ORIGINAL ARTICLE

Cognitive load associations when utilizing auditory display within image-guided neurosurgery

Joseph Plazak[1] · Daniel A. DiGiovanni[2] · D. Louis Collins[2] · Marta Kersten-Oertel[1]

Abstract

Purpose The combination of data visualization and auditory display (e.g., sonification) has been shown to increase accuracy, and reduce perceived difficulty, within 3D navigation tasks. While accuracy within such tasks can be measured in real time, subjective impressions about the difficulty of a task are more elusive to obtain. Prior work utilizing electrophysiology (EEG) has found robust support that cognitive load and working memory can be monitored in real time using EEG data.

Methods In this study, we replicated a 3D navigation task (within the context of image-guided surgery) while recording data pertaining to participants' cognitive load through the use of EEG relative alpha-band weighting data. Specifically, 13 subjects navigated a tracked surgical tool to randomly placed 3D virtual locations on a CT cerebral angiography volume while being aided by visual, aural, or both visual and aural feedback. During the study EEG data were captured from the participants, and after the study a NASA TLX questionnaire was filled out by the subjects. In addition to replicating an existing experimental design on auditory display within image-guided neurosurgery, our primary aim sought to determine whether EEG-based markers of cognitive load mirrored subjective ratings of task difficulty

Results Similar to existing literature, our study found evidence consistent with the hypothesis that auditory display can increase the accuracy of navigating to a specified target. We also found significant differences in cognitive working load across different feedback modalities, but none of which supported the experiments hypotheses. Finally, we found mixed results regarding the relationship between real-time measurements of cognitive workload and a posteriori subjective impressions of task difficulty.

Conclusions Although we did not find a significant correlation between the subjective and physiological measurements, differences in cognitive working load were found. As well, our study further supports the use of auditory display in image-guided surgery.

✉ Marta Kersten-Oertel
marta@ap-lab.ca

1 Gina Cody School of Engineering and Computer Science, Concordia University, EV 3.301, 1455 De Maisonneuve Blvd. W., Montreal, QC H3G 1M8, Canada

2 Montreal Neurological Institute, McGill University, Montreal, QC, Canada

Introduction

Image-guided neurosurgery, or neuronavigation, has been used to visualize the location of surgical tools by tracking and mapping their locations with respect to preoperative models of a patient's anatomy. A common limitation of image-guided neurosurgery is that it requires the surgeon to regularly divert attention away from the patient and toward the neuronavigation system. While there are many potential solutions to this problem, here, we focus on the use of auditory display as a means to mitigate this particular limitation. Auditory displays use sound to communicate information from a computer system to a user. According to Hermann [1] auditory displays use sonification as a means of generating sound signals based

on data input. These sound signals may represent earcons (i.e. distinctive sounds that represent a specific event), audification (i.e. time series data mapped to acoustic pressure), voice messaging and sonification. In this work, we use the traditional definition of sonification which according to Kramer et al. [2] refers to the mapping between data relations (e.g. distances between a surgical probe and target) and an acoustic signal (e.g. sound representation of this distance) for the purpose of facilitating the interpretation of navigating within a neurosurgical volume. More specifically, this paper addresses the methods by which we can assess auditory display as providing a meaningful benefit for clinicians.

Data sonification, a specific type of auditory display, offers unique affordances, relative to data visualization, predominantly because the ear can successfully parse and attend to multiple data streams at once, a phenomenon often referred to as the cocktail party effect [3]. Prior research has investigated the feasibility of utilizing sonified data streams to reduce clinician dependence on visual information [4,5], thereby increasing available attention for other tasks [6] and increasing accuracy within 3D localization tasks [7]. One specific application for sonified data within image-guided surgery systems includes providing information pertaining to depth, a feature that is difficult to visualize on a two-dimensional screen. In this example, auditory display provides continuous and immediate feedback that is not available from a visual display. It bears emphasizing that this particular type of auditory display should not be confused with auditory distance control, which relates to a specific area of research concerned with sound localization [8]. In researching the feasibility of incorporating auditory display within existing systems, it is necessary to have proper metrics to quantify the costs and benefits of each particular implementation; the most common metrics include task-related performance measures (e.g. time or accuracy measurements) and/or subjective judgments (e.g. Likert scale ratings).

