Industrial and Manufacturing Systems Engineering
Publications

Industrial and Manufacturing Systems Engineering

6-2013

Full Virtual Reality vs. Integrated Virtual Reality Training in Welding

Richard T. Stone
Iowa State University, rstone@iastate.edu

Elease McLaurin
Iowa State University, e.mclaurin1@gmail.com

Peihan Zhong
Iowa State University

Kristopher Patrick Watts
Iowa State University, watts.kristopher@gmail.com

Full Virtual Reality vs. Integrated Virtual Reality Training in Welding

Abstract

This study demonstrates that both fully virtual and virtual reality (VR) integrated into real-world training programs are appropriate for use in the domain of welding training, depending on the level of task difficulty. Performance differences were virtually indistinguishable between participants in the fully virtual and the integrated training group at the low and medium weld difficulty levels. At the highest level of difficulty, it became apparent that the VR system was no longer solely sufficient for training. This study also tracked the usage patterns for the visual aids used in the VR simulator. These optional aids were presented to the users as overlays near the image of the weld as it was formed. Patterns observed suggest that the proper selection of certain overlays at certain stages during training was an indicator of success in both groups.

Disciplines

Industrial Engineering | Mechanical Engineering | Systems Engineering

WRC

Full Virtual Reality vs. Integrated Virtual Reality Training in Welding

A comparison is presented of the cognitive skill learning, physical skill learning, and use of training aids between the two types of training

BY R. T. STONE, E. McLAURIN, P. ZHONG, AND K. WATTS

ABSTRACT

This study demonstrates that both fully virtual and virtual reality (VR) integrated into real-world training programs are appropriate for use in the domain of welding training, depending on the level of task difficulty. Performance differences were virtually indistinguishable between participants in the fully virtual and the integrated training group at the low and medium weld difficulty levels. At the highest level of difficulty, it became apparent that the VR system was no longer solely sufficient for training. This study also tracked the usage patterns for the visual aids used in the VR simulator. These optional aids were presented to the users as overlays near the image of the weld as it was formed. Patterns observed suggest that the proper selection of certain overlays at certain stages during training was an indicator of success in both groups.

Introduction

Virtual reality (VR) has the potential to offer occupational training program developers a new tool to help meet the demands for more efficient skill training programs for hazardous environments. Developing some skills in the virtual environment allows for a reduction in material, time, and expert accessibility costs that are associated with traditional training methods. It also allows the novice to learn basic skills in a safer environment (Ref. 1). It has been suggested that VR simulators are effective at producing "pre-trained novices" in that they can teach some learning aspects but not others (Refs. 2–4). However, how the design of the VR simulator influences trainee learning has only received limited attention.

A number of studies have focused on how the fidelity of VR influences training efficiency (Refs. 5, 6). Other studies have only focused on the cognitive skill learning that occurs when using simulators (Refs. 7, 8). Some work has focused on assessing the impact of an augmented reality simulator on hand-eye coordination (Ref. 9). Little work has focused on usage patterns for trainee utilization of real-time feedback features and post-task feedback.

In a previous study, we compared the results of a traditional welding training program, which involved only real-world training, with one that integrated virtual reality training using a simulator with real-world training. From this study, we learned that in the area of welding, integrating virtual reality training into a real-world training program has a number of advantages over traditional training. These advantages include increased weld quality, higher certification rates, reduced training time, improved kinesthetic skill learning, and reduced costs for the simpler welds (Refs. 10, 11).

Follow-up Study Motivation

While conducting our previous study, we observed that before the integrated program trainees progressed to the portion of the training program where they were exposed to the real-world training, there was a significant trend of the integrated trainees achieving the preset mastery level with the VR simulator for the simpler welds. This mastery level was intended to indicate the time at which a trainee was sufficiently prepared to successfully complete the given weld. Based on these trends, it was expected that if the VR training was isolated from the real-world, the results of the integrated trainees would be similar to those of trainees who only had virtual reality training, both in terms of results and usage of the VR simulator features. However, because of the potential carryover effect from the real-world training, the validity of the preset mastery level could not be demonstrated experimentally in this prior study. Nonetheless, the observations during our previous work lead to an interest in the effect of the VR simulator features on trainee learning.

Research Goals

The goal of the present study was to first demonstrate the validity of successful training with the VR simulator. Given that validity, this study explored how trainees used the VR simulator features to learn cognitive and physical skills. These goals were addressed by comparing a fully virtual training program with an integrated training program in terms of the post-training performance of the participants. The performance was defined in terms of pass-fail weld completion rate, physical skill learning, and cognitive skill learning. For this study, the VRTEX®360 welding simulator was selected because it was capable of providing a level of realism and

R. T. STONE (rstone@iastate.edu) is with the Department of Industrial and Manufacturing Systems Engineering, and the Department of Mechanical Engineering, Iowa State University, Ames, Iowa. E. McLAURIN, P. ZHONG, and K. WATTS are with the Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, Iowa.

