

QUALITY AND PATIENT SAFETY

Using an animated patient avatar to improve perception of vital sign information by anaesthesia professionals

D. W. Tscholl^{1,*†}, L. Handschin^{1,†}, P. Neubauer², M. Weiss³, B. Seifert⁴, D. R. Spahn¹ and C. B. Noethiger¹

¹Institute of Anaesthesiology, University and University Hospital Zurich, Zurich, Switzerland, ²Department of Computer Science, University of York, York, UK, ³Leonard N. Stern School of Business, Department of Management and Organizations, New York University, New York, NY, USA and ⁴Department of Biostatistics, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland

*Corresponding author. E-mail: davidtscholl@me.com

† These two authors contributed equally to this work and share first authorship.

Collaborators listed at end of article.

Abstract

Background: Maintaining situation awareness of monitored patients can be challenging because care providers must continually read and integrate multiple waveforms and numerical vital sign values into a mental model of the patient's situation. We developed and evaluated a technology designed to improve perception of vital sign information by presenting patient status as an animated patient avatar.

Methods: After step-wise improvement of the avatar, anaesthesia professionals from two hospitals participated in a comparative study of conventional monitoring. Participants observed identical monitoring scenarios via the two technologies for brief time intervals and afterwards recalled patient status.

Results: Overall, 150 anaesthesia professionals participated in the validation process and 32 participated in the comparative study, completing 128 scenarios, which allowed for 64 direct comparisons. The avatar's inter-rater reliability was high, with Fleiss' kappa of 0.98 (95% confidence interval 0.96–0.99, $P < 0.001$). With the avatar, participants recalled almost twice as many vital signs correctly as with conventional monitoring (9 vs 5, $P < 0.001$). Perceived confidence was improved (2=certain vs 1=uncertain, $P < 0.001$) and perceived workload lowered (task load index 60 vs 76, $P < 0.001$). Participants obtained these results only after watching an educational video explaining the avatar and suggesting quick learnability and potential for real-life usability.

Conclusions: This study provides empirical evidence that an animated avatar offers the opportunity to transmit vital sign information significantly more quickly than conventional monitoring and with improved confidence and reduced cognitive effort. This could help care providers gain situation awareness more efficiently.

Keywords: computer-assisted; diagnosis; patient monitoring; situation awareness

Editor's key points

- Situation awareness is critical in monitoring patient vital signs by anaesthesia professionals.
- A novel technology designed to improve perception of vital sign information, which presents patient status as an animated patient avatar, was compared with conventional monitor data in a simulation scenario.
- With the avatar, participants recalled almost twice as many vital signs correctly compared with conventional monitoring, and perceived confidence was improved and workload lowered.
- Real-life studies are required to confirm that use of an animated avatar can transmit vital sign information more quickly than conventional monitoring.

Situation awareness enables healthcare providers to correctly diagnose patient condition and make informed clinical decisions. This might help care providers to avoid errors and improve patient safety.^{1–7} By definition, it comprises three levels: 1) perception of elements in the environment within a volume of time and space; 2) comprehension of their meaning; and 3) projection of their status in the near future.² In patients connected to a patient monitor, real-time and trend information from the screen and auditory displays, specifically pulse oximeter tone and acoustic alarms, contribute significantly to care providers' mental model of patient status.^{4–6} Therefore, the WHO considers continuous patient monitoring 'extremely important' for the safety of the 313 million surgeries performed worldwide per year.^{8,9} The WHO also sees complications of anaesthesia and surgical care as a leading cause of preventable death and disability, and poor situation awareness and decision-making have been recognised as leading factors causing patient harm.^{1,8,10–14}

The typical design of current conventional monitors does not optimally enable care providers to gain an understanding of the situation in a time-efficient manner because of the way information is presented. Specifically, vital signs are presented individually as waveforms or numbers using a single sensor, single indicator philosophy. Care providers, in a time-consuming process, must read and interpret each waveform or numerical value one after the other to gain situation awareness.^{4,15,16} However, the average time anaesthesiologists spend looking at monitors is only about 2 s, and the need for several glances to detect abnormal values has been reported.^{17,18}

To optimise the information gained from checking a monitor, we developed an information transfer technology named Visual Patient. It was our goal to create an interface that transmits the current status of the vital signs to care providers as quickly as possible and with minimal cognitive effort, which has been described as the goal for successful situation awareness design.² The Visual Patient integrates information about the 11 most commonly used standard vital signs into visualisations in an animated model of the real patient (a patient avatar) using a multiple sensor multiple indicator approach. Furthermore, the technology reduces the complexity of monitoring data by processing vital signs into two or three visualisation conditions. For vital signs that have a too low, safe, or too high status (e.g. arterial pressure) the avatar can display these three visualisation conditions. For vital signs that can be categorised into an unsafe or safe status (e.g. oxygen saturation)

the avatar can display these two conditions. Among the 11 vital signs, 30 different visualisation conditions can be displayed, giving the technology the theoretical capability of displaying a total of 4608 different situations. We used an avatar because previous research has indicated that anatomical representation of information renders a design more intuitive,^{4,15,19} which is important for usability, as a significant problem with prior visualisation technologies was that they were too difficult to learn.⁴ Using anatomical representations is also in line with logical principles concluding that models should have a meaningful relationship with the reality they mirror.²⁰

We completed a comprehensive iterative development process of the avatar and afterwards compared the final version with conventional monitoring. We used a method based on the Situation Awareness Global Assessment Tool,²¹ blanking the screens after 3-s and 10-s intervals and measuring outcome measures that have been shown to be essential for situation awareness, decision-making, and performance.² We hypothesised that after 3-s and 10-s glances, the Visual Patient monitor would: 1) enable participants to perceive more vital signs; 2) improve perceived confidence; and 3) reduce perceived workload.

Methods

Validation and calibration of the avatar

Before this study, we conducted an extensive iterative development and calibration process of the avatar involving 150 individual anaesthesia professionals in two study centres. The avatar version used in the present study reached an inter-rater reliability of >94% for each of its 30 visualisation conditions and was calibrated in a Delphi process involving 11 senior anaesthesia experts to monitor adult ASA physical status 1 and 2 category patients under general anaesthesia. We report the validation and calibration process in full in the supplementary file titled Validation of the avatar. Table 1, Figure 1, Supplementary Video 1 (educational video), and the supplementary file 'Validation and calibration of the avatar' describe the avatar in detail.