While accuracy (e.g. average distance to target, time to reach target, etc.) within such tasks can be measured in real time, subjective impressions about the difficulty of a task are more elusive to obtain. Indeed to evaluate the benefit of different visualization methods, interfaces, simulators, etc., researchers most often perform user studies that look at task-related data such as timing and accuracy in addition to post-experiment surveys on subjective impressions of task difficulty, mental effort, and cognitive workload [9]. Studies have shown, however, that subjective ratings on perceived effort or cognitive workload using such questionnaires as the NASA TLX [10] are not necessarily the best evaluation techniques as they do not always correlate well with task performance measures (e.g. [11,12]). Therefore, measures of workload based on physiological data may be more suitable [13].

Our aim was to utilize electroencephalogram (EEG) recording, as a physiological measure, to monitor participants' cognitive workload, specifically via relative frontal alpha-band activity, while performing 3D neuronavigation tasks. Prior work utilizing EEG has found robust support that cognitive load, working memory, and attention can be monitored in real time using EEG data. For example, in [9], EEG monitoring and response time were used to quantitatively evaluate the effectiveness of different visualization techniques in terms of the burden they place on a viewer's cognitive load.

While several measures of EEG signal related to workload exist, we elected to focus on frontal alpha-band activity, similar to other studies [9,14,15], which has been identified as reflecting cognitive and memory performance [16–18]. In particular, we utilized a specific EEG metric which calculated the relative alpha activity by dividing the absolute linear-scale alpha-band power over the sum of the absolute linear-scale powers in the beta, delta, gamma, and theta bands [19]. Recent work has shown that increases in alpha-band activity reflect the active inhibition of task-irrelevant information, and therefore it is a good indicator of cognitive load, but these changes in alpha activity are coupled to other frequencies and vice versa [20]. Therefore, we elected to extract changes in alpha as it increases relative to other bands to account for this, allowing us to use values that are indicative of real increases in alpha-band activity. We believe this to be the best EEG measure of cognitive workload, as increases in alpha-band activity may co-occur with other large-scale changes in power among the different frequency bands [20].

In order to examine cognitive load associations when utilizing auditory display within image-guided neurosurgery, we designed an experiment utilizing a 3D localization task with auditory, visual, and audio-visual feedback. A priori, and consistent with existing literature, we hypothesized that audio information would improve the ability to locate points within a 3D volume when combined with visual information. Further, we hypothesized that tasks using audio-visual feedback would require less cognitive workload (via both subjective and physiological measurements) than utilizing a singular modality for providing feedback. Finally, we hypothesized that subjective and physiological measurements would be positively correlated.

Methods

In order to interface a neuronavigation system (i.e., IBIS [21]) with an external sound synthesizer, we developed an audio plug-in capable of transmitting Open Sound Control (OSC) messages from an open-source neuronavigation system to an open-source audio programming environment [22]. This plug-in employed "oscpack" [23] for handling OSC packet

manipulation. The audio portion of our system received OSC messages from IBIS, which provided real-time information on the location of a marked surgical probe, the location of a target point within a 3D volume of a computed tomography angiography (CTA) brain scan, and several other messages to help facilitate the control of the experiment (detailed below).

Previous research has utilized a wide variety of sonification types within auditory display (see Black et al. [6] for a review) for the purpose of presenting a continuous stream of information. These sonfications include speech prompts, abstract alarms, relatively abstract sound patterns, and even music-like presentations. In this study, we utilized a simple signal-to-noise manipulation that has been demonstrated to be one of the most effective audio cues in our previous study [24]. This particular sonification involved mapping the real-time Euclidean distance (in three dimensions) between the tip of the surgical probe and a specified target within the volume to the real-time mix of two independent audio signals: a 400 Hz sine tone and white noise. Beyond a preset distance threshold (300 mm), this specific sonification presented white noise, and as the surgical tool approached the target, the sonification faded linearly to a steady-state sine tone. At a distance of 150 mm, the signal consisted of 50% noise and 50% sine tone. Participants were able to determine that they had arrived at the target when the white noise signal was completely absent and only the pure sine tone remained.