KEYWORDS

Virtual Reality (VR)
Real-World Training
Simulator
Integrated Training
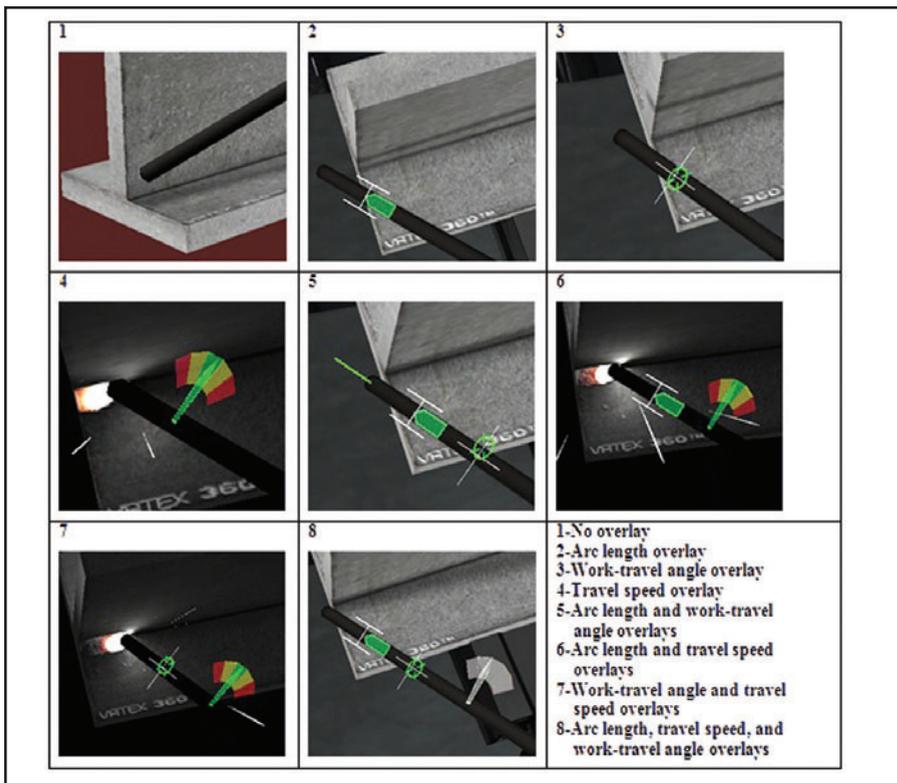Real-Time Feedback

WELDING RESEARCH

Fig. 1 — Images of possible overlay strategies (shown for a 2F weld).

kinesthetic feedback appropriate for the study. Of particular note was the availability of overlays that provided real-time visual feedback to the user. Aviation studies on visual information presented on heads-up displays in the form of overlays have shown that this method of presenting information can be absorbed and used by pilots to improve flight performance, as long as the attention needed to absorb the information presented does not exceed the available attention resource (Refs. 12, 13). For the VRTEX® 360, there were eight possible strategies for the overlay use. These eight strategies are shown in Fig. 1.

The domain of welding was selected due to the complex nature of the physical movements involved and the necessity to hone the specific physical movements for superior execution of welding tasks. According to previous real-world research in the domain of welding (Refs. 14, 15) and feedback gathered from experts, muscles that are of significant importance to welding performance include the deltoid, trapezius, extensor digitorum, and flexor carpi ulnaris. Regarding physical skill learning, it has been demonstrated that the activation and interactions of the muscles serve to distinguish between expert and novice control, ability, and stability during the commission of a task (Ref. 16). Finally, successful welding requires that the welder have a sufficient knowledge base to be able to judge variables related to creating a structurally sound weld.

For this study, it is hypothesized that 1) a fully virtual training program that is comparable to the VR component of an integrated training program will produce comparable results in terms of the kinesthetic and cognitive skills that are acquired, and 2) the selection of the type and number of real-time visual feedback indicators will be linked to the successful training of both the integrated and the VR trainees.

Methods

Front-End Analysis

An ethnographic study was used to define the pedagogical and technological aspects of weld training. In addition, eight expert welders were formally evaluated, particularly in terms of muscle interactions and posture, while conducting the four welds of interest in this study. These data were used to create expert models for comparison with the participants' physical skill learning.

Experimental Materials

A VR welding school and a real-world welding school were constructed on the campus of a Midwestern university. The materials stocked for the real-world school are listed in Appendix A. The VR welding school housed weld booths of the same size and dimension as their traditional counterparts. Each booth contained a new VRTEX® 360 Virtual Reality Arc Welding Trainer with SMAW attachments and multiple sets of welding jackets and gloves. This trainer was chosen due to the fact that it was the highest fidelity VR simulator currently available, and allows users to be fully immersed in a 3D VR environment while conducting welds. For the virtual training system, the user wore a weld helmet with integrated stereoscopic VR screens, used a SMAW weld attachment, of the same size and dimension as a real weld attachment, and used dynamic visual feedback, in the form of overlays, for known variables associated with welding.

Participants

There were 21 male participants randomly assigned to either the integrated training (11 participants) or the VR training (ten participants). The number of participants was initially limited in order to have a student-to-certified welding educator (CWE) ratio that was representative of real-world welding training classes, which generally do not exceed 12 people at a time. It should further be noted that all of the statistical measures utilized are appropriate for use with nonequal sample sizes.