Supplementary video related to this article can be found at <https://doi.org/10.1016/j.bja.2018.04.024>.

Study participants

Study participants were anaesthesiologists (staff and resident physicians) and nurse anaesthetists with a completed specialisation qualification from the anaesthesia departments of the University Hospital of Zurich and the Kantonsspital Winterthur (KSW)—two teaching hospitals in Switzerland performing approximately 30 000 and 10 000 surgical procedures per year, respectively. Participation was voluntary and the participants received no compensation. We included participants who responded to institutional e-mail invitations and asked colleagues to participate according to their availability. The ethics committee of the Canton of Zurich, Zurich, Switzerland reviewed the study protocol and issued a declaration of no-objection (BASEC-No. Req-2016-00103). All participants gave their written informed consent.

Study design

To compare the avatar with conventional monitoring, we conducted a multi-method laboratory study. The methods

Table 1 Descriptions of the visualisations of the avatar, their inter-rater reliability, and calibration according to a Delphi process conducted with experienced anaesthesia specialists described in full in the supplementary file ‘[Validation and calibration of the avatar](#).’ For example, when the input pulse rate of a patient is below 45 beats min⁻¹, the avatar’s body (=element #1) pulsates with a frequency of four per min. This pulsation frequency was perceived as ‘too low’ by 32 of 32 participants. BIS, bispectral index system; mV, millivolt; SpO₂, peripheral oxygen saturation; TOF, train of four.

Vital sign	Visualisation in the avatar	Too low	Safe	Too high
1. Pulse rate	The avatar’s body (element #1) pulsates with a slow, normal, or quick frequency. Inter-rater reliability Calibration (beats min ⁻¹):	4 Pulsations min ⁻¹ 32/32 (100%) <45	78 Pulsations min ⁻¹ 31/32 (97%) 45–100	380 Pulsations min ⁻¹ 32/32 (100%) >100
2. Arterial blood pressure	The avatar’s body (element #1) pulsates just barely, normally or very intensely, extending far beyond the ‘normal-line.’ Inter-rater reliability Calibration (mean arterial blood pressure in mm Hg):	114.5 cm ² Area of element #1 32/32 (100%) <60	225.5 cm ² Area of element #1 31/32 (97%) 60–100	586.2 cm ² Area of element #1 31/32 (97%) >100
3. Central venous pressure	The size of the avatar’s vena cava (element #3) is very small, normal, or very large, extending far beyond the white ‘normal-line.’ Inter-rater reliability Calibration (central venous pressure in mm Hg):	0 cm ² Area of element #3 32/32 (100%) <2	1.7 cm ² Area of element #3 32/32 (100%) 2–14	10.7 cm ² Area of element #3 32/32 (100%) >14
4. Respiratory rate	The avatar’s lungs (element #4) breathe with a slow, normal, or quick frequency and (if data from a CO ₂ sensor is available) the avatar synchronously exhales the CO ₂ cloud (element #5). Inter-rater reliability Calibration (breaths rate min ⁻¹):	3 Respirations min ⁻¹ 32/32 (100%)	27 Respirations min ⁻¹ 32/32 (100%)	225 Respirations min ⁻¹ 32/32 (100%)
5. Tidal volume	The avatar’s lungs extend just barely, normally or far beyond the white ‘normal-line.’ Inter-rater reliability Calibration (tidal volume in ml):	<8 1.8 cm ² Area of element #4 32/32 (100%)	8–20 4.8 cm ² Area of element #4 30/32 (94%)	>20 40.8 cm ² Area of element #4 31/32 (100%)
6. Expiratory CO ₂ concentration	The CO ₂ cloud exhaled by the avatar is just barely visible, reaches a normal extension or extends far beyond the white ‘normal-line.’ Inter-rater reliability Calibration (expiratory CO ₂ in kPa):	<400 0.05 cm ² Area of element #5 31/32 (97%)	400–650 9.3 cm ² Area of element #5 32/32 (100%)	>650 392.4 cm ² Area of element #5 32/32 (100%)
7. Body temperature	There are heatwaves rising from the avatar, or ice crystals visible on its skin. Inter-rater reliability Calibration (body temperature in °C):	<4.2 18 Ice crystals 32/32 (100%)	4.2–5.5 No indicators 32/32 (100%)	>5.5 18 Heatwaves 32/32 (100%)
8. Brain activity	The eyes (element #6) of the avatar are open or closed. Inter-rater reliability Calibration (BIS index):	<36.0 Eyes closed 32/32 (100%)	36.0–37.5 Eyes intermediate 32/32 (100%)	>37.5 Eyes open 32/32 (100%)
9. Peripheral oxygen saturation	The avatar has a healthy skin colour or a purple skin colour. Inter-rater reliability Calibration (SpO ₂ %):	<40 Purple skin 32/32 (100%)	40–50 Normal skin colour 32/32 (100%)	>50 93–100
10. Neuromuscular function	The avatar has extended extremities and extended thumbs or the extremities appear floppy. Inter-rater reliability Calibration (TOF ratio %):	<93 Relaxed body 31/32 (97%)	No relaxation 32/32 (100%)	
11. Electrocardiography ST-segment	The heart muscle of the avatar has a homogenous vivid red colour or a dark purple colour. Inter-rater reliability Calibration (ST segment deviation in mV):	<90 Myocardial hypoxia indicator present 32/32 (100%)	>90 Myocardial hypoxia indicator absent 32/32 (100%)	
		<–0.2, >0.2	–0.2 to +0.2	

applied consisted of the assessment of participants' answers to various research questions after watching scenarios with both technologies. All data were entered using an iPad- (Apple Inc., Cupertino, CA, USA) and iSurvey-based (Harvest Your Data, Wellington, New Zealand) data collection tool.²² [Supplementary Figure S1](#) provides screenshots of this data collection tool.