Procedure

In a controlled pilot experiment, 13 non-expert participants (4 females) who were medical imaging or surgical technology researchers and familiar with neuronavigation navigated a tracked surgical tool to randomly placed 3D virtual locations while being aided by visual, aural, or both visual and aural feedback; both the 3D virtual locations and feedback method were randomly changed for every trial. Visual feedback included a bright yellow point placed on a vessel within a 3D computed tomography angiography (CTA) volume, as well as real-time mapping of the surgical pointer's tip position onto the same surgical volume (as seen in Fig. 1). As described above, audio feedback consisted of a continuous "signal-to-noise" sonification which resulted in a linear transformation from white noise to a pure sine tone as participants approached the target. Audio-visual feedback simply combined the two types of feedback described above.

Participants each completed 24 trials, of which 8 provided audio-only feedback, 8 provided visual-only feedback, and 8 provided both audio and visual feedback. Between every trial, the navigation target was moved to a different location on the surgical volume, and further, the volume was randomly rotated in all three of the x, y, and z axes.

For each trial, continuous pointer data were collected from IBIS at a variable sampling rate of around 20 Hz.

To compensate for sampling rate variation, all data points were time-stamped to allow later data reconstruction; the data were recorded within the Pure Data data flow programming environment [22]. We also collected continuous EEG data from a Muse EEG headband [19]. The Muse headset is equipped with 7 sensors for reading 4 channels, two on the forehead (Fp1 and Fp2) and two behind the ears (TP9 and TP10). Although raw data capture is available, in our study we collected processed (de-noised and decomposed) time series measurements of power spectral density. Specifically, we looked at alpha-band activity relative to beta, theta, and gamma bands, as this EEG feature has been robustly associated with cognitive workload and working memory. Continuous data were broadcast by the Muse at a sampling rate of 10 Hz and were also recorded within Pure Data. It should be noted that the Muse, although a wearable consumer grade EEG device, has been shown to be feasible in a number of common EEG experimental designs, including oddball paradigms, reward learning tasks, and as a marker of cognitive load during performance-based tasks (e.g. [25,26]).

After the completion of the experiment, participants were asked to complete three separate NASA TLX questionnaires [10], one for each of the three experimental conditions (audio, visual, and audio-visual). A priori, we hypothesized that results from the questionnaire would positively correlate with averaged EEG cognitive load measurements taken from each trial.
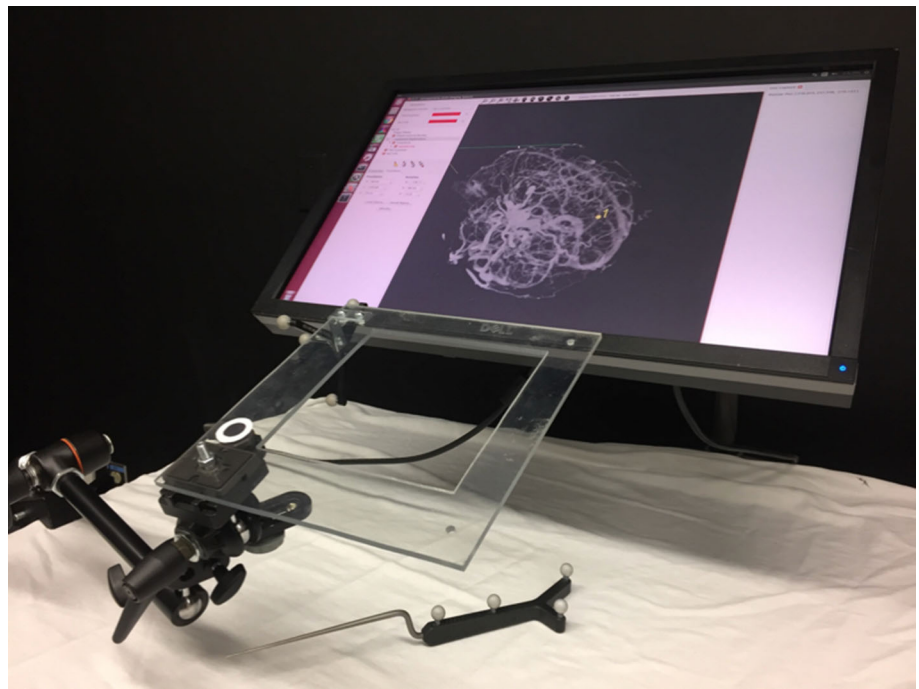
Results

Our main hypotheses stated that adding audio information to visual displays would (1) improve performance within a 3D navigation task, (2) be associated with reduced cognitive load, and (3) be rated as subjectively less demanding than visual trials alone. Below, we explore each of these hypotheses in turn.