All participants had no practical welding exposure and no experience in shielded metal arc welding (SMAW) prior to the beginning of the study. The integrated group had an average age of 41 (SD

Table 1 — List of Overlay Strategies

Overlay Strategies and Assigned Numbers

1	No overlay
2	Arc length overlay
3	Work-travel angle overlay
4	Travel speed overlay
5	Arc length and work-travel angle overlays
6	Arc length and travel speed overlays
7	Work-travel angle and travel speed overlays
8	Arc length, travel speed, and work-travel angle overlays

Fig. 2 — Images of weld types tested.

Fig. 3 — Certifications of each group from each weld type.

= 13.6) years and average height of 70.2 (SD = 2.4) in. The VR group had an average age of 41 (SD = 13.6) years and average height of 70.2 (SD = 2.4) in.

Independent and Dependent Variables

There were two independent variables in this experiment. The primary independent variable was training type with the two levels of integrated training and VR training. The second independent variable was weld type. The four weld types, in order of increasing difficulty, included the 2F (horizontal filet weld), 1G (flat groove weld), 3F (vertical filet weld), and 3G (vertical groove weld). Images of these weld types are shown in Fig. 2.

There were four dependent variables in this investigation: certification rate, physical skill learning, cognitive skill learning, and overlay usage.

Certification rates were determined based on whether or not the welds completed by the participants during the American Welding Society (AWS) welding certification tests were considered acceptable by the certification board. The quality of the welds was judged based on bend tests as well as the dimensions of the weld. For each of the weld quality tests, in addition to determining the acceptability of the weld, an overall weld quality score was assigned which ranged from 0 to 100.

In order to assess physical skill learning, electromyography (EMG) and postural observations were used. EMG data allowed the experimenters to examine the activation of the muscles of interest when participants performed the welding tasks. Further details regarding EMG instrumentation and methods are included in Appendix B. The postures adopted by par-
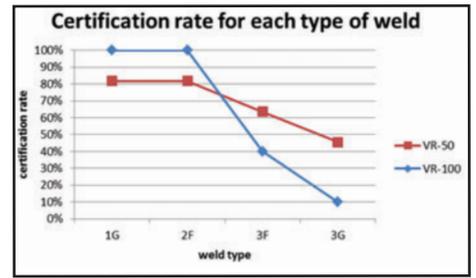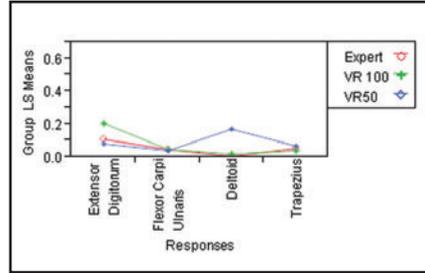

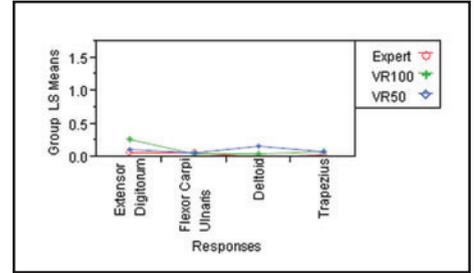Fig. 4 — Muscle responses for 2F weld type for expert, integrated, and VR welders.

Fig. 5 — Muscle responses for 1G weld type for expert, integrated, and VR welders.

ticipants while welding were recorded via video and direct observation by experimenters. The observations were consistent between observers.

In order to assess cognitive skill learning, a survey based on Crook's consideration of Bloom's taxonomy (Ref. 17) was used. Experimenters developed questions to measure cognitive skill learning for each weld type attempted by participants.

In addition to the previous three measures of performance, the visual overlays used for each run with the VR simulator were recorded. The overlays included in the VR welding simulator were as follows: travel speed (the appropriate horizontal speed that the welder should move the electrode holder), work angle and travel angle (the appropriate horizontal and vertical angle the welder should keep be-

tween the electrode and the weld coupon), and arc length (the appropriate distance the tip of the electrode should be from the weld coupon). While the system allowed for travel and work angle as separate overlays, a usage study conducted by the authors indicated that in all cases, participants using work and travel angle used both in equal proportion and were able to utilize them together with no performance impact; hence, for the purpose of this study, they are treated as a single overlay enhancement. As a result, there were eight different combinations of the overlays that could be used. These options are shown in Fig. 1 and described in Table 1, and will henceforth be collectively referred to as overlay strategies, and individually referred to by the assigned number.

