Key Words: Training, Virtual Reality, Welding, EMG, Cognitive Development

Running Head/Short Title: Physical and Cognitive Effects of VRIT

Physical and Cognitive Effects of Virtual Reality Integrated Training

Richard T. Stone, **Kristopher Watts**, **Peihan Zhong**, and **Chen-Shuang Wei**, Iowa State University, Ames, IA

50 WORD DESCRIPTION

This study evaluated VR integrated training in terms of physical and cognitive impact. The experiment revealed that the use of virtual reality integrated training can have significant advantages over traditional training methods in the domain of weld training.

ABSTRACT

Objective: The objective of this study was to evaluate the cognitive and physical impact of virtual reality (VR) integrated training vs. traditional training methods in the domain of weld training. **Background:** Weld training is very important in various industries and represents a complex skill set appropriate for advanced training intervention. As such, there has been a long search for the most successful and most cost-effective method for training new welders. **Methods:** Participants in this study were randomly assigned to one of two separate training courses taught by sanctioned AWS (American Welding Society) CWIs (certified welding instructors); the duration of each course was two weeks. Upon completing the training for a specific weld type, participants were given the opportunity to test for the corresponding certification. Participants were evaluated in terms of their cognitive and physical parameters, total training time exposure, and welding certification awards earned. Each of the four weld types taught in this study represented

distinctively different levels of difficulty and required the development of specialized knowledge and skills. **Results:** This study demonstrated that participants in the virtual reality integrated training group (VR50) performed as well as, and in some cases significantly outperformed, the traditional welding training group (TW). The VR50 group was found to have a 41.6% increase in overall certifications earned over the TW group. **Conclusion:** Virtual reality technology is a valuable tool for the production of skilled welders in a shorter period of time and often with more highly developed skills than their traditionally trained counterparts **Application:** These findings strongly support the use of VR integrated training in the welding industry.

INTRODUCTION

MOTIVATION

Welding is a manufacturing process that is indispensable to industry and is found in every corner of the world. It is a task that requires training and specific skill development. As such, training systems for welders have existed nearly as long as the practice of welding itself. These training technologies range from the very early mechanical welder trainers (US Patent No. 1286529, December 3, 1918), to more complex mechanical systems, and subsequently to computer based training systems such as the CS Wave (Mellet-d'Huart, 2006), ARC+ (Choquet, 2008), and others (White et al., 2010). Computer based systems have most recently led to the use of augmented reality systems in weld training (Park et al., 2007). Despite technological advancements, these training systems all have drawbacks, namely the use of costly welding consumables in training exercises (Wormell et al., 2003).

VR systems can provide the same information as an augmented reality system, but without wasting consumables during the learning process. Additionally, VR systems can provide

visual feedback to the user, indicating proper welding parameters such as travel speed, work angle, arc length, etc., making such systems superior to traditional training. Despite the theoretical implications of success, real world applicability of augmented or VR welding integrated training (a model in which training time is split between virtual and real world environments) remains unproven.

Despite the existence of many studies that examine VR training techniques (e.g. Gallagher and Satava 2002; Grantcharov et al., 2004), there are less in the way of investigations into integrated VR training. As such, the goal of this work is to assess the impact of VR integrated training in the context of absolute performance, training time, and physical and cognitive development. The VR system used in this study provides the user with a fully immersive environment which included manipulation of physical weld tools. The use of feed forward visual overlays and post weld feedback in the VR system allowed the user to improve specific aspects of their welds during training. This level of oversight and guidance is simply not possible during normal weld training due to environmental factors and time constraints. The authors believe that such a tool has great potential to revolutionize the learning process in such a way as to exceed traditional training paradigms.

The authors hypothesize that the use of VR integrated training will result in student performance (certifications and quality of weld) that is equal to or better than that of students trained using traditional methods. Additionally, the authors hypothesize that the use of VR integrated training will greatly reduce overall student training times when compared to traditional training methods. It is expected that the physical development of students trained using VR integrated training will be similar to that of students trained using traditional methods, and that both groups will differ similarly from domain experts. It is further hypothesized that the

cognitive development of students trained using VR integrated training will be similar to that of students trained using traditional methods. Lastly, the authors hypothesize that mental workload (MWL) will not be significantly different between VR integrated and traditionally trained students. The authors believe the findings of this study will be directly applicable to most, if not all, similar integrated VR training regimens particularly if the task has both physical and cognitive components.

RELATED WORK

Very few studies have compared the effectiveness of virtual reality training in the context of a complex skill domain such as welding. Those studies that have been conducted are limited in scope and focus primarily on the VR technology as opposed to the human components (Fast et al., 2004; Mavrikios et al., 2006). There are many studies that have compared VR training to traditional training in fields other than welding (e.g., Ye et al., 1999), particularly in the field of laparoscopic surgery. Physicians have been trained in laparoscopic procedures using the VR Laparoscopic trainer MIST VR (Wilson et al., 1997). Several laparoscopic VR studies have failed to find a significant difference between VR training and traditional training (Munz et al., 2004, Torkington et al., 2000, etc), while others have found that VR training is superior to traditional training in many respects (Gurusamy et al., 2008). Laparoscopic surgery and welding tasks are dissimilar with regards to the overall procedural goals, however, the authors believe that they are similar in the complex nature of physical movement and kinesthetic memory that are essential for execution of both tasks. Jobs requiring regular performance of highly nuanced actions such as suturing or welding necessitate the honing of highly specific physical movements. Welding also requires that the welder have a sufficient knowledge-base to be able to judge variables related to creating a structurally sound weld.