Task accuracy data

There are several metrics that might be used to quantify "improved task performance" for 3D navigation tasks, including the required time to reach a target, the end accuracy of reaching a target, and/or the amount of error required to find a target. For convenience, we utilized the average total distance from target within each 15 s trial to compare feedback across each modality. We calculated the average total trial distance via a simple summation of the distance values between the surgical pointer and the randomly placed 3D virtual location, and then divided by the number of samples recorded within the trial (which was variable depending on several factors discussed above). These "average trial distances" were then used as data points to investigate

Fig. 1 Experimental setup: subjects used a tracked surgical pointer to navigate to a randomly placed point on a 3D cerebral CTA volume using visual, aural or both feedback methods

accuracy as a function of stimulus modality (audio, visual, and audio-visual) via a repeated-measures ANOVA. The results showed a marginally significant effect of condition, $F(2, 12) = 2.894$, $p = .057$; the means for each condition were: audio -42.84 mm, visual -47.82 mm, and audio-visual -20.57 mm, A Tukey post hoc test found marginally significant differences between the visual and audio-visual condition ($p = .06$); the difference between audio and audio-visual conditions was skewed in the predicted direction, but not found to be significant ($p = .16$). These results were consistent with the hypothesis that audio-visual feedback resulted in greater accuracy relative to visual-only feedback and are plotted in Fig. 2. Figure 3 visualizes the results as time series data calculated from averaged trial data for all participants as a function of stimulus modality.

Physiological data

In seeking to understand the relative cognitive load placed on participants during each trial, we utilized a single EEG metric (alpha-band change relative to theta, delta, and gamma band change) that could be easily collapsed into an average. For each trial, we calculated the average physiological working cognitive load across the entire 15 s second trial (roughly 300 data points per trial). These average trial observations were then analyzed via a repeated-measures ANOVA in order to investigate the effect of modality on participants' cognitive load throughout the task. The results showed a significant effect for condition, $F(2, 12) = 4.24$, $p = .0154$; the means for each condition were: audio $-.136$, visual
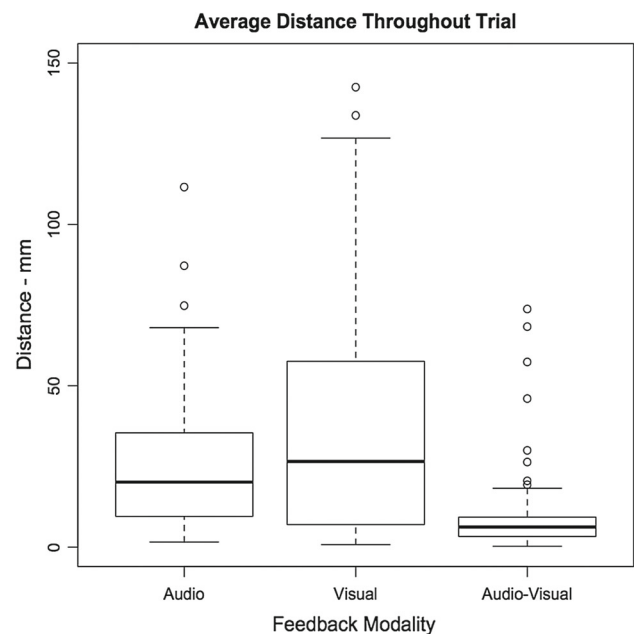

Fig. 2 Audio-visual feedback resulted in greater accuracy relative to visual-only feedback and aural feedback

