

Research Letter

Assessment of Radiation Therapy Technologists' Workload and Situation Awareness: Monitoring 2 Versus 3 Collocated Display Monitors



Joseph K. Nuamah, PhD,^a Prithima R. Mosaly, PhD,^{a,b,*} Robert Adams, RTT, ED,^a Kathik Adapa, MBBS, MPH,^{a,b} Bhisham S. Chera, MD,^a Lawrence B. Marks, MD,^a and Lukasz M. Mazur, PhD^{a,b}

^aDepartment of Radiation Oncology, University of North Carolina at Chapel Hill, North Carolina; and ^bSchool of Information and Library Sciences, University of North Carolina at Chapel Hill, North Carolina

Received 17 April 2020; revised 7 September 2020; accepted 14 September 2020

Abstract

Purpose: This study aimed to assess the effect of monitoring 2 versus 3 collocated displays on radiation therapist technologists' (RTTs) workload (WL) and situation awareness (SA) during routine treatment delivery tasks.

Methods and Materials: Seven RTTs completed 4 simulated treatment delivery scenarios (2 scenarios per experimental condition; 2 vs 3 collocated displays) in a within-subject experiment. WL was subjectively measured using the National Aeronautics and Space Administration (NASA) Task Load Index, and objectively measured using eye activity measures. SA was subjectively measured using the SA rating technique, and objectively measured using the SA global assessment technique. Two-tailed paired *t* tests were conducted to test for differences in means when parametric assumptions were satisfied, otherwise Wilcoxon signed-rank tests were conducted. A .05 level of significance was applied to all statistical tests.

Results: No statistically and clinically significant differences were observed between monitoring 2 versus 3 monitors on eye tracking measures (blink rate: 9.4 [4.8] vs 9.6 [4.0]; task evoked pupillary response: 0.16 [0.14] vs 0.21 [0.15]; NASA Task Load Index: 34.7 [19.8] vs 35.3 [20.4]; SA rating technique: 19.3 [6.2] vs 19.5 [7.0]; and SA global assessment technique scores: 100 [0] vs 100 [0]).

Conclusions: Our preliminary findings suggest that monitoring 3 collocated displays by 1 RTT does not impact WL and SA compared with monitoring 2 collocated displays. Only 2 of many possible configurations were investigated. If institutions removed the 3rd display based on the results of this study, there could be unforeseen error(s) if that display helped in situations not assessed in this study.

© 2020 Published by Elsevier Inc. on behalf of American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Sources of support: The research was supported in part by grant 1R03HS025258 from the Agency for Healthcare Research and Quality.

Disclosures: The authors declare no conflicts of interest. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Health Care Research and Quality.

* Corresponding author: Prithima R. Mosaly, PhD; E-mail: prithimareddy@gmail.com

<https://doi.org/10.1016/j.adro.2020.09.012>

2452-1094/© 2020 Published by Elsevier Inc. on behalf of American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Prior research suggests that the performance of health care providers, including radiation oncology professionals, often depends on their level of workload (WL) and situation awareness (SA).¹⁻⁵ Radiation therapy technologists (RTTs) are front-line professionals who have

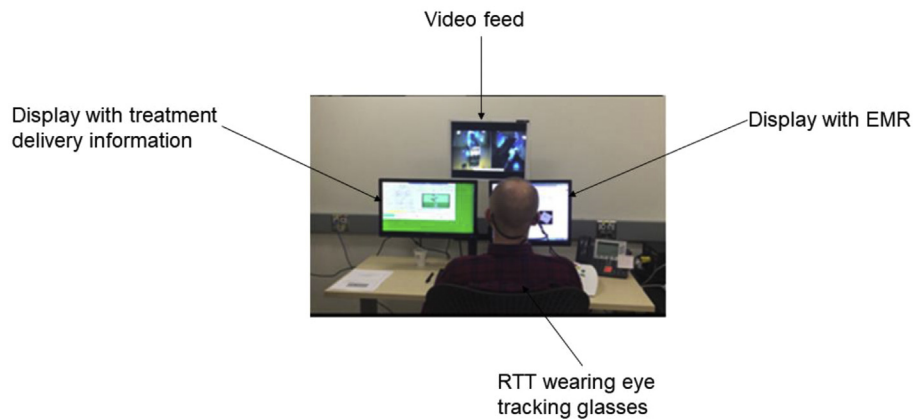


Figure 1 Radiation therapist technologist monitoring 3 collocated displays. The display dedicated to live-stream video of a patient undergoing treatment is turned off in the 2-display configuration.

the challenging task of monitoring the delivery of ionizing radiation to patients, where WL and attention-related errors can cause patient harm. RTTs need to attend to multiple types of information concurrently (eg, setup, imaging, patient information) that are usually displayed in various ways, including across several computer displays. For example, depending on vendors' implementation, RTTs are often set up to monitor patient information organized on 2 or 3 displays. In some departments, 1 RTT monitors 2 collocated displays (1 display for electronic medical records [EMRs], and 1 display for treatment delivery) with another RTT monitoring 1 display dedicated to the live-stream video of the patient undergoing treatment.