In order to gauge the effectiveness of VR integrated training with respect to traditional training systems, the cognitive, physical, and performance impacts should all be addressed. The primary performance measure of interest in welding training is the certification of the welders. Past studies have measured welds based on specific criteria (Casalino et al., 2004; Kannan et al., 2006; Reyes-Valdés et al., 2006), which are part of the American Welding Society (AWS) certification requirements. For this reason, all student welders in this study were trained under an AWS Certified Welding Instructor CWI and were given the opportunity to certify through the AWS. Establishing the effectiveness of a training tool in the context of its primary goal (weld certifications) is a valid metric for establishing training impact. Further, the CWI judged welds produced in this study based on the quality of the weld, and graded each individual weld based on the AWS criteria for weld quality.

However, certifications themselves do not explain how human development is affected by the use of VR systems. For this reason, it is critically important to evaluate the cognitive impact associated with VR training environments (Wickens and Baker, 1995; Znbaka et al., 2005). Most studies comparing VR and traditional training have looked only at the performance and cognitive aspects associated with training (e.g., Munz et al., 2004; Seymour et al., 2002). These prior studies did not address the fact that a human's physical or kinesthetic memory development can be an important developmental factor, depending on the type of training being conveyed. Kinesthetic memory refers to the extent that the human body can recall its movements and postures. Through the use of sensory-motor learning and development of kinesthetic memory, a person can accomplish specific physical movements without thinking about how his/her body's parts should move (Ebert, 2009). Interviews with 16 professional

welders and welding engineers revealed that they felt sensory-motor learning was a very important component of successful welding.

Studies have demonstrated differences in muscle activity between expert and novice welders (Keir et al., 2004). According to previous research (Beauchamp et al., 1997; Kadefors, 1976) and feedback/observations gathered from experts, muscles that are of significant importance to welding performance include the deltoid, trapezius, extensor digitorum, and the flexor carpi ulnaris. The activation and interactions of these muscles serve to define expert welder control, ability, and stability during the commission of a weld.

METHOD

Pre-Study

Prior to the experiment, ethnographic evaluation and expert interviews, as well as observational studies and limited controlled experiments were performed in order to gather information on welding and weld training. The principal investigator attended a formal weld school and earned several welding certifications/qualifications. In addition, several other investigators were trained and have achieved welding certifications/qualifications. Investigators conducted 16 interviews with experienced welders, welding instructors, and welding engineers. A detailed biomechanical evaluation of eight expert welders performing the weld types discussed in this paper was also conducted. The eight experts were fitted with EMG electrodes and asked to perform all of the weld tests examined in this paper. This pilot data was used to refine the study protocol in order to ensure technical accuracy and applicability to the field of welding.

Experimental Materials and Setting

A traditional welding facility was constructed on the Iowa State University Campus (see figure 1). Specifically, a 740 sq. foot room housed six welding booths. Each booth was equipped with the following: a new Lincoln Electric Power MIG 350MP welder with SMAW (stick metal arc welding) attachments, two auto adjusting welding helmets, welding jackets and gloves, power grinders, slag hammer, wire brushes, welding table, quenching buckets, and other miscellaneous welding equipment. Consumables, welding plates, and electrodes were also made available to the trainees. The welding facility was outfitted with 1200 flat stock plates, 500 groove plates, and 8,000 7018 electrodes.

One floor below the traditional welding facility, a VR weld training facility was constructed. A 500 sq. foot room was converted to house weld booths of the same size and dimension as their traditional counterparts (see figure 1). Each booth contained a new VRTEX 360 VR Welding Trainer with SMAW attachments and welding jackets and gloves. The VRTEX 360 trainer was chosen because it is the highest fidelity VR simulator currently available, and allows for users to be fully immersed in a 3D VR environment while conducting welds. The VR helmet on the VRTEX 360 features stereo sound, with stereovision SVGA (800X600 at 24 bit depth with 200:1 high contrast) goggles that provide a 40° diagonal field of view, allowing for very realistic image sizes (105" at 12'). Although the system does not simulate heat or friction forces associated with welding, it does require the user to use and manipulate weld tools in the same way they would in real world welding. The VR tables were set at the same height as the traditional tables. Hence the virtual training systems created a very high fidelity situation, in which the user wore a weld helmet, used a SMAW weld attachment, (of the same size and dimension as a real weld attachment) and could hold placed weld stock (the plates) as would a real world welder. In

addition, the VR system has an advanced physics engine that allowed for the execution of real time welds that were virtually indistinguishable from welds made in the real world.

In this study two groups were assigned. One was trained with 100% traditional training (TW), while the other spent 50% of their time in traditional training and 50% of their time in the virtual environment (VR50). A 50:50 ratio was chosen because it gave experimenters ample time to collect detailed information and because interviews with welding instructors indicated that 50% VR integration represented the higher end of implementation that they would be comfortable accepting for use in near term welding schools.

Participants

There were 22 participants in total (21 males and one female). None of the participants had any prior practical welding exposure, nor experience in SMAW welding prior to the beginning of the study. All participants committed to 80 training hours over two weeks. The participants were randomly assigned to either the TW or the VR50 group. The TW group had an average age of 44 (SD=13) years, average height of 70.2 (SD=3.9) inches, and average weight of 215.3 (SD=26.3) pounds. The VR50 group had an average age of 41 (SD=13.6) years, average height of 70.2 (SD=2.4) inches, and average weight of 228.6 (SD=46.7) pounds.

Independent and Dependent Variables

The primary independent variable in this experiment was training type at two levels, representing the type of interface being tested: Traditional Welding (TW) and 50% Virtual Reality (VR50).   There were five dependent measures in this investigation: performance (certifications awarded and quality of weld), total training time, cognitive development, MWL,

and physical development. The primary performance measure was the certifications awarded, evaluated for each of four different weld types, including the 2F (horizontal filet weld), 1G (flat groove weld), 3F (vertical filet weld), and 3G welds (vertical groove weld) (see Figure 2). Each of these welds represents a subsequent and significant increase in level of difficulty over the previously listed weld.