$-.131$, and audio-visual $-.129$. A Tukey post hoc test found a significant difference between the audio and audio-visual condition ($p = .01$), whereas the difference between the visual and audio-visual condition was not significant ($p = .18$). These results were not consistent with the hypothesis that audio-visual feedback resulted in lower deduced working cognitive loads relative to visual-only feedback
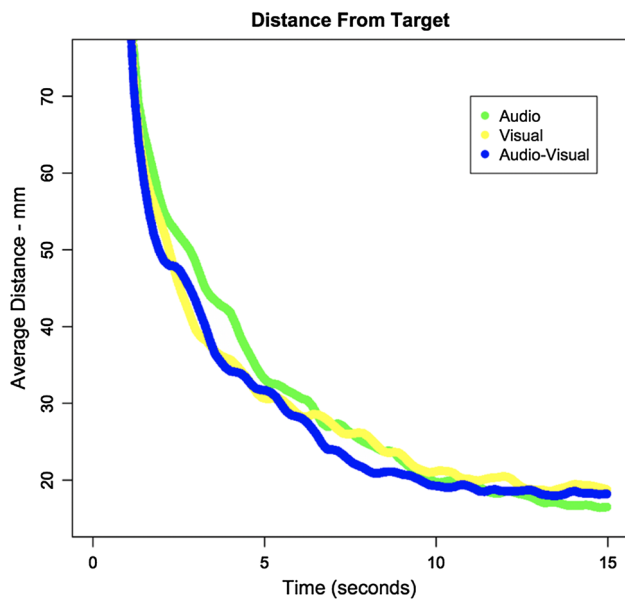
Fig. 3 Distance from target over time calculated from averaged trial data for all participants as a function of stimulus modality

Fig. 4 For cognitive load there was a significant difference between the audio and audio-visual conditions but no significant difference between the visual and audio-visual condition
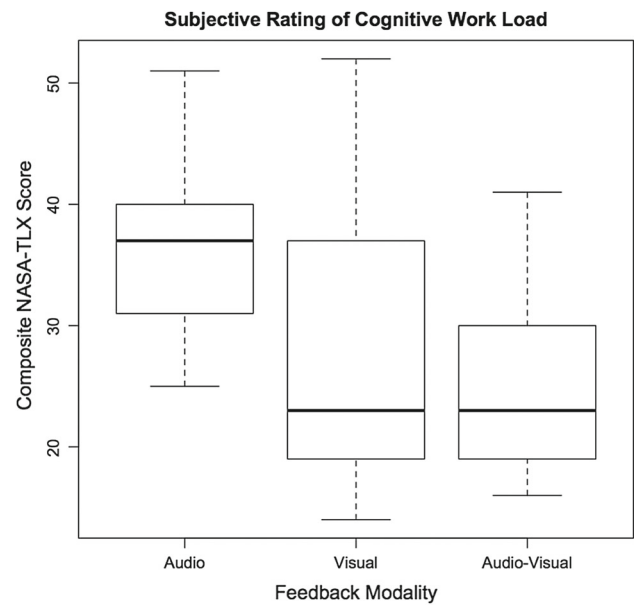

Fig. 5 For subjective cognitive workload (using the NASA TLX questionnaire) results showed significant differences between the audio and visual condition, as well as the audio and audio-visual conditions but no significant difference between the visual and audio-visual conditions
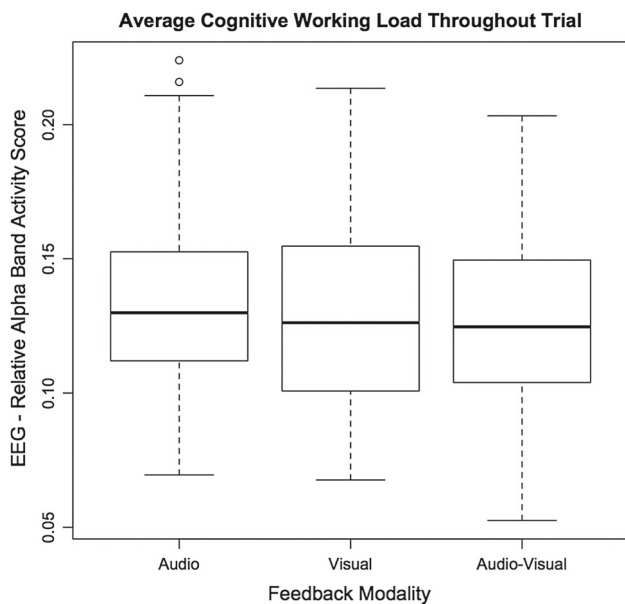
Subjective data