The tasks are particularly challenging at the start of a patient's course of therapy because the information is relatively new and must be carefully overseen. However, little is known regarding how many displays an RTT should monitor during treatment delivery. Herein, we assessed the impact of the number of information displays on RTTs' WL and SA during routine treatment delivery tasks. Specifically, we compared RTTs' WL and SA while monitoring 2 versus 3 collocated displays. In principle, increasing the number of collocated displays could increase visual attention switches and WL, resulting in decreased SA.⁶

Methods and materials

Study participants and setting

A convenience sample of 7 RTTs (3 men and 4 women; all with >3 years of experience), incentivized with a \$100 gift card, participated in this internal review board-approved study, conducted in the Human Factors Laboratory within the Department of Radiation Oncology at the University of North Carolina at Chapel Hill. Simulated assessments were performed using an emulator

and workstations that closely replicated RTTs' typical working environment, consisting of 2 versus 3 collocated information display configurations (Fig 1). For the 2 displays, these included the EMR and treatment delivery information (assuming that the 2nd RTT [actor in the experiment; professional RTT] would monitor the 3rd [not collocated] display with video feed of patient undergoing treatment). For the 3 displays, these included the EMR, treatment delivery information, and video of the patient undergoing treatment. In this configuration, 1 RTT would assume responsibility of monitoring all 3 displays.

Data collection

Each participant was given no time limit to complete 2 simulated scenarios in each display configuration. The scenarios were chosen based on the common incident themes observed in our department's incident learning database. The scenarios were randomized between the 2 configurations to minimize the effect of (any) task difficulty, although per design there were no differences in difficulty between the simulated scenarios.

Quantification of workload

The National Aeronautics and Space Administration Task Load Index (NASA-TLX),⁷ a widely used subjective WL measure in radiation oncology research,^{1-3,5} was used to quantify participants' perceived WL. Participants completed the NASA-TLX questionnaire at the end of each scenario. Each participant's rating of each of the 6 NASA-TLX dimensions (mental demand, physical demand, temporal demand, frustration, effort, and performance) was combined with the weight given to that dimension by that participant to obtain the participant's composite WL score for each scenario. Blink rate and task-evoked pupillary response, which are 2 widely validated eye activity measures,^{2,8} were used to quantify participants' physiological WL. Eye tracking glasses by

Table 1 Results of subjective and objective workload measures

Workload measures	Configuration		Statistical test	P value	Effect size	95% CI
	2-display	3-display				
Subjective workload						
Global NASA Task Load Index	34.79 (19.85)	35.33 (20.44)	Wilcoxon signed-rank test	.84	0.10*	0.41-0.43
Mental demand	48.21 (25.91)	51.07 (28.77)	Two-tailed paired <i>t</i> test	.14	0.10*	−0.65 to 0.85
Physical demand	18.57 (17.59)	22.50 (17.73)	Wilcoxon signed-rank test	.29	0.16†	−0.27 to 0.54
Temporal demand	23.21 (19.96)	23.57 (18.75)	Wilcoxon signed-rank test	.47	0.03*	−0.36 to 0.42
Performance	32.86 (30.17)	37.86 (32.62)	Wilcoxon signed-rank test	.61	0.06*	−0.36 to 0.46
Effort	34.64 (24.37)	39.29 (26.01)	Wilcoxon signed-rank test	.15	0.11*	−0.32 to 0.51
Frustration	38.93 (31.51)	37.86 (33.09)	Wilcoxon signed-rank test	.48	−0.01*	−0.42 to 0.41
Objective workload						
Blink rate	9.44 (4.83)	9.67 (4.00)	Two-tailed paired <i>t</i> test	.44	0.06*	−0.38 to 0.47
Task-evoked pupillary response	0.16 (0.14)	0.21 (0.15)	Wilcoxon signed-rank test	.94	0.26†	−0.18 to 0.61

Abbreviations: CI = confidence interval; NASA = National Aeronautics and Space Administration.

* Negligible.

† Small.

Natural Gaze (Sensomotor Inc, Munich, Germany), worn by participants while they engaged with the displays, were used to record eye activity data.