The primary cognitive development measure in this study was based on Bloom's taxonomy. Bloom's taxonomy is a widely used method of measuring cognitive development (Bloom et al., 1956). In order to further simplify the taxonomy, Crook's consideration of Bloom's taxonomy (Crooks, 1988) was used. Crook's is similar to the traditional Bloom's taxonomy, but the questions posed are grouped into three categories: knowledge, understanding and application, and higher mental processing. This is a condensation of Bloom's original six categories. For this study, experimenters developed specific questions and tests to evaluate cognitive development in each weld type attempted by participants.

The MWL assessment used was the NASA Task Load Index (NASA TLX; Hart and Staveland, 1988; Hart and Wickens, 1990), which is a multidimensional subjective measure. NASA TLX has been used to compare the MWL before VR training with MWL after training (Sheik-Nainar et al., 2005; Stefanidis et al., 2007). In fact, out of the thousands of studies that use NASA TLX, 6% focus on evaluation of "virtual or augmented vision" (Hart, 2006).

The physical development measures used in this study included electromyography (EMG) feedback for the deltoid, trapezius, extensor digitorum, and the flexor carpi ulnaris muscles. Electrodes were placed on participants' skin in parallel with the fibers of these muscles according to recommendations of Zipp (1982) and Perotto (1996). EMG readouts allowed the experimenters to specifically examine the interactions between these muscles during

performance of welding tasks. A maximum voluntary contraction (MVC) was performed in order obtain a baseline value representative of the maximum voluntary muscle exertion in each participant. In order to measure the MVC for the trapezius and deltoid muscles, participants abducted their arms at the shoulder joint in the coronal plane at 90° against a stationary force. In order to gather the MVC for the extensor digitorum, the participants were asked to perform an extension of the wrist against a stationary object while the participants' extended arm (abducted about the shoulder in the sagittal plane) was held horizontally in front of them. Finally, in order to gather the MVC of the flexor carpi ulnaris, the participants were asked to squeeze a handle in order to achieve a power grip. This was achieved while the participants' extended arm (abducted about the shoulder in the sagittal plane) was held horizontally in front of them.

Experimental Tasks and Procedure

This experiment was conducted over a five week period; each group (VR50 or TW) was active over a defined two week period, with one inactive week separating the two active periods to ensure that there was no contact between VR50 and TW participants. Before participants arrived at the test site, they were randomly assigned to either the VR50 or the TW group. 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. In the traditional welding school (TW group), participants were trained in the application of welding techniques starting with the simplest weld (2F), and proceeding through the most difficult weld (3G). The time allotted to teaching each weld was fixed, and included formal lectures and practical lab training conducted by a fully certified AWS CWI. Following the training for each weld type, participants were given a single weld certification test piece. All weld tests (2F, 3F,

1G or 3G) were performed in the presence of the CWI. Once completed, test pieces underwent an on-site visual inspection by the CWI. If the test piece passed visual inspection, it was then sent to an independent laboratory for structural testing. Certification or failure for the participant was based on the results of this structural testing. During practical lab training, participants were fitted with electrodes, enabling experimenters to record EMG data while participants conducted welds. The EMG recorded activity of the previously discussed muscles of interest. During practical lab training periods, the CWI evaluated practice welds to determine whether or not a participant was ready to be tested prior to the end of their total allotted training time. If the total allotted time had been used, the participants were required to test regardless of whether or not the CWI had determined the participant ready to test. Immediately following the certification test for each of the four weld types, participants were given a written cognitive test related to the welding unit used and the weld type they had just performed. Further, a NASA TLX test was also administered to measure MWL.

In the VR integrated welding training (VR50 group), the experiment was conducted in the same basic manner as in the TW group. Both groups were given the same total allotted training time for each weld type. The major difference between traditional training and VR integrated training was in the training system itself. Participants in the VR50 group spent only 50% of their time training (lectures and practical lab training) under the direction of a CWI for each weld type. The remaining 50% of their time was spent training on the VRTEX 360 system. During this time the VR system itself served as the instructor by providing feedback after every weld and by providing optional use of visual overlay that would guide the user to improve key aspects of their weld such as travel speed, work/travel angle, and arc length. During VR training time, the participants (in pairs) used the VRTEX 360 to conduct virtual welds of each of the four

weld types. If a participant was able to earn a machine generated quality score of 85% at least twice for a weld, he/she was permitted to discontinue their VR training time early. If the VR50 participant reached the total time allotted for VR training (equal to 50% of the TW group's total allotted training time) he/she would be moved to the real world training regardless of whether or not a VR training score of 85% or better had been achieved. In this study it was rare for a participant to fail to achieve a VR training score of ≥85%; this situation occurred in only eight instances. Once participants had moved on from VR to real world training they were not allowed to return to the virtual training. Further, they were only allowed to spend as much time in real world training as they had used in the VR training; this rule was instituted in order to ensure the 50/50 training time ratio.

RESULTS

Certification and Quality

The number of certifications achieved by participants for each of the four weld types was tallied for each of the two training conditions. Figure 3 shows the total number of certifications for each weld type by training condition (a maximum of 11 certifications was possible for any given weld type).Due to the categorical nature of the data, chi square analysis was used to look for significant differences between groups for each of the four weld type. The chi square results were 2.2, 3.667, 1.692, 0.188 for 2F, 1G, 3F, 3G respectively, all of which were less than 3.841 ($\chi^2_{1,0.05}$).The number of certifications earned for the VR50 and TW groups were not found to be significantly different given $\alpha = 0.05$ for the 2F, 1G, 3F, or 3G weld types. While no significance was found via the chi square test, the result for 1G weld outcomes was very close to 3.841, which indicated that the VR50 group had a practically higher certification rate than did the TW group. It should be

noted that in all cases the VR50 group showed a descriptively higher rate of certifications than the TW group did.

In addition, each of the weld tests was accompanied by an overall quality score, ranging from 0 to 100. These scores were based on structural variables (bend tests) as well as visual parameters (dimensions/measurements of the weld). The mean quality score for both groups are shown in Table 1. Inferential analysis did not show a significant difference in quality between the two groups in terms of the 2F, 3F and 3G weld types. However the VR50 group was found to have significantly higher quality scores in terms of the 1G weld type ($T_{0.05, 1, 20}$ = -2.237, P = 0.037).