As with previous studies, we utilized a standardized post-experiment questionnaire in order to quantify participants' impressions about the difficulty of the experimental task across the three different modalities. Specifically, we used the NASA TLX test in order to probe each participant about the required mental, physical, and temporal demands for each modality, as well as their overall performance, frustration level, and amount of effort. Each of these factors was rated on a 10-point scale using an online version of the form. After completion, ratings from all six measurements were used to calculate (by summation) an index of subjective cognitive working load for each of the three modalities used in the experiment. Again using a repeated-measures ANOVA, we analyzed the effect of modality on subjective ratings of working load. The test found a significant main effect for modality, $F(2, 12) = 12.38$, $p < .001$; the means for each condition were: audio -36.31, visual -27.62, and audio-visual -24.77. A Tukey post hoc test found significant differences between the audio and visual condition ($p < .001$) as well as the audio and audio-visual conditions ($p = .004$), but no significant difference between the visual and audio-visual conditions. These results were not consistent with the hypothesis that audio-visual feedback resulted in reduced subjective working load relative to visual-only feedback (see Fig. 5).

A unique facet of our experimental design was the ability to correlate two different measures of cognitive working load (both physiological and subjective) within a simulated

(see Fig. 4). For this significant difference, higher relative alpha-band EEG activity may be viewed as a marker of a greater cognitive workload, and thus, the data showed that audio-only trials demanded greater cognitive resources when compared to audio-visual trials.
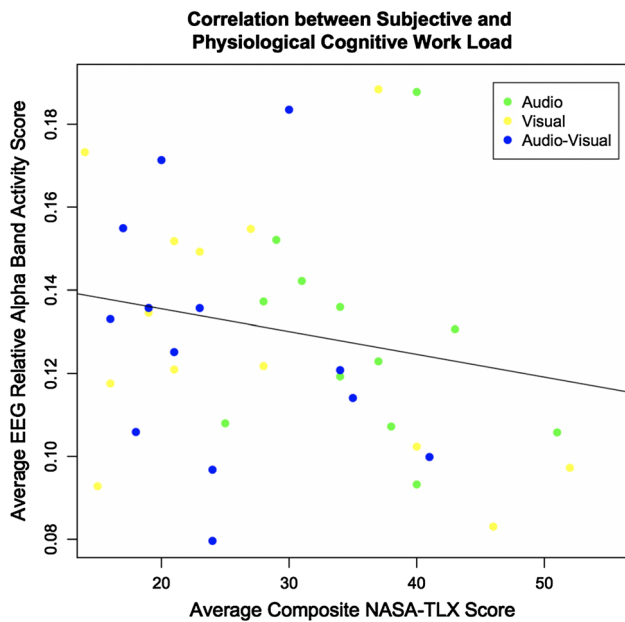
Fig. 6 The overall correlation between average EEG score and composite NASA TLX per subject was not statistically significant

surgical task. The motivation was to explore the efficacy of basic EEG recordings as a marker of cognitive working load for future work in human–computer interaction research. To examine the correlation between collected EEG data and the results from the NASA TLX questionnaire, we compared the average EEG measurement (across all trials) against the average TLX scores on a per person basis. We plotted these data in a two-dimensional space to show average EEG activity as a function of NASA TLX composite score (see Fig. 6). The overall correlation between average EEG score and composite NASA TLX was not statistically significant, $F(1, 11) = 1.039$, $p = .33$, $r = -.18$; the implications of which are discussed below.

Discussion and conclusions

Consistent with previous work on auditory display within image-guided surgery, our findings provide further support for the hypothesis that combining auditory distance cues with existing visual information may result in greater accuracy when locating a given target in 3D. To the contrary, when participants were asked to rate the perceived difficulty of navigating via our 3 stimulus modality types, their responses did not indicate that audio-visual navigation was significantly easier than visual navigation alone (although the results were skewed in this predicted direction). A similar pattern emerged within the reported physiological data, which did not reveal any significant difference between visual and audio-visual feedback, but did show results skewed in the predicted direc-

tion. Despite similar trends between these two markers of task difficulty, the two variables (i.e. subjective reporting and physiological data) were not found to be positively correlated on an individual basis.

The aim of this research was not to diminish the utility, or discourage the use of, a posteriori questionnaires within UX studies. There are a large number of well-tested tools for assessing user experience (e.g. SUS [27], UMUX [28], UMUX-Lite [29]), and even some tools which are specifically tailored to assess auditory user interfaces (e.g. BUZZ [30]). In seeking to build better real-time systems, we designed our experimental paradigm to compare a known real-time metric of cognitive workload with an established a posteriori assessment tool, as any known correlations between such metrics might potentially facilitate better research and application design.