Table 2 — Data Analysis Summary for Certification Rates

	2F	1G	3F	3G
χ^2	1.053	1.053	1.818	3.810
Prob > χ^2	0.305	0.305	0.178	0.05

Table 3 — Summary of Weld Quality Data

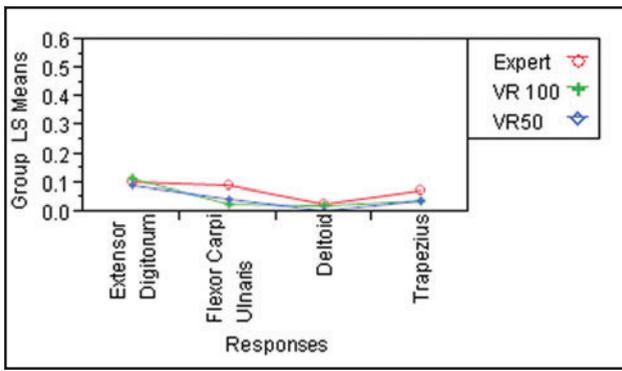| Weld Type	2F			1G			3F			3G		
Group	Mean	S.D.	p-Value	Mean	S.D.	p-Value	Mean	S.D.	p-Value	Mean	S.D.	p-Value
VR50	92	9		88	10		81	16		61	24	
			0.5916			0.9248			0.2670			0.2787
VR100	89.7	3		89.5	2		71.8	19		53	19	

Fig. 6 — Muscle responses for 3F weld type for expert, integrated, and VR welders.
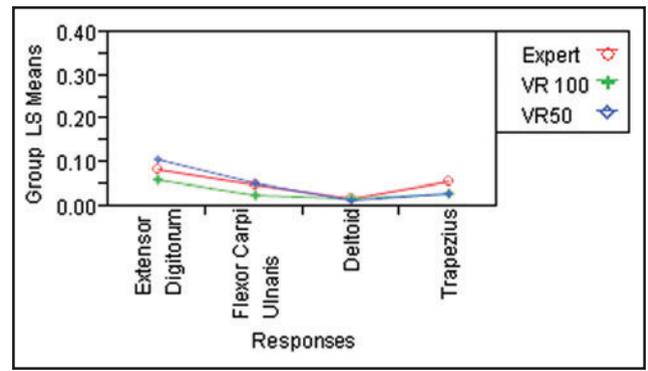

Fig. 7 — Muscle responses for 3G weld type for expert, integrated, and VR welders.
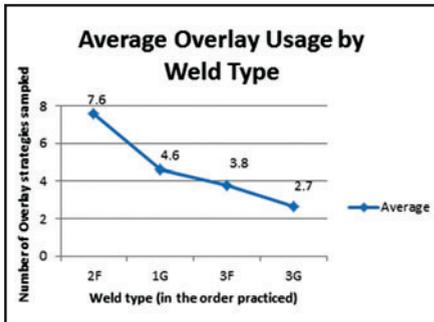

Fig. 8 — Trend of sampling as weld training progressed.

The authors specifically wanted to know how users would interact with the systems in a real-world setting. As a result, the experimental design included features that would allow them to use the system in a manner consistent with real training (for example, allowing large amounts of free practice). The experimenters noted in an earlier experiment a pattern of usage for the overlays. One of the goals of the present experiment was to see if that pattern would reemerge. To artificially force balance among conditions would not have allowed for these questions to be properly answered.

Experimental Procedure

Prior to experimentation, all participants were given informed consent followed by individual screening tests to ensure that they possessed normal visual acuity, depth perception, and hearing. Both the integrated and VR training groups were given take-home training materials and instructional videos on welding to supplement their study opportunities at home. Access to materials was regulated and no significant difference in exposure or opportunity was noted between participants by experimenters.

Also, in order to minimize the usability of the system as a factor, all participants were given training on the VR welding simulator system and system features prior to the experiments. The training ensured that all participants understood the system, the features of the system, and how to access those features. In particular, each overlay was explained in terms of its relationship to welding.

Participants in the integrated training group spent half of their time training (in pairs) with the VRTEX 360® system, and during this time, the VR system served as the instructor by providing weld quality feedback after every weld and by providing optional visual overlays that would guide the user to improve key aspects of their weld. The remaining half of their time was spent in traditional welding training under the direction of an AWS CWE/CWI® who presented lectures and supervised the participants when they practiced the weld types with real welding machines. Before changing from practic-

Table 4 — Analysis of Muscle Interactions

Weld Types	MANOVA Results	Expert vs. Integrated	Expert vs. VR	Integrated vs. VR
2F	$F(8, 42) = 2.2509$ $P = 0.0423$	$F(4, 13) = 2.3230$ $P = 0.1114$	$F(4, 12) = 1.9458$ $P = 0.1669$	$F(4, 14) = 3.0576$ $P = 0.0526$
1G	$F(8, 38) = 2.6760$ $P = 0.0195$	$F(4, 12) = 5.5532$ $P = 0.0091$	$F(4, 11) = 7.6919$ $P = 0.0033$	$F(4, 12) = 1.2126$ $P = 0.3557$
3F	$F(8, 44) = 1.7561$ $P = 0.1122$	N/A	N/A	N/A
3G	$F(8, 44) = 1.2657$ $P = 0.2858$	N/A	N/A	N/A

Table 5 — Portion of T-Test Results for Crook's Test Analysis

Crook's Consideration of Bloom's Taxonomy T-Test Values

Test Question Type Score	VR Mean Mean Score	Integrated SD	VR SD	Integrated Ratio	T	P Value
3G Knowledge	3.5	4.36	0.71	1.12	2.0868	0.0506
3G Comprehension	0.7	0.55	0.67	0.52	0.5901	0.5621
3G Application	2.1	2.36	0.32	0.67	1.1271	1.1271
3G Analysis	1.6	2.36	0.7	0.81	2.3027	0.0328

ing a given weld type with the VR welding machine to practicing with a real welding machine, the participants had to earn a simulator-generated quality score of 85% at least twice for the weld type. Once participants had moved on to the real-world training, they were not allowed to return to the virtual training. Further, they could only use as much time as they had used in the VR training for the real-world training aspect in order to keep the 50/50 split.