Table 1. *Bloom's T Values.*

weld type	2F		1G		3F		3G	
group	Mean	SD	Mean	SD	Mean	SD	Mean	SD
VR50	92	9	88	10	81	16	61	24
TT	86	20	70	26	73	23	62	23

Training Time

The total amount of time utilized by participants to complete their training and certification test for each weld type was averaged for both experimental groups. Figure 4 shows the mean task times for both training conditions by weld type. Analysis revealed that the VR50 group had significantly lower training times as compared to the TW group across all weld types 2F ($T_{0.05, 1, 20}$= 6.367, $P <= 0.001$), 1G ($T_{0.05, 1, 20}$= 3.468, P = 0.002), 3F ($T_{0.05, 1, 20}$= 3.340, P = 0.003), and 3G ($T_{0.05, 1, 20}$= 3.015, P = 0.007).

Cognitive Development

Cognitive development was assessed across four categories (knowledge, comprehension, application, and analysis), each representing a different aspect of cognitive capability. Figure 5 shows the mean scores within each of the four categories of cognitive development by weld type for both groups. In order to determine whether the experimental groups were significantly different with regards to cognitive development, a T-test was conducted for each question type within each weld type (2F,3F,1G, and 3G). The T-test compared the knowledge scores, comprehension scores, application scores, and analysis scores for the groups within each weld type. The results of this analysis can be seen in Table 2. Each of the tests was conducted with an $\alpha = 0.05$, and degrees of freedom of 20 (with the exception of the 2F test, which only has 19 degrees of freedom since one participant was given a test in the wrong order).

Table 2. *Crook's consideration of Bloom's Taxonomy T Values.*

Crook's Consideration of Bloom's Taxonomy T-Test Values						
Test Question Type	TW Mean Score	Mean Score	TW SD	VR50 SD	T Ratio	P Value
2F Knowledge	4	3.18	1.15	0.87	-1.84197	0.0811
2F Comprehension	0.6	1.18	0.52	0.6	2.362453	0.029*
2F Application	1.4	1.18	0.97	0.75	1.113528	0.2794
2F Analysis	2	2.636	1.33	0.81	1.337024	0.197
1G Knowledge	3.64	3.45	1.03	0.82	-0.45883	0.6513
1G Comprehension	2.18	2.27	0.75	1.01	0.239732	0.813
1G Application	1.82	2.18	0.87	0.6	1.135924	0.2694
1G Analysis	0.27	0.55	0.47	0.69	1.088214	0.2894
3F Knowledge	1.45	1.64	0.69	0.5	0.707107	0.4877
3F Comprehension	1.64	1.81	0.92	0.4	0.597614	0.5568
3F Application	2.64	2.1	0.5	0.7	-2.09529	0.0491*
3F Analysis	0.81	1.9	0.75	1.14	2.656845	0.0151*
3G Knowledge	3.45	4.36	1.04	1.12	1.976424	0.0621
3G Comprehension	1.18	0.55	0.87	0.52	-2.07322	0.0513
3G Application	2	2.36	0.63	0.67	1.304656	0.2068
3G Analysis	2	2.36	0.45	0.81	1.304656	0.2068

Our results indicate four instances of significance within the scores of the Crook's taxonomy. Two of these results are higher for the VR50 group, and two are higher for the TW group. The cognitive development areas with higher VR50 scores include 2F comprehension and 3F analysis. The cognitive development areas with a higher TW scores are 3F application and 3G comprehension.

Physical Development

Physical development was assessed with respect to the average muscle activity expressed as a percentage of maximum voluntary contraction (MVC) for the interaction of the four muscles

of interest (deltoid, trapezius, extensor digitorum, and the flexor carpi ulnaris muscles) within each of the weld types for expert welders and both of the experimental groups (VR50 and TW). Figure 6 shows the muscle activity interaction profiles for each of the four weld types.

A multivariate analysis of variance (MANOVA) was used to account for the multiple dependent variables (4 different muscles) that define kinesthetic memory development in this study. The experts' physical development was taken as the standard or optimal approach, therefore a result showing no significant difference between each of the training groups and the expert group would indicate desirable physical development in the trainees.

The results of the analysis for the 2F weld type revealed a significant difference between three conditions (expert, VR50, and TW) (F (8, 34) =2.4057, P = 0.0356). However, further analysis shows there was no significant difference between any two conditions (VR50 and TW (F (4, 10) =2.9188, P = 0.0772); expert and VR50 (F (4, 13) =2.3230, P =0.1114); expert and TW (F (4, 8) =2.1801, P=0.1617).

The outcomes for the 1G weld type also revealed a significant difference between the three conditions (F (8, 42) =3.3556), P =0.0046). Further analysis revealed that the expert and TW groups (F (4, 13) = 1.2238, P = 0.3480) did not differ from one another, whereas the VR50 group was found to be significantly different from both the TW (F (4, 14) = 3.6085, p = 0.0319) and the expert welder (F (4, 12) = 5.5532, P = 0.0091) conditions.

The results of the 3F weld type showed no significant difference between the three conditions (F (8, 46) =1.1931, P = 0.3240). The outcomes from the 3G analysis also showed no difference between any of the three conditions (F (8, 44) =0.5230, P = 0.8327).

Mental Workload

Overall MWL ratings were averaged from ratings of four welding types for the VR50 and TW groups. A series of T-tests showed no significant differences between training conditions in terms of MWL. The mean MWL rating was descriptively lower for the VR50 group when compared to the TW group, however there was no significant difference found between the two groups when considering the four welding types.

DISCUSSION

The main issues addressed in this study were, (1) evaluation of the ways in which VR integrated training (VR50) affected the overall training performance compared to traditional training (TW), and (2) comparison of the cognitive and physical developments in VR integrated weld training with those in traditional weld training. These issues will now be discussed by addressing the hypotheses posed at the beginning of this paper.