We collected data with participants sitting in front of an Aspire V15 Nitro laptop computer (ACER, Inc., Taipei, Taiwan) inside a quiet room at the two study centres; the only other person present was the data collector. Before the data collection sessions, participants completed a survey about personal information and, to obtain a standardised method of instruction, watched a 6-min educational video ([Supplementary Video 1](#)) explaining the avatar. This video and all others in this study were shown in ultra-high resolution (3840×2160 pixels) at 60 frames per second on the 15-inch screen of the laptop.

After the educational video, participants were familiarised with the interface of the conventional monitor, a simulation emulating a GE Datex Ohmeda monitor (General Electric Company, Boston, MA, USA) recorded using the app SimMon (Castle 2 Andersen ApS Frederiksværksgade 30 3400 Hillerød, Denmark). [Figure 1b](#) presents a picture of the study setup and the conventional monitor layout. Each participant rated four videos in sequence. These videos consisted of one 3-s and one 10-s scenario, each of which was shown twice, once with either technology. These scenarios were taken from a pool of four total scenarios, which are outlined in [Supplementary Table S1](#). To blind participants to the fact that they rated the same scenarios twice, videos were shown in interchanging order, beginning with a random first scenario. [Supplementary Figure S2](#) provides a flowchart explaining this procedure.

The vital signs in the scenarios were randomly created. This randomisation was ideal for measuring vital sign perception without interference because it makes it impossible to guess the status of vitals based on the state of others, as would be possible with pattern recognition in scenarios designed to measure higher levels of situation awareness.

The Visual Patient videos showed only the avatar without any numbers or waveforms, and the conventional videos showed only the conventional interface with the usual numbers and waveforms ([Fig. 1b](#)). The conventional scenarios included a standard audio signal with frequency and pitch for heart rate and oxygen saturation, respectively, and contained only unequivocally safe or unsafe values. The scenarios were pre-recorded and played back as videos to enable equal study conditions for all participants. After the 3-s and 10-s intervals, the screen of the laptop went black and participants were asked to answer the questions about the scenario on the iPad. We chose the 3-s time interval because Ford and colleagues¹⁷ reported anaesthetists' median look duration at a patient monitor to be about 2 s. As a second interval, we chose 10 s because, in real-life measurements, we found it to be the maximum time clinicians take to scan a monitor during patient care.

Outcome measures

Vital sign perception

To compare the number of vital signs that could be perceived correctly with both technologies, participants were asked to recall the status of each vital sign after each video.

The choices were four nominal scale-level measurement options: 'too low', 'safe', 'too high', or 'no recall'; 11 was the

maximum score, 0 the minimum. The iPad asked the vital signs in randomised order on a single screen to minimise interference from ordering effects and the limits of working memory.

The rationale for measuring vital sign perception was that it forms an integral part of anaesthesia providers' situation awareness level I, the first of the three hierarchical levels of situation awareness. Situation awareness errors have been reported to be responsible for up to 80% of anaesthesia incidents and level I failures have been identified as the largest single category of situation awareness errors in anaesthesia with 42%¹² and 38%,¹³ respectively.

Confidence

To compare the participants' confidence in their vital sign perception with both technologies, they were asked to indicate their certainty that their assessment is correct for each vital sign after each video. Choices were four nominal scale-level measurement options: from 0='very uncertain' to 3='very certain'. For each participant and scenario, a median with interquartile range was calculated.

The rationale for measuring confidence was that uncertainty is a psychological stress factor and situation awareness and confidence work together for a person to decide. Research has shown that if situation awareness is good and confidence in that situation awareness is high, a person is more likely to achieve a good outcome. If with equally good situation awareness, the person has a low level of confidence in that situation awareness, they will most likely not act on it, choosing to gather more information or behave protectively and thus be ineffective.^{2,23}