The participants in the VR training group spent all of their time learning in the same VR environment as did the members of the integrated group. For each weld type, participants in the VR group were allowed to practice in the VR environment for approximately the same amount of time that their integrated counterparts did. For example, the average participant in the integrated group spent 6.3 h in their VR training for the 3G weld, so the training limit for those in the VR group was set at 6.3 h with a 15% tolerance. This method ensured that the way in which the participants utilized the VR welding simulator and the VR instructional features would be directly comparable between groups.

Following the training for each weld type, participants were given their one and final weld certification test piece. They performed their prescribed test (2F, 3F, 1G, or 3G) in the presence of the CWI/CWE. Once completed, the test pieces underwent a visual inspection by the CWI/CWE on site. If the test piece passed visual inspection, it was then sent to an independent laboratory for weld quality testing. Certification or failure for the participant was based on the results of this testing. During pretest practice and the final testing plate, participants were fitted with electrodes so that experimenters could record EMG data while participants conducted their welds. The EMG data were collected during the middle phase of their welding (at this time their posture was stabilized) and averaged over 5 s (the weld usually took 30 s). Immediately following each certification test for all four weld types, participants were given a written cognitive survey related to the welding unit used and the weld type they had just performed.

Results

The performance measures for this study were weld certification rates, physical skill learning, and cognitive skill learning. Also, the relationship between the use of the visual overlay feedback and the welding performance was explored.

Certification Rate

As summarized in Table 2, the Chi-square test found no significant difference between the two groups across all weld types except for 3G. For 3G, $\chi^2_{0.05,\ 1} = 3.810$ and p = 0.05, indicating the integrated group had significantly more 3G certifications than the VR group. As a descriptive trend, the VR group had more certifications than the integrated group for the easier weld types (1G and 2F), while the integrated group had more certifications for the more difficult weld types (3G and 3F). The certification rates achieved by participants in the two groups for each of the four weld types are shown in Fig. 3. In addition, the mean weld quality score for both groups are shown in Table 3. No significant difference was found in quality between the two groups for any of the weld types.

Physical Skill Learning

Physical skill learning was assessed with respect to the average muscle activity expressed as a percentage of maximum voluntary contraction (MVC) for the four muscles of interest (deltoid, trapezius, extensor digitorum, and flexor carpi ulnaris muscles). The normalized muscle activities for these four muscles form a pattern of four dependent variables. A multivariate analysis of variance (MANOVA) was used to identify any interactions between the pattern formed by the experts and those of both of the experimental groups (integrated is VR 50 and VR is VR 100) for each of the weld types. Figures 4–7 show the muscle activity interaction profiles for each of the four weld types.

The results of the MANOVA for the 2F weld type show a significant difference between the expert, integrated, and VR groups. However, post-hoc MANOVA pairwise comparison tests re-veal no significant difference between any two conditions. This may have been due to the decrease in the degrees of freedom. The 1G weld type MANOVA revealed that there was a significant difference between the three conditions. Post-hoc MANOVA pairwise comparison tests revealed that the integrated and VR groups did not differ from one another. However, both the integrated and VR groups were found to be significantly different from the expert group. Author observations confirmed that both the VR and the integrated group adopted an altered posture which increased body stabilization. MANOVA results of the 3F and 3G weld types showed that there was no significant difference between the three conditions. These results are summarized in Table 4.

Cognitive Skill Learning

Cognitive skill learning was measured across four categories of the Crook's consideration for Bloom's taxonomy. These categories are, in order of increasing understanding, knowledge, comprehension, application, and analysis. In order to determine if the integrated and VR groups were significantly different, it was necessary to first determine if the data were normal and had homogeneity of variance. After it was determined the two groups had homogeneity of variance and were normal, a T-test was conducted for each question type for each weld type. The results can be seen in Table 5. Each test was conducted using $\alpha = 0.05$ and 19 degrees of freedom. The results indicate only one instance of significance within the Crook's taxonomy. For the 3G analysis level, the integrated group per-

Table 6 — ANOVA Results for Significance in Weld Types

Weld Type	F Ratio	F-Value
1G	6.9069	0.0020
2F	0.6275	0.6498
3F	11.2458	0.0001
3G	0.7102	0.5048

Table 7 — Post-Hoc Test Results for Significance in Overlay Usage

Weld Type	Overlay	Level			Mean Weld Quality	Weld Type	Overlay	Level			Mean Weld Quality
1G	5	A			93.8	3F	4	A			95.0
	4	A	B		92.0		5	A			93.0
	1	A	B		88.3		2	A	B		87.0
	3		B		84.7		6	A	B	C	78.0
	8			C	74.0		8			C	64.5
							3		B	C	63.0
							1			C	57.8