Quantification of situation awareness

The SA rating technique (SART),⁹ a widely used subjective technique,⁹⁻¹¹ was used to quantify participants' perceived SA. The SART questionnaire, administered after the trial, uses 3 dimensions: Attentional demand, attentional supply, and understanding to measure participants' SA. Participants rated each dimension on a 7-point rating scale (1 = low; 7 = high). For each participant and scenario, a composite SART score was computed using the following formula: $SART = U - (D - S)$, where *U* summed the understanding representing familiarity and quality/quantity of the information being processed, *D* summed the attentional demand representing instability, variability, and complexity of the situation, and *S* summed the attentional supply representing perceived readiness for the activity, spare mental capacity, good concentration, and proper division of attention. The SA global assessment technique (SAGAT),¹² validated in many domains including health care,¹³⁻¹⁵ was used to objectively quantify participants' SA. During each simulated scenario, participants were asked a number of SA-related questions, with at least 2 questions related to each of the displays (eg, Who are we treating? Does the patient have bolus? What field are you on right now? What side of the patient is the gantry on?), and were given up to 10 seconds to respond to each question. Correct versus incorrect responses were coded as 1 and 0, respectively, and then averaged to form a composite score. The 2nd RTT did not answer any SAGAT-related questions during the scenarios.

Statistical analysis

A Shapiro-Wilk test was used to test for normality. Two-tailed paired *t* tests were conducted when normality assumptions were satisfied, otherwise Wilcoxon signed-rank tests were conducted. A .05 level of significance was applied to all statistical tests. The analyses were performed using R software, version 3.6.2.¹⁶

Results

Workload (NASA-TLX, blink rate, task-evoked pupillary response)

There were no statistically and clinically significant differences in NASA-TLX while indicating higher scores for the 3 versus 2 display configuration, except for the frustration subdimension of the NASA-TLX (Table 1).

Situation awareness (SART and SAGAT)

There were no statistically and clinically significant differences on the outcome measures while indicating higher scores for the 3 versus 2 display configuration, except for the attentional demand subdimension of SART (Table 2).

Discussion

To our knowledge, no prior study has investigated the effects of 2 versus 3 collocated displays on RTTs' WL and SA. The results suggest that the addition of the 3rd

Table 2 Results of subjective and objective SA measures

	Configuration		Statistical test	P value	Effect size	95% Confidence interval
	2-display	3-display				
Subjective SA						
Composite SA rating technique	19.36 (6.20)	19.50 (7.04)	Two-tailed paired <i>t</i> test	.94	0.02*	−0.73 to 0.78
Attentional demand	3.10 (1.28)	2.76 (1.17)	Wilcoxon signed-rank test	.09	−0.16†	−0.53 to 0.25
Attentional supply	4.12 (0.78)	4.23 (0.92)	Wilcoxon signed-rank test	.86	0.06*	−0.36 to 0.47
Understanding	6.07 (1.37)	6.36 (0.93)	Wilcoxon signed-rank test	.97	0.05†	−0.34 to 0.43
Objective SA						
SA global assessment technique	100.00 (0.00)	100.00 (0.00)				

Abbreviation: SA = situation awareness.

* Negligible.

† Small.

collocated display monitor appeared to have no significant impact on RTTs' WL or SA during treatment delivery. Although measures associated with 3 displays were higher compared with 2 displays, the differences were negligible and thus neither statistically significant nor clinically relevant.

Specifically, in both experimental conditions, the global and individual dimensions of the NASA-TLX instrument indicated acceptable levels of WL (global NASA-TLX ≈ 35), and scores >55 have been associated with high perceived WL and degradation in performance across many domains, including radiation therapy.^{1,17} This finding was further confirmed by both objective WL measures (eg, blink rate: ≈ 9 blinks/minute, suggesting a WL level comparable with a routine reading task¹⁸; task-evoked pupillary response: average increase of 0.16–0.21 mm, suggesting a WL level comparable with engagement of 2–3 chunks of information¹⁹, which is below the 7 ± 2 chunks of information that most individuals can hold and recall in their working memory without degradation of performance).²⁰

Similarly to WL, in both experimental conditions the composite ($U - [D - S] \approx 19$) and individual dimensions of the SART instrument suggested reasonable levels of SA (eg, average understanding of the situation: ≈ 6 of 7 [7 = optimal score]; average attentional demand: ≈ 3 of 7 [1 = optimal score]; and average attentional supply: ≈ 4 of 7 [7 = optimal score]). This finding was further confirmed by the objective measure of SA (SAGAT), indicating that RTTs were able to correctly answer all SA-related questions.