The lack of a significant correlation (in either direction) between the subjective and physiological measurements of cognitive workload should not be misinterpreted as evidence of an absent effect, yet despite this null result, and because this result was largely unexpected, a bit of additional discussion is warranted. While there are countless reasons why these two markers for the same phenomenon might not be correlated, we offer two purely hypothetical explanations. First, one might assume that the subjective data are flawed, and if so, that participants tend to underestimate the true difficulty of a task when such ratings are made a posteriori. This scenario could explain the slight negative correlation reported above if participants used the subjective rating portion of the experiment to mitigate the perception of poor performance (i.e., higher physiological workload results in the need for more subjective padding of the true task difficulty). Indeed one of the shortcomings of the NASA TLX is that participants tend to correlate their task performance with their workload ratings. Secondly, one might assume that the physiological data are flawed, perhaps because of noisy signals. In this scenario, we would expect the relationship between physiological and subjective data to be random. Arguably, this seems to be a more plausible explanation for the reported data, and it would be possible to test via a small number of replication studies.

There were several limitations of this study that deserve special mention. In particular, in seeking to understand and develop a system for expert users, our convenience sample of novice users may fail to capture some of the nuances involved in a 3D target localization task. Secondly, the reported results represent a relatively small number of subjects, and it seems likely that a larger study would better reveal a more accurate picture of several marginally significant effects reported here. Thirdly, the provided analysis of time series data represents only a fraction of what could be gleaned from even this small dataset; additional analyses could potentially highlight if the physiological signals collected in this study were pat-

terned. Finally, limitations imposed by our low-cost EEG unit present real challenges to interpreting the data; additional work should be done to verify the physiological markers of interest on both consumer and medical grade equipment.

In future work, we plan to examine whether a stronger effect can be found by taking EEG recordings from a more sophisticated device with a wider array of electrodes. This would allow us to examine EEG signal change posteriorly as well as anteriorly, providing a more comprehensive examination of changes across the scalp. Increasing the number and quality of sensors would also facilitate additional analyses, such as event-related potentials (ERP) analysis, potential signal coherence across the scalp, and changes in signal power distribution.

While this specific study was more focused on a methodological question (i.e. how to quantify the potential benefits of auditory display) than a specific implementation (i.e. new sonification types), we wish to share one insight gleaned from a participant debriefing. Within existing neuronavigation systems that employ auditory display, and within the system utilized for this research, the data that feed the auditory signal demands careful design. In both our audio and audio-visual trials, the audio signal represented a 3-dimensional Euclidean distance from the target. This strategy makes sense when only auditory feedback is used, but it results in redundant information when employed within audio-visual settings. Ideally, audio information that is intended to be used in conjunction with visual information should not duplicate information that could be obtained visually. In an ideal audio-visual setup, visual information might be used to navigate the x and y dimensions, whereas audio information might be used only for z-dimension navigation.

In investigating the use auditory display within neuronavigation systems, the trends presented within the data continue to warrant additional research. In replicating similar experimental designs, and obtaining similar results, our aim was to contribute to the vast number of ideas about when, where, and how auditory display could best be deployed within neuronavigation systems. Unfortunately, the results presented here regarding cognitive workload as a metric for quantifying the value of auditory display systems were mixed, but nonetheless, offered a number of further questions for researchers in this field to address. If we wish to truly demonstrate that auditory display is useful within the operating room, then we must first and foremost verify that the tools we use to measure its utility are valid and robust.