The authors first hypothesized that the use of VR integrated welding training would result in student performance (number of certifications) that was equal to or better than that of students trained using traditional methods. Results of this study revealed that the overall performance for three of the four weld types (2F, 3F, and 3G) was not distinguishable between the training groups, given a CI of 95%. Hence, for three weld types, overall performance is similar for both VR50 and TW groups with some descriptive benefits favoring VR training over traditional training. These findings are in agreement with those of several previous VR studies (e.g., Munz et al., 2004; Torkington et al., 2000). Of particular interest was the finding that the certifications for 1G weld type shows that the VR50 group outperformed the TW group to the point of practical significance. As mentioned earlier, the absolute number of 1G certifications earned was nearly significant, in

addition to the fact that the quality scores for this weld type were found to be significantly higher for the VR50 group. The results of this study support the original hypothesis, indicating that students using VR integrated training have certifications equal to (2F, 3F, and 3G welds) or better than (1G weld) those of students trained using traditional methods.

The authors' second hypothesis stated that the use of VR integrated training would greatly reduce student training times when compared to traditional training methods. The training times for participants in this study showed significant reductions in favor of the VR50 group across all weld types. Participants in the VR50 group acquired task-critical skills faster than their TW group counterparts, thus reducing the time investment needed to achieve competency levels necessary for testing. These time-based findings are highly consistent with previous studies in VR training (e.g., Lapointe and Robert, 2000; Jordan et al., 2000; Sabha Ganai et al., 2007). The reduction in training time requirements in VR50 students is believed to be due in part to the fact that the VRTEX 360 system gives constant feedback to users as they train, informing them which specific aspects of their welds required attention. In traditional training, students must take time to seek out instruction from the CWI; further time is expended as the CWI performs visual inspection and analysis of a weld and then provides the student with feedback. Post-experimental interviews confirmed that participants in the VR50 group were much more likely to seek feedback from the VR system than the instructor due to the fact that they felt the information given by the system was "delivered in a more timely manner" than that which could be gain from a shared instructor (the CWI). In addition, the VR environment allowed users to perform welds in less time by reducing the set-up time (time spent transporting and placing test plates in the work area), thus allowing the welder in the VR environment to complete more welds in the allotted training time. In accordance with the authors' hypothesis,

VR integrated weld training greatly reduced student training times when compared to traditional weld training methods in this study.

The third hypothesis theorized that the physical development of welders trained using VR integrated weld training would be similar to that of welders trained using traditional methods; both groups of newly trained novice welders were expected to differ similarly from expert welders. Given the relatively short training period (two weeks), the difficulty of the task, and the precision required to perform a passing weld, the authors did not expect to find a statistically significant difference for any of the weld types in terms of kinesthetic memory or physical development.

In the case of the 2F weld type, the physical development of the TW and VR50 participants were not distinguishable from one another, nor were they distinguishable from expert welders. This finding was understandable given the relatively simple nature of the 2F weld. In the case of the 1G weld type, the physical development of participants in the VR50 group was significantly different from that of expert welders, while there was no significant difference between the TW group and the expert group. The VR50 participants utilized a modified approach to performing the 1G weld which led to increased shoulder abduction and consequently a significant higher deltoid activity. This distinctive approach did not appear to be utilized by the expert welders nor was it taught by CWIs. The reason for this difference in approach is likely multi-factorial in nature.

The VR machine encourages exploration and experimentation in learning the welding procedure. It allows the user to actively evaluate their progress in real time and adjust their techniques accordingly. VR training also allows the user to start over and abandon poor welds without concern over wasted consumable materials. These attributes embolden users and

encourage exploration into new welding procedures and techniques without fear of accruing excessive expenses. This freedom to experiment combined with real-time system feedback led users of the VR system in this study to adopt a previously untaught approach to welding (specifically, a minor increase in shoulder abduction appeared to be advantageous to optimal travel angle and arc control). Post-experiment video review and interview sessions regarding the use of the exaggerated should abduction revealed that nearly all the VR50 participants were using visual overlays to perfect the first weld pass in the 1G position. This level of detail encouraged them to adopt an increased angle (resulting in increased shoulder abduction) beyond that suggested by the CWI.

The specific approach taken by the VR50 participants allowed for greater overall control in flat position welds, thus increasing the students' chances of making a structurally sound weld (as demonstrated with the higher 1G quality/certifications) by way of more consistent travel angle and arc length. However, considering the higher muscle activity level than the expert group, it can be conclude that the VR50 participants compromised their physical efficiency to achieve a sound weld, which, were they to practice more real welding and gradually master and refine their welding technique, might be discarded.

The significant difference of muscle activity for 3F between VR50 group and the other two groups was also due to different postures adopted by welders. Four of 11 participants in the VR50 group used a physical prop (either a can or even the grinder) to support their elbow, which resulted in lower muscle activity level than the other two groups. However, most often in real welding, welders will not be provided with any such physical supports due to the complexity of typical welding environments. It is difficult for a VR system to mimic every condition or environment. However the VR system provides trainees with general instruction on performing

the task and is useful in forming a basis of understanding and skill that can then be applied and modified as needed in the real world environment.

In the case of the 3G weld type (most difficult weld assessed in this study), the physical development of participants in the TW and VR50 groups were not distinguishable from one another, nor was either group found to be distinguishable from the expert welders. This outcome is better understood when the difficulty of this weld is considered.

In the case of the 3F and 3G weld types (both vertical welds, the two more difficult welds assessed in this study), the physical development of participants in the TW and VR50 groups were not distinguishable from one another, nor was either group found to be distinguishable from the expert welders. This outcome is better understood when the nature of vertical welds is examined. Vertical welds tend to leave less room for error, and as such are less forgiving of experimental techniques and deviation of physical approach. As a result, novice welders in both training groups adopted a more limited physical strategy; these strategies were highly similar to those used by expert welders.

