Wallace MT. The effects of visual training on multisensory temporal processing. Exp Brain Res 2013;225:479–89

8. Powers AR 3rd, Hevey MA, Wallace MT. Neural correlates of multisensory perceptual learning. J Neurosci 2012;32:6263–74
9. Stevenson RA, Fister JK, Barnett ZP, Nidiffer AR, Wallace MT. Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. Exp Brain Res 2012;219:121–37
10. Powers AR 3rd, Hillock AR, Wallace MT. Perceptual training narrows the temporal window of multisensory binding. J Neurosci 2009;29:12265–74
11. Alais D, Cass J. Multisensory perceptual learning of temporal order: audiovisual learning transfers to vision but not audition. PLoS One 2010;5:e11283
12. Shams L, Wozny DR, Kim R, Seitz A. Influences of multisensory experience on subsequent unisensory processing. Front Psychol 2011;2:264
13. Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. Brain Res Cogn Brain Res 2003;17:447–53
14. Stein BE, London N, Wilkinson LK, Price DD. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. J Cogn Neurosci 1996;8:497–506
15. Nelson WT, Hettinger LJ, Cunningham JA, Brickman BJ, Haas MW, McKinley RL. Effects of localized auditory information on visual target detection performance using a helmet-mounted display. Hum Factors 1998;40:452–60
16. Wilkinson F, James TW, Wilson HR, Gati JS, Menon RS, Goodale MA. An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. Curr Biol 2000;10:1455–8
17. Diederich A, Colonius H. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. Percept Psychophys 2004;66:1388–404
18. Hershenson M. Reaction time as a measure of intersensory facilitation. J Exp Psychol 1962;63:289–93
19. Reason J. Human error: models and management. BMJ 2000;320:768–70
20. Cooper JB, Newbower RS, Kitz RJ. An analysis of major errors and equipment failures in anesthesia management: considerations for prevention and detection. Anesthesiology 1984;60:34–42
21. Cooper JB, Newbower RS, Long CD, McPeek B. Preventable anesthesia mishaps: a study of human factors. Anesthesiology 1978;49:399–406
22. Weinger MB, Reddy SB, Slagle JM. Multiple measures of anesthesia workload during teaching and nonteaching cases. Anesth Analg 2004;98:1419–25
23. Gaba DM. Human error in anesthetic mishaps. Int Anesthesiol Clin 1989;27:137–47
24. Gaba DM, DeAnda A. A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. Anesthesiology 1988;69:387–94