Compliances with ethical standard

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Hermann T (2008) Taxonomy and definitions for sonification and auditory display. International Community for Auditory Display, Chicago
2. Kramer G, Walker B, Bonebright T, Cook P, Flowers JH, Miner N, Neuhoff J (2010) Sonification report: status of the field and research agenda. Technical Report. International Community for Auditory Display, 1999. http://www.icad.org/websiteV2.0/References/nsf.html. Accessed Jan 2019
3. Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25(5):975–979
4. Hansen C, Black D, Lange C, Rieber F, Lamadé W, Donati M, Hahn HK (2013) Auditory support for resection guidance in navigated liver surgery. Int J Med Robot Comput Assist Surg 9(1):36–43
5. Willems PWA, Noordmans HJ, van Overbeeke JJ, Viergever MA, Tulleken CAF, van der Sprenkel JB (2005) The impact of auditory feedback on neuronavigation. Acta Neurochir 147(2):167–173
6. Black D, Hansen C, Nabavi A, Kikinis R, Hahn H (2017) A Survey of auditory display in image-guided interventions. Int J Comput Assist Radiol Surg 1–12:1665
7. Bork F, Fuers B, Schneider AK, Pinto F, Graumann C, Navab N (2015) Auditory and visio-temporal distance coding for 3-dimensional perception in medical augmented reality. In: IEEE international symposium on mixed and augmented reality (ISMAR), pp 7–12
8. Wendt F, Zotter F, Frank M, Höldrich R (2017) Auditory distance control using a variable-directivity loudspeaker. Appl Sci 7(7):666
9. Anderson EW, Potter KC, Matzen LE, Shepherd JF, Preston GA, Silva CT (2011) A user study of visualization effectiveness using EEG and cognitive load. Comput Graph Forum 30(3):791–800
10. Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv Psychol 52:139–183
11. Hankins TC, Wilson GF (1998) A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. Aviat Space Environ Med 69(4):360–367
12. Naismith LM, Cheung JJ, Ringsted C, Cavalcanti RB (2015) Limitations of subjective cognitive load measures in simulationbased procedural training. Med Educ 49(8):805–814
13. Annett J (2002) Subjective rating scales: science or art? Ergonomics 45(14):966–987
14. Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Rev 29:169195
15. Yuksel BF, Oleson KB, Harrison L, Peck EM, Afergan D, Chang R, Jacob RJ (2016) Learn piano with BACh: an adaptive learning interface that adjusts task difficulty based on brain state. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 5372–5384

16. Michels L, Moazami-Goudarzi M, Jeanmonod D, Sarnthein J (2008) EEG alpha distinguishes between cuneal and precuneal activation in working memory. Neuroimage 40(3):1296–1310
17. Jensen O, Gelfand J, Kounios J, Lisman JE (2002) Oscillations in the alpha band (912 Hz) increase with memory load during retention in a short-term memory task. Cereb Cortex 12(8):877–882
18. Jensen O, Tesche CD (2002) Frontal theta activity in humans increases with memory load in a working memory task. Eur J Neurosci 15(8):1395–1399
19. Muse Developers (2015). http://developer.choosemuse.com/research-tools. Accessed Nov 2018
20. Roux F, Uhlhaas PJ (2014) Working memory and neural oscillations: alphagamma versus thetagamma codes for distinct WM information? Trends Cogn Sci 18(1):16–25
21. Drouin S, Kochanowska A, Kersten-Oertel M, Gerard IJ, Zelmann R, De Nigris D, Hall JA, Sinclair D, Petrecca K, Del Maestro R, Collins DLIBIS (2017) An OR ready open-source platform for image-guided neurosurgery. Int J Comput Assist Radiol Surg 12(3):363–378
22. Puckette M (1996) Pure data: another integrated computer music environment. In: Proceedings of the second intercollege computer music concerts, pp 37–41
23. Bencina R (2006) oscpack [computer software]
24. Plazak J, Drouin S, Collins L, Kersten-Oertel M (2017) Distance sonification in image-guided neurosurgery. Healthc Technol Lett 4(5):199–203
25. Krigolson OE, Williams CC, Norton A, Hassall CD, Colino FL (2017) Choosing MUSE: validation of a low-cost, portable EEG system for ERP research. Front Neurosci 11:109
26. Negi S, Mitra R (2018) EEG metrics to determine cognitive load and affective states: a pilot study. In: Proceedings of the 2018 ACM international joint conference and 2018 international symposium on pervasive and ubiquitous computing and wearable computers, pp 182–185
27. Brooke J (1996) SUS—a quick and dirty usability scale. Usability Eval Ind 189(194):4–7
28. Berkman MI, Karahoca D (2016) Re-assessing the usability metric for user experience (UMUX) scale. J Usability Stud 11(3):89–109
29. Lewis JR, Utesch BS, Maher DE (2013) UMUX-LITE: when there's no time for the SUS. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 2099–2102. ACM
30. Tomlinson BJ, Noah BE, Walker BN (2018) BUZZ: an auditory interface user experience scale. In: Extended abstracts of the 2018 CHI conference on human factors in computing systems, p LBW096. ACM