Table 8 — Dominant Usage of Overlays and Associated Participant Pass Percentages

	Overlay	1	2	3	4	5	6	7	8
2F	Number Dominant Used	8	0	2	6	5	3	0	0
	Percentage Passed	87.5	N/A	100	100	100	100	N/A	N/A
1G	Number Dominant Used	3	0	3	1	11	0	0	3
	Percentage Passed	100	N/A	100	100	100	N/A	N/A	66.7
3F	Number Dominant Used	5	1	1	1	8	1	0	4
	Percentage Passed	20	100	0	100	100	100	N/A	0
3G	Number Dominant Used	9	0	0	0	8	0	0	4
	Percentage Passed	22.2	N/A	N/A	N/A	50	N/A	N/A	0

formed significantly better than the VR group. For these T-tests, the alpha criterion was not adjusted because each level was independent and as such an adjustment such as a Bonferonni adjustment would not be applicable, even if MANOVA tests had been used to analyze the data (Refs. 18, 19, 20).

Overlay Usage

Also of interest for this study was determining if the use of the visual overlays had any impact on the performance measures. There were eight possible strategies for the overlay usage. For each of the VR welds completed by the participants, a report was generated by the system, which included a listing of any overlays used and the weld quality score. These reports were used to identify and analyze any patterns in terms of overlay usage strategy and weld quality.

It was first determined if the integrated group used the overlays during their VR training differently than the VR group. For each of the four weld types, a series of T-tests was performed to compare the percentage of usage for each of the overlay strategies for the two groups (VR and integrated). It should be noted that the alpha criterion was adjusted using the Bonferroni method. The results of these T-tests indicated the two groups were not significantly different in their use of the overlays. As a result, for the following analyses, the data for the two groups could be combined. An ANOVA was used to determine if the use of any of the overlay strategies during the welding training had a significant impact on the weld quality score for the real-world test welds submitted for certification. This test indicated that the weld quality of the 1G and the 3F weld type were significantly impacted by the choice of overlay strategy. The results are shown in Table 6.

Next, Tukey post-hoc test was conducted using the T distribution at $\alpha = 0.05$ to determine which overlay(s) were significant and if they produced more or lesser quality welds. The results for this test are shown in Table 7. For both the 1G and 3F weld types, overlay strategies 4 and 5 produced high mean weld quality scores (high is >90), while overlay strategies 1, 3, and 8 produced low mean weld quality scores (low is <90).

However, examining the quality scores in isolation presents an incomplete picture. The dominant overlay strategy used for each participant for each weld type needs to be considered. Examining these data in conjunction with the pass rate will allow for greater transparency of what overlay strategy was the most successful for producing passing welds. Table 8 shows the dominate strategy by weld type as well as the percentage of people who used that overlay type who passed their weld. By examining Table 8, and keeping in mind the previously determined information that overlays 4 and 5 were the most successful overlays, it can observed that for the two significant trials (1G and 3F), overlay 5 was much more widely used. As a result, it can be concluded that overlay 5 was the most used and successful strategy.

Another trend observed was a decrease in the sampling of overlay strategies as the participants practiced more difficult weld types. Figure 8 shows the number of overlay strategies sampled as the participants progressed through the VR welding training.

Discussion

For the comparison of performance measures between integrated and VR welding training, it was hypothesized that 1) a fully virtual training program that is comparable to the VR component of an integrated training program will produce comparable results in terms of the kinesthetic and cognitive skills that are acquired, and 2) the selection of the type and number of real-time visual feedback indicators will be linked to the successful training of both the integrated and the VR trainees. A discussion of the results of this study in light of these hypotheses follows for each of the weld types.

2F Weld Type

The 2F weld type was the simplest type of weld to complete for this study. The analysis of the certifications obtained by participants in both groups revealed no significant difference. This indicated that for the simplest weld, the VR training was overall as effective as the real-world training. Since there was no interaction of the muscle activity patterns for the VR and integrated groups when compared to the experts, interfacing with the physical VR simulator tools was sufficient to develop similar physical skills, in terms of motor control, as experts. The results also showed that the reference materials that the VR group had access to (welding CD, welding texts, and information presented through the VR simulator interface) were sufficient to produce equivalent cognitive skill learning as what the integrated group gained with the welding lectures and the real-world welding exposure.

1G Weld Type

The 1G weld type was a medium difficulty weld. The analysis of the certifications obtained by participants in both groups revealed no significant difference. This indicated that for this medium difficulty weld, the VR training was overall as effective as the real-world training. The results showed that the reference materials that the VR group had access to were sufficient to produce equivalent cognitive skill learning to what the integrated group gained with the welding lectures and the real-world welding exposure.

There was a significant difference in the muscle activity patterns for the VR and integrated groups as compared to the experts, indicating that interfacing with the physical VR simulator tools did not develop the same physical skills, in terms of motor control, as experts. However,

these differences exist because both the VR and integrated groups adopted the same, more stable posture for completing the weld. It should be noted that, although the integrated group completed real-world welding after the VR training, the real-world training did not change the pattern of using this alternate posture for completing the 1G weld.