When considering the 2F, 3F, and 3G weld positions, no significant difference between the VR50 and TW groups was apparent in terms of physical response. In addition, both groups of novice welders differed from, or were equally similar to expert welders in a statically similar manner. These findings were in accordance with the authors' third hypothesis, however a discordant finding arose in evaluating the physical development outcomes of the 1G weld type. Results clearly indicate that the VR50 students adopted an alternative and favorable strategy when compared to the traditional welders when performing the 1G weld. Furthermore within the 1G weld type, the VR50 group was significantly distinctive from the expert welder group whereas the TW group was not. Hence the third hypothesis was not robust across all weld types.

This finding indicates that VR integrated weld training systems give the trainees more freedom to develop their own approach of welding, which might not be practical in real welding. Providing the trainees with proper movement restriction will be useful to help them learn a standard posture more quickly and easily.

The fourth hypothesis presented in this paper stated that the cognitive development of students trained using VR integrated weld training would be similar to that of students trained using traditional methods. In most cases cognitive development was not found to be significantly different between the VR50 and TW groups. However, it should be noted that two cognitive development areas favored the VR50 group, while similarly, two cognitive development areas favored the TW group.

Two of the differences in cognitive development were related to the 3F weld type. The cognitive evaluation of the 3F weld type indicated that the VR50 group had significantly higher analysis development but the TW group had significantly higher application development. The knowledge development and comprehensions development was not distinctly different for either group. This indicates a difference in the focus of learning between groups for this weld type, however one cannot conclude that this constitutes a greater overall cognitive development for either group.

The remaining two differences in cognitive development were found in the 2F and 3G weld types. Specifically the VR50 group was found to have significantly higher comprehension development for the 2F weld type and the TW group was found to have significantly higher comprehension development for the 3G weld type. These differences in development did not appear to cause a significant downstream differentiation in analysis development or application development for either group. Once again, it is not possible to conclude that a greater overall

cognitive development existed for either group. In addition, there appears to be no correlation between cognitive development and certification, indicating that students trained using the virtual integrated system had effectively the same level of cognitive development as did students trained using traditional methods. Results of this study revealed that the overall scores for cognitive development did not appear to be significantly better for either training group. This finding is in agreement with the original hypothesis.

The fifth and final hypothesis proposed that MWL would not be significantly different between welders trained using VR integrated versus traditional systems. The results of the MWL analysis in this study clearly showed no difference between participants in the VR50 and TW groups.

Study Limitations and Implications for Future Work

The primary limitation of this study lies in the fact that the optimal ratio of VR to real world (traditional) training is unknown. The findings of this study show that 50% VR integration shows some benefit, however this does not necessarily indicate that this ratio is optimal. Another limitation of this study involves the length of training time allotted. Two weeks is a realistic amount of time for a condensed weld school attempting to teach the 2F, 1G, 3F, and 3G weld types, however an investigation of welders in training over a longer period of time could provide additional information to be used in defining how weld complexity influences student attitudes and interactions with VR versus real world welds.

In the authors' future work, a 100% VR weld training school will be examined and the progress of all students tracked and tested with an AWS CWI for all the weld types discussed in

this paper. This proposed study will help further define the boundaries and potentials of modern VR technology.

CONCLUSIONS

The results of this study strongly support the use of VR integrated training. VR integrated training was shown to lead to comparable cognitive development when compared to traditional weld training. In addition, kinesthetic memory development in students using VR integrated weld training was demonstrated to be different to that found in students using traditional weld training for certain weld types (specifically 1G). This difference in kinesthetic development was a contributing factor in the superior performance outcomes produced by participants in the VR integrated group. This study provides evidence (in terms of actual earned AWS certifications) that VR integrated weld training can lead to performance outcomes that are as good as or in some cases superior with a relative significance to traditional weld training. Furthermore, this study demonstrated that the training of qualified welders can be accomplished in a significantly shorter period of time, and with a marked reduction in the use of costly consumables when VR integrated training methods are employed. Perhaps most importantly, these findings may be applicable to a variety of other fields such as medical or military training. The authors propose that future design of VR training systems consider how the system itself directs both cognitive and physiological development of the user. The results of this study have demonstrated that for tasks requiring both physical and cognitive development, VR integrated training is a valuable instructional technology.

ACKNOWLEDGMENTS

KEY POINTS

- VR integrated welding is an effective method for weld training

- The physical development differences between traditional and VR integrated welding are in no way detrimental to task performance.

- There are no significant differences in cognitive development between the traditional and VR integrated students.

REFERENCES

Beauchamp, Y., Marchand, D., & 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 joints. In *Proceedings of the annual international Occupational Ergonomics and Safety Conference*, 197–200.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives: Cognitive Domain*. New York, NY: David Mckay.

Cave, H. (1918). US Patent No. 1286529. U. S. P. a. T. Office.

Choquet, C. (2008). Arc+: Todays virtual reality solution for welders. In *International Conference of Safety and Reliability of Welded components in Energy and Processing Industry on the occasion of the 61st IIW Annual Assembly Conference.*

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

Ebert, A., Deller, M., Steffen, D., & Heintz, M. (2009). "Where Did I Put That?" – Effectiveness of Kinesthetic Memory in Immersive Virtual Environments. *Proceedings of the 13th International Conference on Human-Computer Interaction*, San Diego, CA.

Fast, K., Gifford, T., & Yancey, R. (2004). Virtual Training for Welding. In the *3rd IEEE and ACM International Symposium on Mixed and Augmented Reality,* 298–299.

Gallagher, A.G., & Satava, R.M.(2002). Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. *Surgical Endoscopy*, 16(12): 1746-1752.

Ganai, S., Donroe, J.A., St Louis, M.R., Lewis, G.M., & Seymour, N.E.(2007). Virtual-reality training improves angled telescope skills in novice laparoscopists. *The American Journal of Surgery,* 193(2): 260-265.