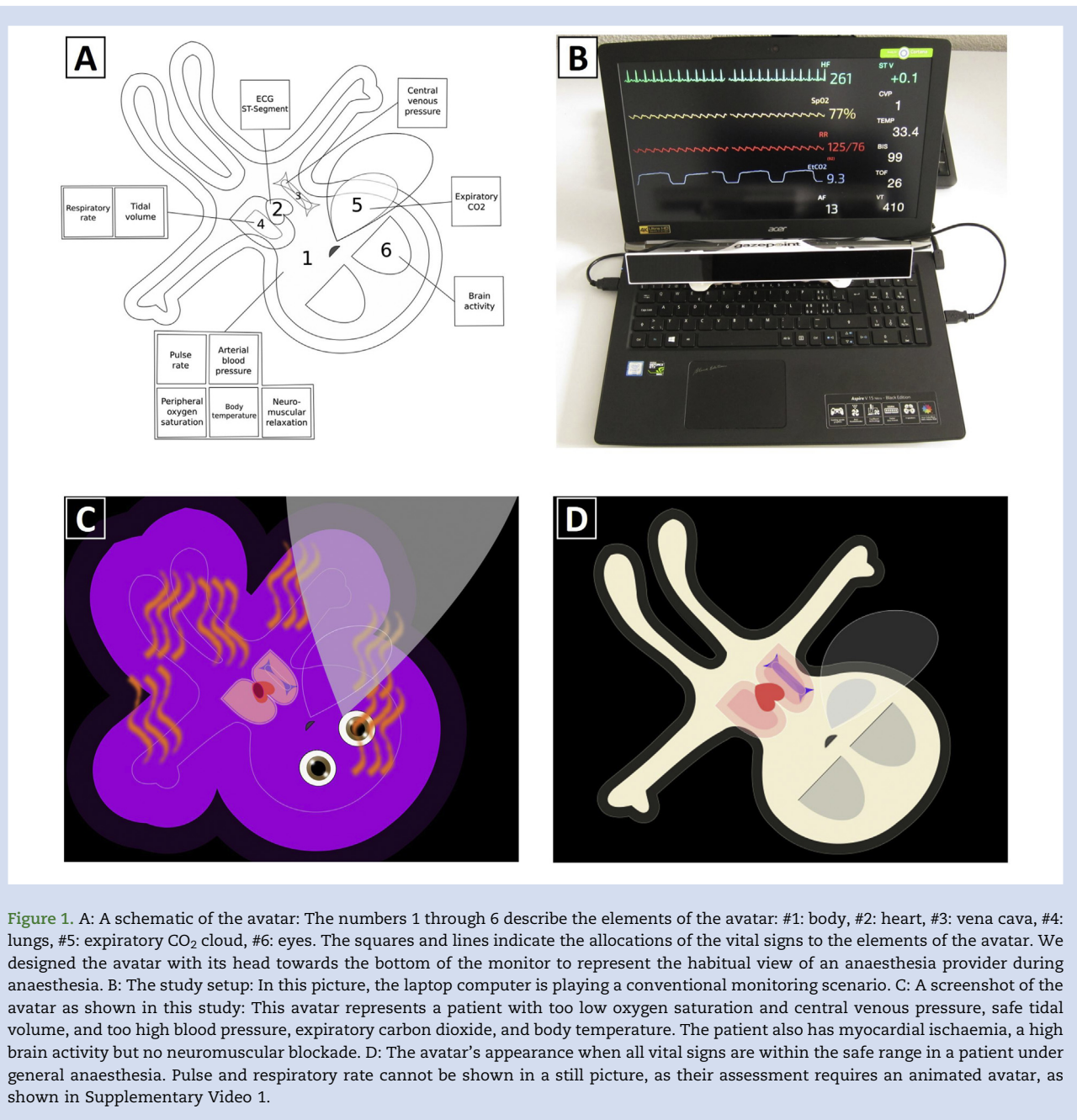
Workload

To compare the perceived workload with both technologies, participants completed the NASA Task Load Index (TLX) questionnaire after each scenario to provide a validated measure of perceived workload.²⁴ Generally, the NASA-TLX score has six questions, but the question regarding physical demand can be removed if a task does not present physical demand,²⁵ as in this study. The raw NASA-TLX score was calculated from the ratings that participants gave to its five questions on scales from 0 to 100 in intervals of 5. Maximum and minimum scores were 100 and 0, respectively.

The rationale for measuring perceived workload was that high cognitive workload has been described as a factor reducing capacity to process information and thereby the generation of situation awareness,³⁴ and subsequently decision-making²⁶ and task performance.^{27,28} Stress causes people to pay less attention to peripheral information, become more disorganised in scanning information, succumb to attentional tunnelling, and deciding without taking into account all available information, termed premature closure.²

Statistical analysis

We used two-tailed paired t-tests to compare correctly perceived vital signs, perceived confidence, and workload with either technology. To investigate whether prior experience with the avatar affected the results, we performed a subgroup analysis comparing the results of the participants who had their first contact with the avatar during the present study with the results of the participants who had also taken part in the validation process. For this analysis, we used Mann-Whitney U-tests. A P-value <0.05 was considered statistically significant.



The sample size for this study was based on an *a priori* power analysis for a paired t-test. Assuming a standard deviation of differences of 0.81, as determined by a pilot study, eight participants per scenario were required to detect a difference of one correctly recalled vital at the 0.05 significance level with a power of 80%.

Results

Table 2 outlines the study and participant characteristics. There were 32 anaesthesia professionals who each rated two scenarios twice, once with Visual Patient and once with conventional monitoring, resulting in 128 rated scenarios, which allowed for 64 direct comparisons.

Using the avatar significantly improved all three outcome measures (vital sign perception, perceived confidence, and workload) in all scenarios. Except for a single participant, in only one of the 64 comparisons, all participants recalled more vital signs using the avatar.

In both 3-s Visual Patient scenarios, participants perceived more vital signs and, at the same time, rated workload lower than in both 10-s conventional scenarios, suggesting that 3 s of avatar monitoring transferred more information about vital sign status with lower perceived workload than 10 s of conventional monitoring. Figures 2–4 present the results for vital sign perception, perceived confidence, and workload for all scenarios on an individual participant level. Supplementary Figure S3 shows the results on a scenario level.

Table 2 Study and participant characteristics. Visual Patient naïve participants=participants who had no prior contact with the avatar during the iterative development process. GE, General Electric Company, Boston, MA, USA; IQR, interquartile range; VP, Visual Patient.

Study centre	University Hospital Zurich	Kantonsspital Winterthur	Total
Duration of study (days)	20	2	22
VP naïve participants, n/N (%)	0/16 (0)	15/16 (94)	15/32 (47)
Participants, N	16	16	32
Scenarios rated, n	64	64	128
Direct comparisons, n	32	32	64
Senior staff members and staff members, n (%)	6 (37)	6 (37)	12 (37)
Residents, n (%)	4 (25)	4 (25)	8 (25)
Nurse anaesthetists, n (%)	6 (37)	6 (37)	12 (37)
Female/male participants, n (%)	7 (44)/9 (56)	10 (62)/6 (37)	17 (53)/15 (47)
Median (IQR) age group of participants (yr)	25–34 (25–34 to 35–44)	35–44 (25–34 to 45–54)	35–44 (25–4 to 45–54)
Median (IQR) anaesthesia experience group of participants (yr)	5–10 (5–10 to >10)	1–5 (1–5 to 5–10)	5–10 (1–5 to >10)
Median (IQR) experience with conventional patient monitors from different monitoring manufacturers	2 (2–3)	2 (1–4)	2 (1–3)
Manufacturer participants reported most experience with	Dräger Medical (16/16)	GE Healthcare (12/16)	Dräger Medical (26/32)
Median (IQR) duration of data collection session (min)	32 (28–35)	35 (32–41)	33 (30–39)