3F Weld Type

The 3F weld type was also a medium difficulty weld. The analysis of the certifications obtained by participants in both groups revealed no significant difference. This indicated that for this medium difficulty weld, the VR training was overall as effective as the real-world training. Also, since there was no difference in the muscle activity patterns for the VR and integrated groups as compared to the experts, interfacing with the physical VR simulator tools was sufficient to develop similar physical skills, in terms of motor control, as experts. The results showed that the reference materials that the VR group had access to were sufficient to produce equivalent cognitive skill learning to what the integrated group gained with the welding lectures and the real-world welding exposure.

3G Weld Type

The 3G weld type was the most complex weld to complete for this study. The analysis of the certifications obtained by participants in both groups did reveal a significant difference, with fewer certifications for the VR group. This information indicates that the VR training was not sufficient for training the participants in how to correctly complete this complex weld. The integrated group had a higher certification rate than the VR group, indicating that the fidelity of the VR simulator played a role in the decreased efficiency of the training. Also, although the integrated group had a higher certification rate than the VR group, the number of certifications earned by the integrated group for the 3G weld was significantly less than the number earned for the other three weld types, indicating that the amount of real-world training time also played a role in the effectiveness of the training.

Also, there was a significant difference in the cognitive skill learning, with the integrated group having more understanding for analysis, the highest level of development. These results suggests that to have a more complete understanding of the more complex welds, the VR simulator is not sufficient. However, since there was no difference in the muscle activities for the VR and integrated as compared to the experts, interfacing with the physical

VR simulator tools was sufficient to develop similar physical skills, in terms of motor control, as experts.

Overlay Strategies

Regarding the relationship between the strategic use of the visual overlay feedback and the welding performance, it was observed that some overlay strategies led to better quality scores than others. The overlay strategies that consistently led to higher quality scores were 4 and 5. The overlay strategies that consistently led to lower quality scores were 8, 3, and 1. Furthermore, it was concluded that overlay 5 was the most widely used and successful strategy.

The trend indicated by these results shows that as the complexity of the weld increased, the participants who used more overlays, and thus increased the amount of feedback, tended to have improved performance. This general strategy has been shown to be successful in other studies (Ref. 21), up to a point. As participants continued to increase complexity, it was expected that they would reach a "tipping point" where performance no longer increased, but rather decreased. This effect was observed in this study and the tipping point was three overlays.

In addition, selecting an appropriate overlay strategy was most important for the medium difficulty welds. For the simplest weld, the overlays were not particularly necessary. For the complex weld, the selection of the overlays became less relevant because the fidelity of the VR simulator in accurately representing the welding conditions was limited. The trend observed for the sampling of the overlay strategies reflects this distinction. For the simplest weld, the sampling was increased because no strategy was truly more effective than the others; however, the participants were just beginning the VR training so they were more likely to explore the different options available. For the most complex weld, the sampling was greatly diminished because again no strategy was truly more effective than the other; however, the participants were now accustomed to the VR simulator and thus had no motivation to try other overlay strategies.

Conclusions

The results of this study have shown that VR and integrated training programs are both appropriate for use in the domain of weld training depending on the level of task difficulty. The differences between the VR and integrated groups were virtually indistinguishable at the low- and medium-weld difficulty levels. It was only at the highest level of difficulty that it became apparent that the VR system was no longer sufficient and required supplemen-

tation from real-world training. It is important to point out that the visual overlay usage in both groups followed similar trends. Both groups showed a trend of decreased sampling as the training progressed.

The worst welders tended to be more erratic in their selection and often would attempt to utilize too many overlay features (such as using travel speed, work-travel angle, and arc length) at the same time. Utilizing this many overlays at once resulted in participants no longer being able to give sufficient amounts of their visual attention resource to the actual weld bead being created. Hence, these individuals failed to properly transfer skills when placed in real-world environments.

The results of this study have demonstrated the advantages and limitations of fully virtual and integrated training in terms of feedback usage, performance, cognitive skill, and physical skill learning.

References

1. Lucas, J., Thabet, W., and Worlikar, P. 2007. Using virtual reality (VR) to improve conveyor belt safety in surface mining. *24th W78 Conference Maribor 2007 & 5th ITCEDU Workshop & 14th EG-ICE Workshop: Bringing ITC knowledge to work*. Maribor, Slovenia, pp. 431–438.

2. Gallagher, A. G., Ritter, E. M., Champion, H., Higgins, G., Fried, M. P., Moses, G., Smith, C. D., and Satava, R. M. 2005. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Annals of Surgery* 241(2): 364–72.

3. Van, S. K. R., Ritter, E. M., and Smith, C. D. 2006. The pretrained novice: using simulation-based training to improve learning in the operating room. *Surgical Innovation* 13(3): 198–204.