Grantcharov, T. P., Kristiansen, V.B., Bendix, J., Bardram L., Rosenberg J., & Funch-Jensen, P. (2004). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery,* (91): 146–150.

Gurusamy, K., Aggarwal, R., Palanivelu, L., & Davidson, B.R. (2008). Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. *The British Journal of Surgery,* 95(9): 1088-1097.

Hart, S.G., Wickens, C.D.(1990). *Manprint, an approach to systems integration*. H.R. Booher (Ed.).Van Nostrand Reinhold, New York.

Hart, S. G. (2006). Nasa-Task Load Index (Nasa-TLX); 20 Years Later. *Human Factors and Ergonomics Society Annual Meeting Proceedings,* 50: 904-908.

Hart, S. G*., &* Staveland, L. E*.(*1988)*.* Development of a NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*. Human Mental Workload*(*pp.* 139–83)*,* P. S.Hancock, & N.Meshkati(Ed.)*. Amsterdam: North-Holland.*

Herberts, P., & Kadefors, R. (1976). A Study of Painful Shoulder In Welders. *Acta Orthopaedica,*47: 381-387.

Jordan, JA., Gallagher, A.G., McGuigan, J., McGlade, K., & McClure, N.(2000).A Comparison between Randomly Alternating Imaging, Normal Laparoscopic Imaging, and Virtual Reality Training in Laparoscopic Psychomotor Skill Acquisition. *The American Journal of Surgery,* 180(3): 208-211.

Kannan, T., & Murugan, N. (2006). Effect of flux cored arc welding process parameters on duplex stainless steel clad quality. *Journal of Materials Processing Technology,* 176(1-3): 230-239.

Keir, P.J., & 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.

Lapointe, J. F., & Roberts, J. M. (2000). Using VR for efficient training of forestry machine operators. *Education and Information Technologies,* 5(4): 237-250.

Mavrikios, F., Karabatsou, V., Fragos, D., & Chryssolouris, G.(2006). A prototype virtual reality-based demonstration for immersive and interactive simulation of welding processes. *International Journal of Computer Integrated Manufacturing, 19*(3), 294-300.

Mellet-d'Huart (2006). A Model of (en) Action to Approach embodiment: a Cornerstone for the design of virtual environments for learning Virtual reality. *Journal of Virtual reality,* 10(3-4): 253-269.

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

Park, M., Schmidt, L., Schlick, C., & Luczak, H. (2007). Design and Evaluation of an Augmented Reality Welding Helmet. . *Human Factors and Ergonomics in Manufacturing,* 17(4): 317–330.

Perotto, A. (1996). *Anatomic Guide for the Electromyographer: The Limbs and Trunk*. Springfield, IL: Charles C. Thomas.

Reyes-Valdés, F.A. , Torres-Treviño, L. , Lopez-Cortéz, V.H. , de la Cruz-Márquez, F., & Briceño-Ramirez, F. (2006). Correlation between gmaw process and weld quality parameters: A neural network approach applied in the automotive industry. *FABTECH and AWS 2006.* Atlanta, Georgia.

Seymour, N.E., Gallagher, A. G., Roman, S. A., O'Brien, M.K., Bansal, V.K., Andersen, D.K., & Satava, R.M. (2002). Virtual Reality Training Improves Operating Room Performance. *ANNALS OF SURGERY,* 236(4): 458–464.

Sheik-Nainar, M. A., Kaber, D.B. , Chow, MY. (2005). Control gain adaptation in virtual reality mediated human–telerobot interaction. *Human Factors and Ergonomics in Manufacturing & Service Industries,* 15(3): 259-274.

Stefanidis, D., Korndorffer, J. R. Jr, Markley, S., Sierra, R., Heniford, B. T., Scott, D. J. (2007). Closing the gap in operative performance between novices and experts: does harder mean better for laparoscopic simulator training? .*Journal of the American College of Surgeons,* 205: 307-313.

Torkington J, Smith, S.G., Rees, B.I., & Darzi, A. (2001). Skill transfer from virtual reality to a real laparoscopic task. *Surgical Endoscopy,* 15: 1076–1079.

White, S.A., Prachyabrued, M., Chambers, T.L., Borst, C.W., & Reiners, D. (2010). Low-cost simulated MIG welding for advancement in technical training. *Virtual Reality*, 1-13.

Wickens, C. D., & Baker, P. (1995). Cognitive issues in virtual reality. *Virtual environments and advanced interface design* (pp. 514-541) T. A. F. Woodrow Barfield III (Ed.). New York: Oxford University Press.

Wilson, M.S., Middlebrook, A., Sutton, C., Stone, R., & McCloy, R.F.(1997) MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance. *Ann R Coll Surg Engl,* 79: 403–404.

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the presence questionnaire. *Teleoperators and Virtual Environments,* 14: 298–312.

Witmer, B. G., & SINGER, M.J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Teleoperators and Virtual Environments,* 7(3): 225-240.

Wormell, D., & Foxlin, E.(2003). Advancements in 3D Interactive Devices for Virtual Environments. *ACM International Conference Preceeding Series; Proceedings of the Eurographics Workshop on Virtual Environments*, 39:47–56, 2003.

Ye, N., Banerjee, P., Banerjee, A., & Dech, F.(1999). A comparative study of assembly planning in traditional and virtual environments*. IEEE Transactions On Systems Man And Cybernetics, Part C—Applications and Reviews,* 29(4): 546-555.

Zanbaka, C. A., Lok, B. C., Babu, S. V., Ulinski, A. C., and Hodges, L. F. (2005). Comparison of path visualizations and cognitive measures relative to travel techniques in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics,* 11: 694-705.

Zipp, P. (1982). Recommendations for the standardization of lead positions in surface electromyography. European *Journal of Applied Physiology and Occupational Physiology,* 50(1): 41-54.

ABOUT AUTHORS

Richard T. Stone Ph.D.