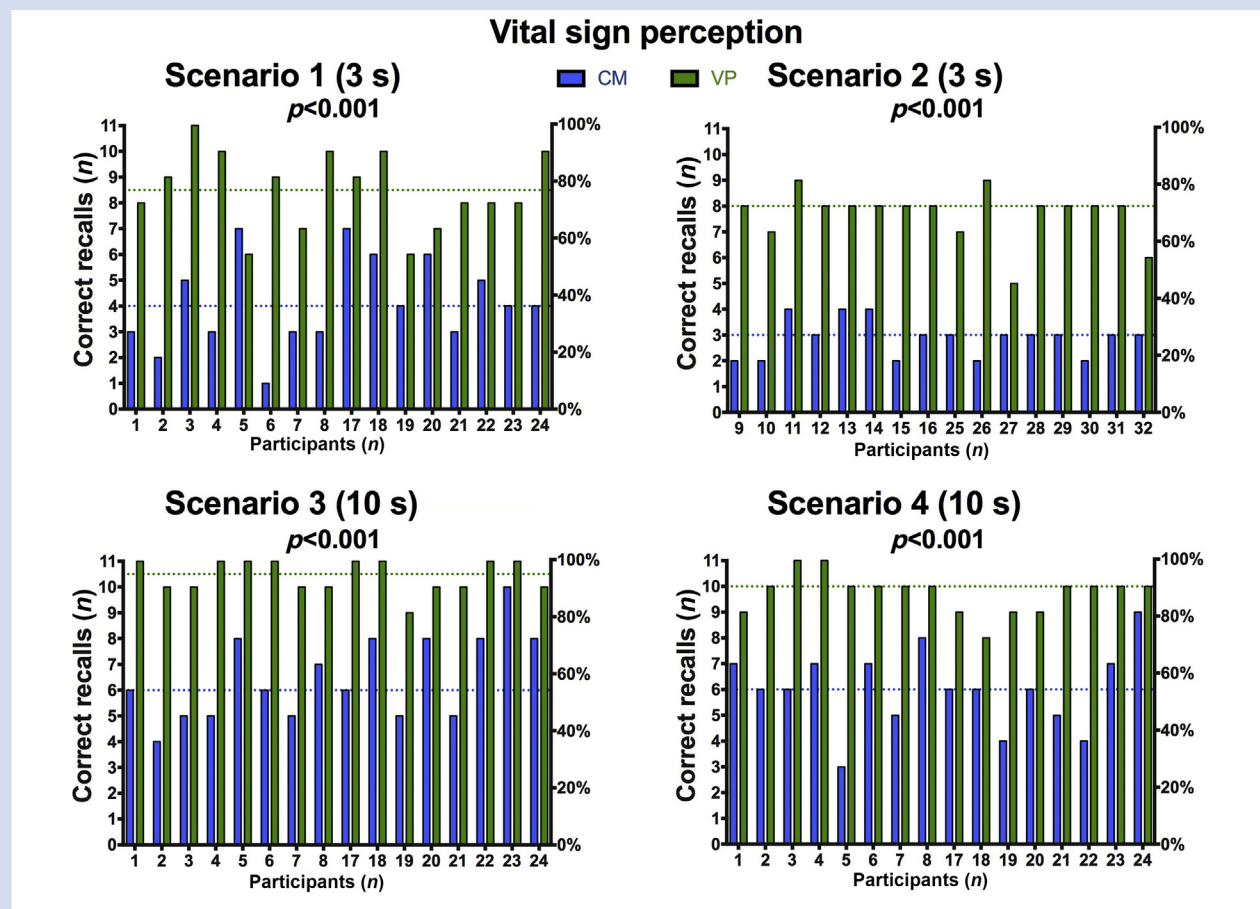


Figure 2. Visual Patient (VP) compared to conventional monitoring (CM): Correctly recalled vital signs on an individual participant level. N for each scenario was 16. The dotted lines indicate the medians. Participants 1–8 (USZ) and 17–24 (KSW) rated scenarios 1 and 3, participants 9–16 (USZ) and 25–32 (KSW) rated scenarios 2 and 4. USZ = University Hospital Zurich (study centre 1), KSW = Kantonsspital Winterthur (study centre 2).