This study has several limitations (eg, 7 RTTs from 2 academic institution; all with >3 years of experience), which does not allow for us to generalize our findings

across all RTTs and institutions. Specifically, we investigated only 2 of many possible configurations while recognizing that patient monitoring may consist of digital camera data from ≥ 1 angles, motion management interfaces, and possibly real-time surface or radiographic imaging. All these components may be organized on a number of displays, as several windows on the same display, or partially integrated with the treatment console, depending on different vendors' implementation. Thus, future research with a larger sample size of RTTs from different institutions and with varying levels of experience and multiple configurations is warranted. Our results could be biased by the standard configuration used in our institution, which includes 3 displays, with 1 RTT monitoring 2 collocated displays (1 display for EMR and 1 display for treatment delivery), and another RTT monitoring the live video of the patient undergoing treatment.

Conclusions

Nonetheless, these are noteworthy finding suggesting that monitoring 3 collocated displays by 1 RTT does not affect the WL and SA compared with monitoring 2 collocated displays.

Acknowledgments

The authors are grateful to the University of North Carolina Health System and the Lineberger Comprehensive Cancer Center for their support, and to the dosimetry

students for their help in developing the study scenarios and assisting in running the laboratory experiments.

References

1. Mazur LM, Mosaly PR, Hoyle LM, Jones EL, Chera BS, Marks LB. Relating physician's workload with errors during radiation therapy planning. *Pract Radiat Oncol*. 2014;4:71-75.
2. Mosaly PR, Guo H, Mazur L. Toward better understanding of task difficulty during physicians' interaction with electronic health record system (EHRs). *Int J Hum Comput Interact*. 2019:1-9.
3. Mazur LM, Mosaly PR, Hoyle LM, Jones EL, Marks LB. Subjective and objective quantification of physician's workload and performance during radiation therapy planning tasks. *Pract Radiat Oncol*. 2013;3:e171-e177.
4. Mazur LM, Mosaly PR, Moore C, Marks L. Association of the usability of electronic health records with cognitive workload and performance levels among physicians. *JAMA Netw Open*. 2019;2, e191709.
5. Mazur LM, Mosaly PR, Moore C, et al. Toward a better understanding of task demands, workload, and performance during physician-computer interactions. *J Am Med Inform Assoc*. 2016;23:1113-1120.
6. Vidulich MA, Tsang PS. The confluence of situation awareness and mental workload for adaptable human-machine systems. *J Cogn Eng Decis Mak*. 2015;9:95-97.
7. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: *Human Mental Workload*. Vol 52. Advances in Psychology. Amsterdam, the Netherlands: Elsevier; 1988:139-183.
8. Krejtz K, Duchowski AT, Niedzielska A, Biele C, Krejtz I. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PLoS One*. 2018;13:e0203629.
9. Taylor RM. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In: Salas E, ed. *Situational Awareness*. Abingdon, United Kingdom: Routledge; 2017:111-128.
10. Vincent L, Higham H, Greig P, et al. SC25 Using the situation awareness global assessment tool (SAGAT) and the situation awareness rating technique (SART) using standardised scenarios in simulation training for intensive care teams. *BMJ Simul Technol Enhanced Learning*. 2018;4:A26.2-A27.
11. Salmon PM, Stanton NA, Walker GH, et al. Measuring situation awareness in complex systems: Comparison of measures study. *Int J Ind Ergon*. 2009;39:490-500.
12. Endsley MR. Measurement of situation awareness in dynamic systems. *Hum Factors*. 1995;37:65-84.
13. Gardner AK, Kosemund M, Martinez J. Examining the feasibility and predictive validity of the SAGAT tool to assess situation awareness among medical trainees. *Simul Healthc*. 2017;12:17-21.
14. Morgan P, Tregunno D, Brydges R, et al. Using a situational awareness global assessment technique for interprofessional obstetrical team training with high fidelity simulation. *J Interprof Care*. 2015;29:13-19.
15. Rosenman ED, Dixon AJ, Webb JM, et al. A simulation-based approach to measuring team situational awareness in emergency medicine: A multicenter, observational study. *Acad Emerg Med*. 2018;25:196-204.
16. R: The R project for statistical computing. Available at: <https://www.r-project.org/>. Accessed February 24, 2020.
17. Grier RA. How high is high? A meta-analysis of NASA-TLX global workload scores. *Proc Hum Factors Ergon Soc*. 2015;59:1727-1731.
18. Doughty MJ. Consideration of three types of spontaneous eyeblink activity in normal humans: During reading and video display terminal use, in primary gaze, and while in conversation. *Optom Vis Sci*. 2001;78:712-725.
19. Mosaly PR, Mazur LM, Marks LB. Quantification of baseline pupillary response and task-evoked pupillary response during constant and incremental task load. *Ergonomics*. 2017;60:1369-1375.
20. Miller GA. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol Rev*. 1956;63:81-97.