4. Di, Giulio E., Fregonese, D., Casetti, T., Cestari, R., Chilovi, F., D'Ambra, G., Di, Matteo G., and Delle, F. G. 2004. Training with a computer-based simulator achieves basic manual skills required for upper endoscopy: a randomized controlled trial. *Gastrointestinal Endoscopy* 60(2): 196–200.

5. Kenyon, R. V., and Afenya, M. B. 1995. Training in virtual and real environments. *Annals of Biomedical Engineering* 23: 445–455.

6. Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., and Penn, P. R. 2000. Training in virtual environments: transfer to real-world tasks and equivalence to real task training. *Ergonomics* 43(4): 494–511.

7. Munz, Y., Kumar, B. D., Moorthy, K., Bann, S., and Darzi, A. 2004. Laparoscopic virtual reality and box trainers: is one superior to the other. *Surgical Endoscopy* 18: 485–494.

8. Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., and Satava, R. M. 2002. Virtual reality training improves operating room performance. *Annals of Surgery* 236(4): 458–464.

9. Park, M., Schmidt, L., Schlick, C., and Luczak, H. 2007. Design and evaluation of an augmented reality welding helmet. *Human Factors and Ergonomics in Manufacturing and Serv-*

ice Industries 17: 317–330. doi: 10.1002/hfm.20077.

10. Stone, R. T., Watts, K., and Zhong, P. 2011. Virtual reality integrated welder training. *Welding Journal* 90(7): 136.

11. Stone, R. T., Watts, K. P., Zhong, P., and Wei, C.-S. 2011. Physical and cognitive effects of virtual reality integrated training. *Human Factors* 53(5): 558–572.

12. Calhoun, G. L., Draper, M. H., Abernathy, M. F., Patzek, M., and Delgado, F. 2005. Synthetic vision system for improving unmanned aerial vehicle operator situation awareness [5802-21]. *Proceedings SPIE the International Society for Optical Engineering*, 5802, 219–230.

13. Yeh, M., Merlo, J. L., Wickens, C. D., and Brandenburg, D. L. 2003. Head up versus head down: The costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors* 45: 390–407.

14. Beauchamp, Y., Marchand, D., and Galopin, M. 1997. Impact of the use of welding guns equipped with a fume extraction nozzle on muscular activation, psychophysical perception and quality of welded joint. *Proceedings of the Annual International Occupational Ergonomics and Safety Conference*, Washington, DC: IOS Press, pp. 197–200.

15. Herberts, P., and Kadefors, R. 1976. A study of painful shoulder in welders. *Acta Orthopaedica* 47: 381–387.

16. Keir, P. J., and MacDonell, C. W. 2004. Muscle activity during patient transfers: a preliminary study on the influence of lift assists and experience. *Ergonomics* 47(3): 296–306.

17. Crooks, T. 1988. Assessing student performance. *Green Guide No. 8*. Higher Education Research and Development Society of Australasia (HERDSA), Kensington.

18. Ertepinar and Geban. 1996. Effect of instruction supplied with the investigative-oriented laboratory approach on achievement in a science course. *Education Research* (38.3): 333–341.

19. Gokhale, A. 1995. Collaborative learning enhances critical thinking. *Journal of Technology Education* (7.1): 22–30.

20. Mevarech, Z., and Werner, S. 1985. Are mastery learning strategies beneficial for developing problem solving skills? *Higher Education* (14.4): 425–432.

21. Kaber, D., Kim, S-H., Kaufmann, K., Alexander, A., Steltzer, E., and Hsiang, S. 2009. Modeling the effects of HUD visual properties, pilot experience and flight scenario on a multi-dimensional measure of clutter. Hampton, Va., NASA Langley Research Center.

Appendix A

List of Materials for Real-World Welding School

1) Lincoln Electric Power MIG 350MP welding machine with SMAW (shielded metal arc welding) attachments
2) Two auto darkening welding helmets
3) Multiple sets of welding jackets and gloves
4) Power grinders
5) Slag hammer
6) Wire brushes
7) Welding table
8) Quenching buckets
9) Flat stock plates
10) Groove plates
11) 7018 electrodes
12) Runoff tabs
13) Consumables

Appendix B

Notes Regarding Collection of EMG Data

To collect EMG data, equipment by FlexComp Infiniti Cl. by Thought Technology Ltd. was used. The sample rate was 2048 samples/s. The sensor used was EMG MyoScan-Pro Sensor, and the electrode was T3402M-Triode by the same company. The EMG feedback signal was filtered, rectified, and smoothed automatically by the software packaged with the FlexComp Infiniti Cl hardware.

Maximum voluntary contractions (MVC) were performed in order to obtain a baseline for the maximum the participants were willing to exert their muscles. For the MVC for the trapezius and deltoid, the participants abducted their arms at the shoulder joint in the coronal plane at 90 deg against a stationary force. For the MVC for the extensor digitorum, the participants were asked to perform an extension of the wrist against a stationary object while the they held their extended arm (abducted about the shoulder in the sagittal plane) horizontally in front of them. Finally, for the MVC of the flexor carpiularis, the participants was asked to squeeze a handle in order to achieve a power grip. This was achieved while the participant's extended arm (abducted about the shoulder in the sagittal plane) was held horizontally in front of them.