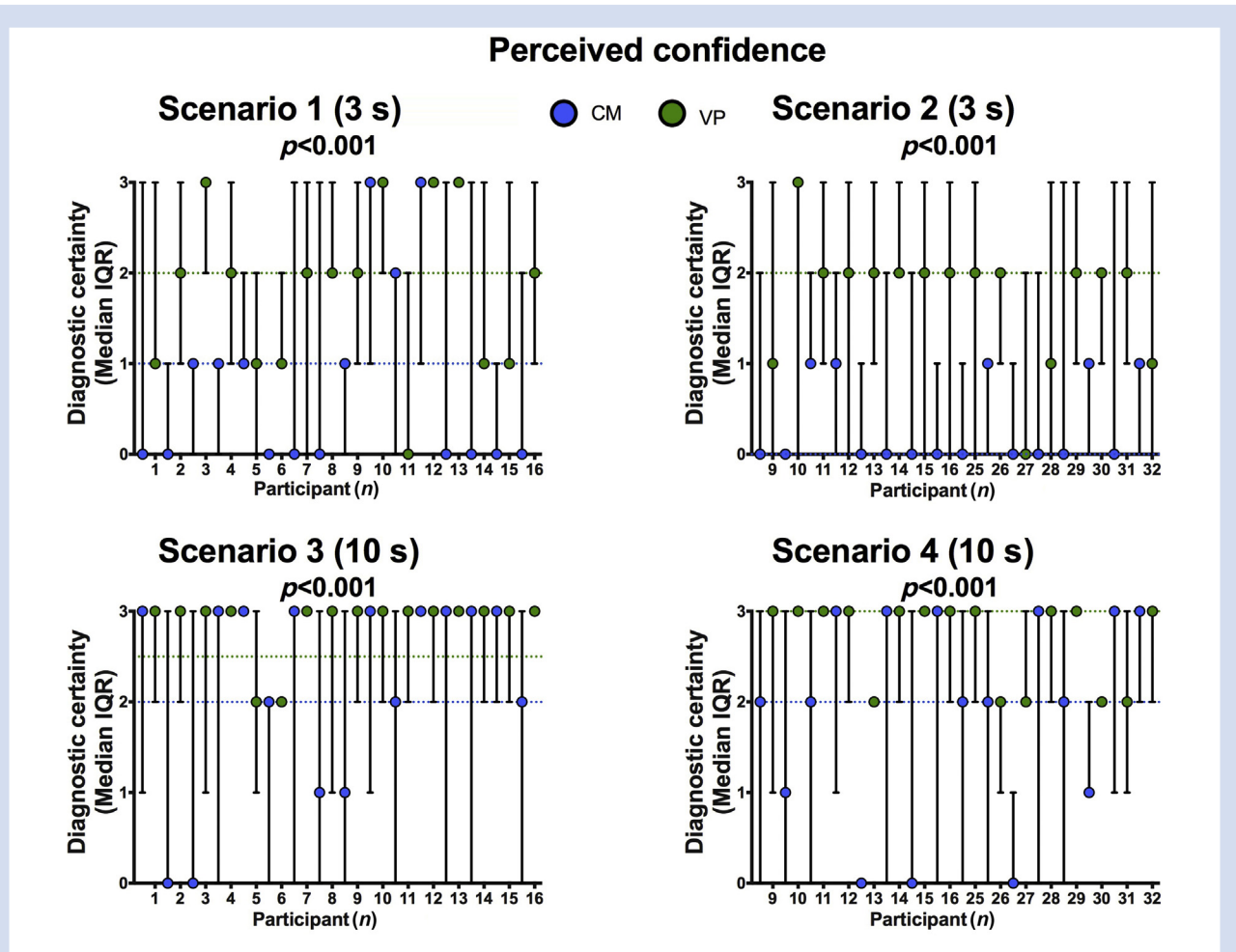


Figure 3. Visual Patient (VP) compared to conventional monitoring (CM): Perceived confidence on an individual participant level. N for each scenario was 16. The dotted lines indicate the medians. Diagnostic certainty: 0 = very uncertain, 1 = uncertain, 2 = certain, 3 = very certain. Participants 1–8 (USZ) and 17–24 (KSW) rated scenarios 1 and 3, participants 9–16 (USZ) and 25–32 (KSW) rated scenarios 2 and 4. USZ = University Hospital Zurich (study centre 1), KSW = Kantonsspital Winterthur (study centre 2).

There were no significant differences in the results of participants who had their first contact with the avatar during the present study and participants who had been introduced to the Visual Patient technology during the validation process already. [Supplementary Table S2](#) outlines the results of the subgroup analysis.

Discussion

We compared a newly developed patient monitoring information transfer technology that creates an avatar of the monitored patient with conventional patient monitoring. We found that during both longer 10-s and shorter 3-s glances, which anaesthesia providers use to scan patient monitors in real-life patient care, the Visual Patient technology almost doubled the number of perceived vital signs, improved care provider confidence, and reduced perceived workload. Participants achieved these results after only watching an educational video explaining the avatar, which suggests quick learnability and potential for real-life usability.

Improving the number of vital signs that care providers can perceive (situation awareness level I) when scanning a

monitor is relevant because integrated with other information sources (e.g. feeling the patient's pulse, looking at the surgical field, etc.) it helps care providers develop a mental model of the patient's situation.^{4–6} Improving the perception of critical information has been associated with improved patient outcomes.^{29–31}

Improving confidence in vital sign perceptions that are correct and reducing perceived workload are relevant because uncertainty and workload are psychological stress factors that bind cognitive resources which could otherwise be used to gain higher levels of situation awareness and make decisions. The benefits of the new monitor would likely be highest when cognitive resources are decreased (e.g. in high-workload situations such as emergency care, fatigued or less experienced care providers).^{14,32–34}

The basis for the participants' impressive performance using the avatar was most likely a combination of several design factors. We based the development of the avatar on recommendations for a successful situation awareness interface as described by Endsley and Jones,² and applied their principles regarding user-centred design. The avatar's three main design features are direct, anatomical

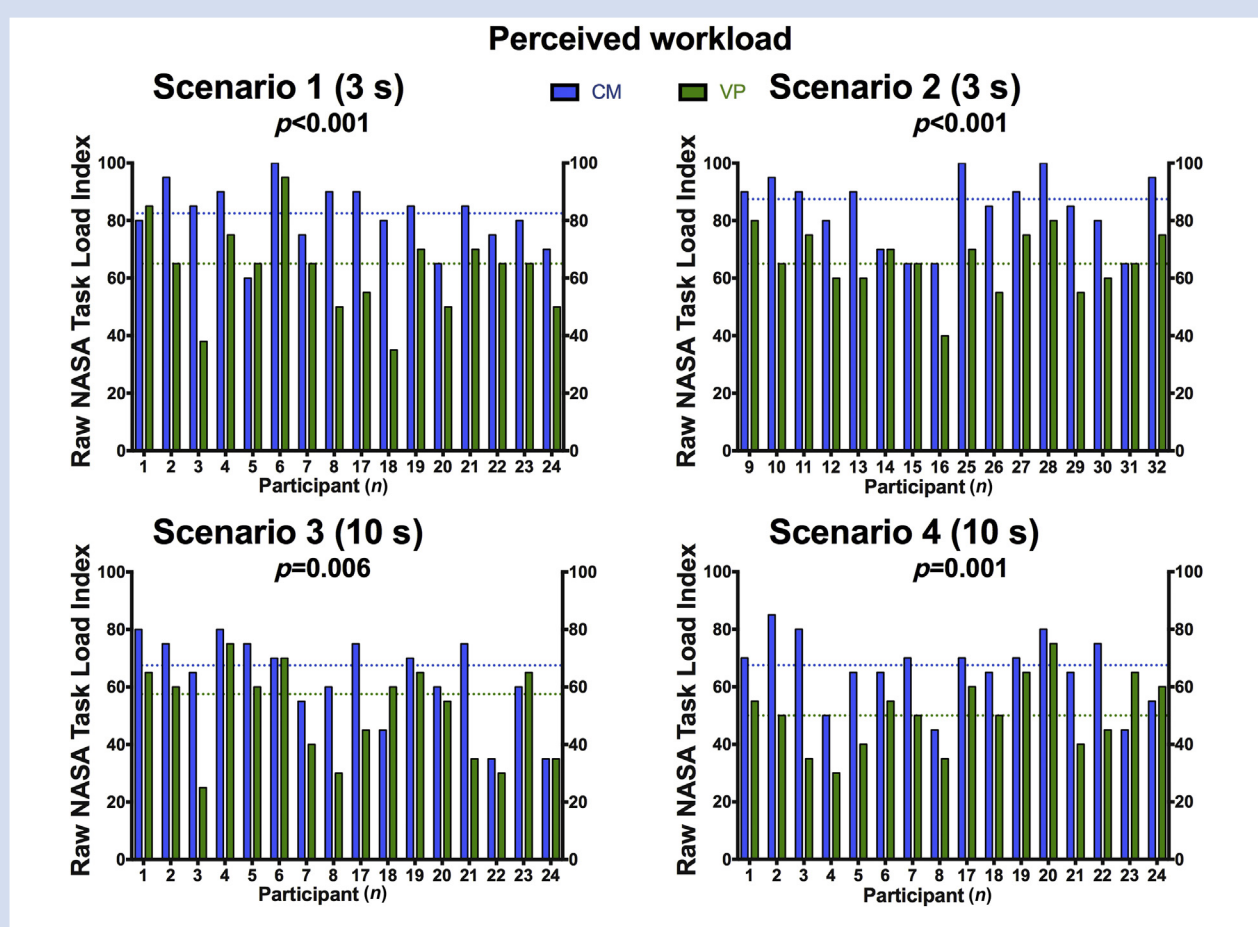


Figure 4. Visual Patient (VP) compared to conventional monitoring (CM): Perceived workload on an individual participant level. N for each scenario was 16. The dotted lines indicate the medians. Participants 1–8 (USZ) and 17–24 (KSW) rated scenarios 1 and 3, participants 9–16 (USZ) and 25–32 (KSW) rated scenarios 2 and 4. USZ = University Hospital Zurich (study centre 1), KSW = Kantonsspital Winterthur (study centre 2).

representation of the vital sign information, multiple sensors and multiple indicators approach, and reduction of information complexity. The avatar first translates vital sign data into information (i.e. a parameter is too high or too low vs a number). It then presents this information in several inter-rater reliable and easy-to-learn anatomical, direct visualisations (i.e. visualisations modelled after the phenomena that generate the data, e.g. high heart rate=quick pulsation of the body, low oxygen saturation=cyanotic skin colour). These design characteristics eliminate the need for the care providers to calculate information from lower level data.^{2,35} When a care provider observes an avatar with a quick and intense pulsation, cyanotic skin colour, floppy limbs, and snowflakes, they can diagnose at a single glance that the patient has tachycardia, hypertension, desaturation, hypothermia, and neuromuscular block. To obtain the same 'picture' using conventional monitoring, the care provider needs to sequentially visually fixate and mentally categorise the meaning of five individual waveforms or numerical values on various positions of the screen.

While the results of this study indicate several advantages of Visual Patient monitoring over conventional monitoring,

the conventional interface retains some advantages over the avatar. Specifically, reducing the complexity of information to ease understanding, as accomplished by the avatar, can reduce its level of precision,³⁵ and detection of a deviation in the avatar might have to be followed up in its precise numerical form on the conventional monitor. Furthermore, unlike conventional monitoring, the current version of the avatar cannot display trends, which are often early warning signs before vital signs are too high or too low. Therefore, we envision the avatar as a complementary technology to conventional patient monitoring, its purpose being to support care providers to perceive critical information already available in the system more efficiently.

Limitations

This study has limitations. We collected data in a laboratory setting using separately presented and simulated avatar and conventional monitoring scenarios. Therefore, it remains unclear whether the improvements found would increase or decrease when the two technologies are presented together in real-life situations (e.g. in an operating room), where

additional factors, such as team communication,^{14,36} checklist use,^{29–31,37,38} and various alarm systems can impact the perception of vital sign information.

We focused on information perception (level I situation awareness) and factors affecting situation awareness, but did not study the effects of the avatar on higher levels of situation awareness and decision-making. However, the three levels of situation awareness, decision-making, and performance are interconnected, and the most frequent situation awareness problems are level I errors, where available information is not recognised.^{12,13} Further studies investigating whether the results expand to higher situation awareness levels are warranted, especially under high cognitive workload, when situation awareness levels II and III errors occur because of working memory limitations. Such a study should ideally be performed as a high-fidelity simulation accounting for the additional factors outlined above.

The calibration we used adapts the avatar for monitoring adult ASA 1 and 2 patients within a tight range, determined safe by experienced anaesthesia experts. Other calibrations (e.g. infants and paediatric patients, and those with critical illness) are necessary and will follow.

The possibility of selection bias cannot be excluded because selection of participants was not randomised and depended on their response to institutional e-mail invitations. However, this was a two-centre study, participants were gender-, profession-, and experience-balanced, and we presented the identical scenarios in Visual Patient and conventional monitoring to each participant, enabling a direct comparison of the two technologies.

Conclusions

We showed that an animated avatar provides an opportunity to transmit the status of a patient's vital signs to care providers more quickly, and with improved confidence and reduced cognitive effort than conventional monitoring. Such an avatar might serve as a valuable situation awareness add-on to conventional monitoring.

Authors' contributions

Study design: D.W.T., L.H., P.N., M.W., B.S., D.R.S., C.B.N.
 Evaluated original study data: D.W.T., L.H., B.S., C.B.N.
 Data collection: L.H., C.B.N.
 Data analysis: D.W.T., L.H., B.S., D.R.S., C.B.N.
 Drafted the article and approved the final version: D.W.T., L.H., P.N., M.W., B.S., D.R.S., C.B.N.
 Archiving study files: D.W.T.

Collaborator contributions

Organisation of data collection at KSW: M.T.G.

Acknowledgements

The authors thank I. Barshi, Human Systems Integration Division, NASA Ames Research Center, Mountain View, CA, USA, F. Drews, Department of Psychology, University of Utah, Salt Lake City, UT, USA, N. Bienefeld-Seall, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland, and A. Kaserer, University and University Hospital Zurich for their contributions to this project and all participants at the two study sites.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.bja.2018.04.024>.

Declarations of interest

The University of Zurich owns the intellectual property rights to the technology described in this manuscript and registered 'Visual Patient' as a trademark. As inventors, the authors D.W.T., P.N., and C.B.N. may receive royalties in the event of commercialisation.

Funding

Institute of Anaesthesiology, University and University Hospital Zurich: Institutional funding. University of Zurich: Proof of concept funding (UZ16/288POC).

Collaborators

M.T.Ganter.

References

1. World Health Organization Patient Safety Group. *Human factors in patient safety: review of topics and tools*, Geneva. 2009
2. Endsley MR, Jones DG. *Designing for situation awareness: an approach to user-centered design*. Second Edn. Boca Raton, FL, USA: CRC Press, Inc.; 2011
3. Endsley MR. Toward a theory of situation awareness in dynamic systems. *Hum Factor*. 1995; **37**: 32–64
4. Schulz CM, Endsley MR, Kochs EF, Gelb AW, Wagner KJ. Situation awareness in anesthesia: concept and research. *Anesthesiology* 2013; **118**: 729–42
5. Gaba DM, Howard SK, Small SD. Situation awareness in anesthesiology. *Hum Factor*. 1995; **37**: 20–31
6. Fioratou E, Flin R, Glavin R, Patey R. Beyond monitoring: distributed situation awareness in anaesthesia. *Br J Anaesth* 2010; **105**: 83–90
7. Reason J. Human error: models and management. *BMJ* 2000; **320**: 768–70
8. World health organization guidelines for safe surgery 2009: *safe surgery saves lives*. Geneva. 2009
9. Weiser TG, Haynes AB, Molina G, et al. Size and distribution of the global volume of surgery in 2012. *Bull World Health Organ* 2016; **94**: 201–9F
10. Caplan RA, Vistica MF, Posner KL, Cheney FW. Adverse anesthetic outcomes arising from gas delivery equipment: a closed claims analysis. *Anesthesiology* 1997; **87**: 741–8
11. Way LW, Stewart L, Gantert W, et al. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Ann Surg* 2003; **237**: 460–9
12. Schulz CM, Burden A, Posner KL, et al. Frequency and type of situational awareness errors contributing to death and brain damage: a closed claims analysis. *Anesthesiology* 2017; **127**: 326–37
13. Schulz CM, Krautheim V, Hackemann A, Kreuzer M, Kochs EF, Wagner KJ. Situation awareness errors in anesthesia and critical care in 200 cases of a critical incident reporting system. *BMC Anesthesiol* 2015; **16**: 4

14. Helmreich RL. On error management: lessons from aviation. *BMJ* 2000; **320**: 781–5
15. Drews FA, Westenskow DR. The right picture is worth a thousand numbers: data displays in anesthesia. *Hum Factor*. 2006; **48**: 59–71
16. Rayner K. Eye movements in reading and information processing: 20 years of research. *Psychol Bull* 1998; **124**: 372–422
17. Ford S, Birmingham E, King A, Lim J, Ansermino JM. At-a-glance monitoring: covert observations of anesthesiologists in the operating room. *Anesth Analg* 2010; **111**: 653–8
18. Loeb RG. Monitor surveillance and vigilance of anesthesia residents. *Anesthesiology* 1994; **80**: 527–33
19. Wachter SB, Agutter J, Syroid N, Drews F, Weinger MB, Westenskow D. The employment of an iterative design process to develop a pulmonary graphical display. *JAMIA* 2003; **10**: 363–72
20. Wittgenstein L. *Tractatus logico-philosophicus*. London: Routledge & Kegan Paul; 1922
21. Endsley MR. *Direct measurement of situation awareness: validity and use of SAGAT*. Situation awareness analysis and measurement. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 2000. p. 147–73
22. Tscholl DW, Weiss M, Spahn DR, Noethiger CB. How to conduct multimethod field studies in the operating room: the ipad combined with a survey app as a valid and reliable data collection tool. *JMIR Res Protoc* 2016; **5**: e4
23. Endsley MR, Jones WM. *Situation awareness information dominance & information warfare*. Wright-Patterson AFB OH 45433–47022. 1997
24. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv Psychol* 1988; **52**: 139–83
25. Hart SG. Nasa-task load index (NASA-TLX); 20 years later. *Proc Hum Factor. Ergon Soc Annu Meet* 2006; **50**: 904–8
26. Fothergill S, Neal A. The effect of workload on conflict decision making strategies in air traffic control. *Proc Hum Factor. Ergon Soc Annu Meet* 2008; **52**: 39–43
27. Prottengeier J, Petzoldt M, Jess N, et al. The effect of a standardised source of divided attention in airway management: a randomised, crossover, interventional manikin study. *Eur J Anaesthesiol* 2016; **33**: 195–203
28. McKinley J, Dempster M, Gormley GJ. ‘Sorry, I meant the patient’s left side’: impact of distraction on left–right discrimination. *Med Educ* 2015; **49**: 427–35
29. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**: 491–9
30. de Vries EN, Prins HA, Crolla RM, et al. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med* 2010; **363**: 1928–37
31. Starmer AJ, Spector ND, Srivastava R, et al. Changes in medical errors after implementation of a handoff program. *N Engl J Med* 2014; **371**: 1803–12
32. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns’ weekly work hours on sleep and attentional failures. *N Engl J Med* 2004; **351**: 1829–37
33. Dawson D, Reid K. Fatigue, alcohol and performance impairment. *Nature* 1997; **388**: 235
34. Wallace DJ, Angu DC, Barnato AE, Kramer AA, Kahn JM. Night time intensivist staffing and mortality among critically ill patients. *N Engl J Med* 2012; **366**: 2093–101
35. Degani A, Jorgensen CC, Iverson D, Shafto M, Olson L. *On organization of information: approach and early work*. Moffett Field, CA, USA: NASA Ames Research Center; 2009
36. Kolbe M, Burtscher MJ, Wacker J, et al. Speaking up is related to better team performance in simulated anesthesia inductions: an observational study. *Anesth Analg* 2012; **115**: 1099–108
37. Arriaga AF, Bader AM, Wong JM, et al. Simulation-based trial of surgical-crisis checklists. *N Engl J Med* 2013; **368**: 246–53
38. Tscholl DW, Weiss M, Kolbe M, et al. An anesthesia pre-induction checklist to improve information exchange, knowledge of critical information, perception of safety, and possibly perception of teamwork in anesthesia teams. *Anesth Analg* 2015; **121**: 948–56

Handling editor: H.C. Hemmings Jr