Richard T. Stone Ph.D. AEP is an assistant professor at Iowa State University in the departments of Industrial and Manufacturing Systems Engineering, and Mechanical Engineering. Dr. Stone received his Ph.D. in industrial engineering from the State University of New York at Buffalo in 2008. Dr. Stones' current research focuses on enhancing human performance with consideration to cognitive, biomechanical, physical, environmental and technological factors. Current domain areas included augmented reality and tele-robotic systems and human power generation design.

Kristopher Watts

Kristopher Watts, BS is a second year Ph.D. student studying under Dr. Stone and Dr. Peters in the department of Industrial and Manufacturing Systems Engineering at Iowa State University. His research interest includes cognitive engineering, applied biomechanical engineering, and manufacturing systems engineering.

Peihan Zhong

Peihan Zhong, BS is a second year Ph.D. student studying under Dr. Stone in the department of Industrial and Manufacturing Systems Engineering at Iowa State University. Her research interest includes human performance engineering and human-robot interaction (HRI).

Chen-Shuang Wei

Chen-Shuang Wei, MS is a third year Ph.D. student studying under Dr. Stone in the department of Industrial and Manufacturing Systems Engineering at Iowa State University.   Her research interest includes Human Machine Training Systems and Visualization.

List of Figures

Figure 1. ISU welding school: (left) Traditional welding booth; (right) Virtual welding booth.

Figure 2. The four certifiable weld positions in this study, depicted in order of increasing difficulty.
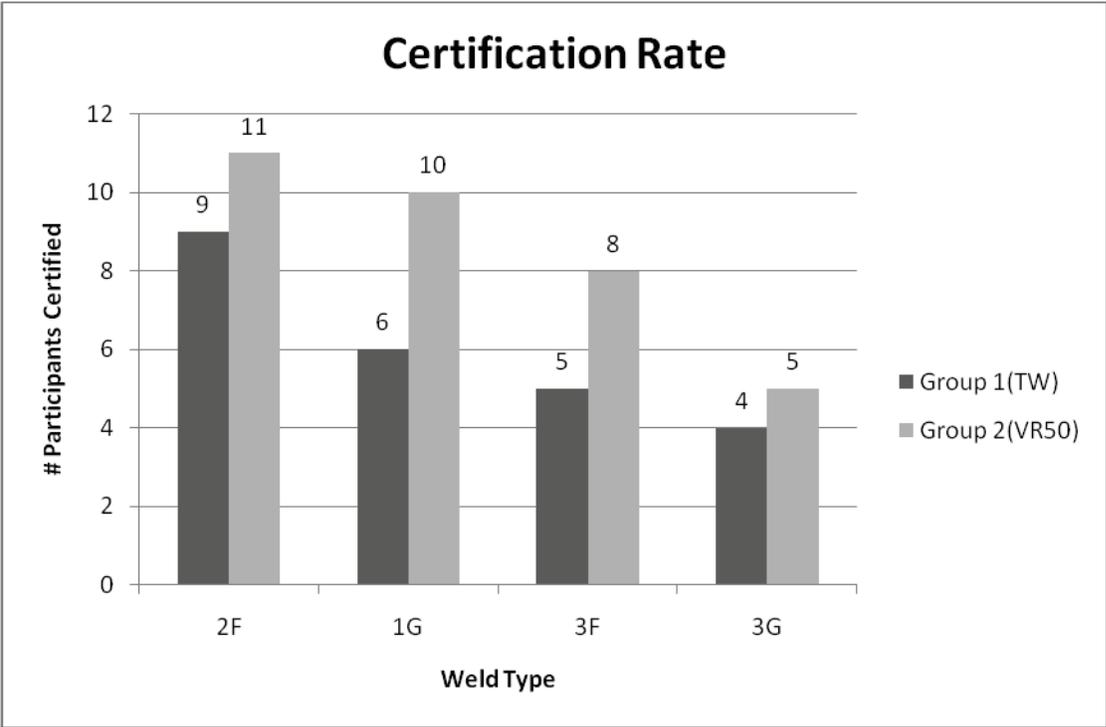
Figure 3. Number of certification awarded by weld type (in order of increasing difficulty).
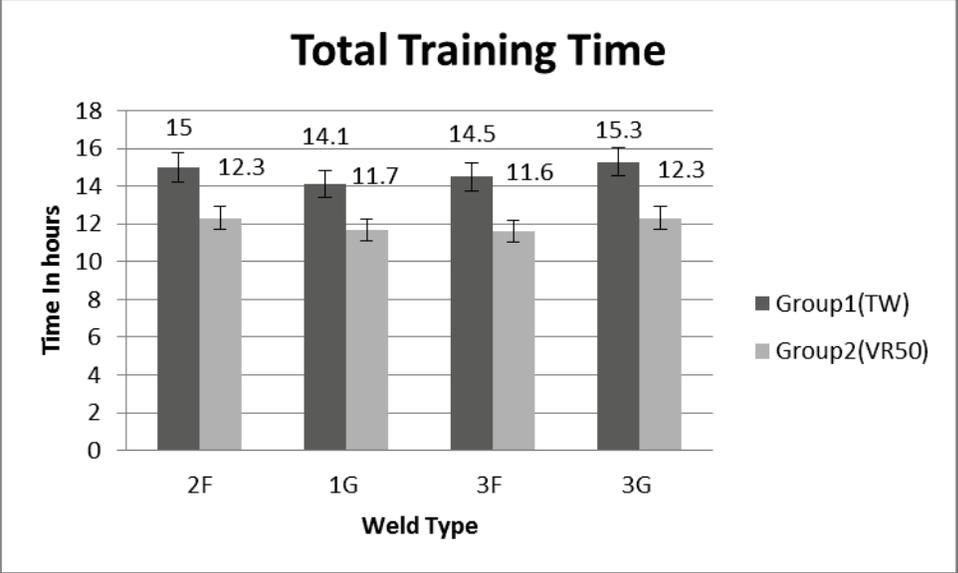
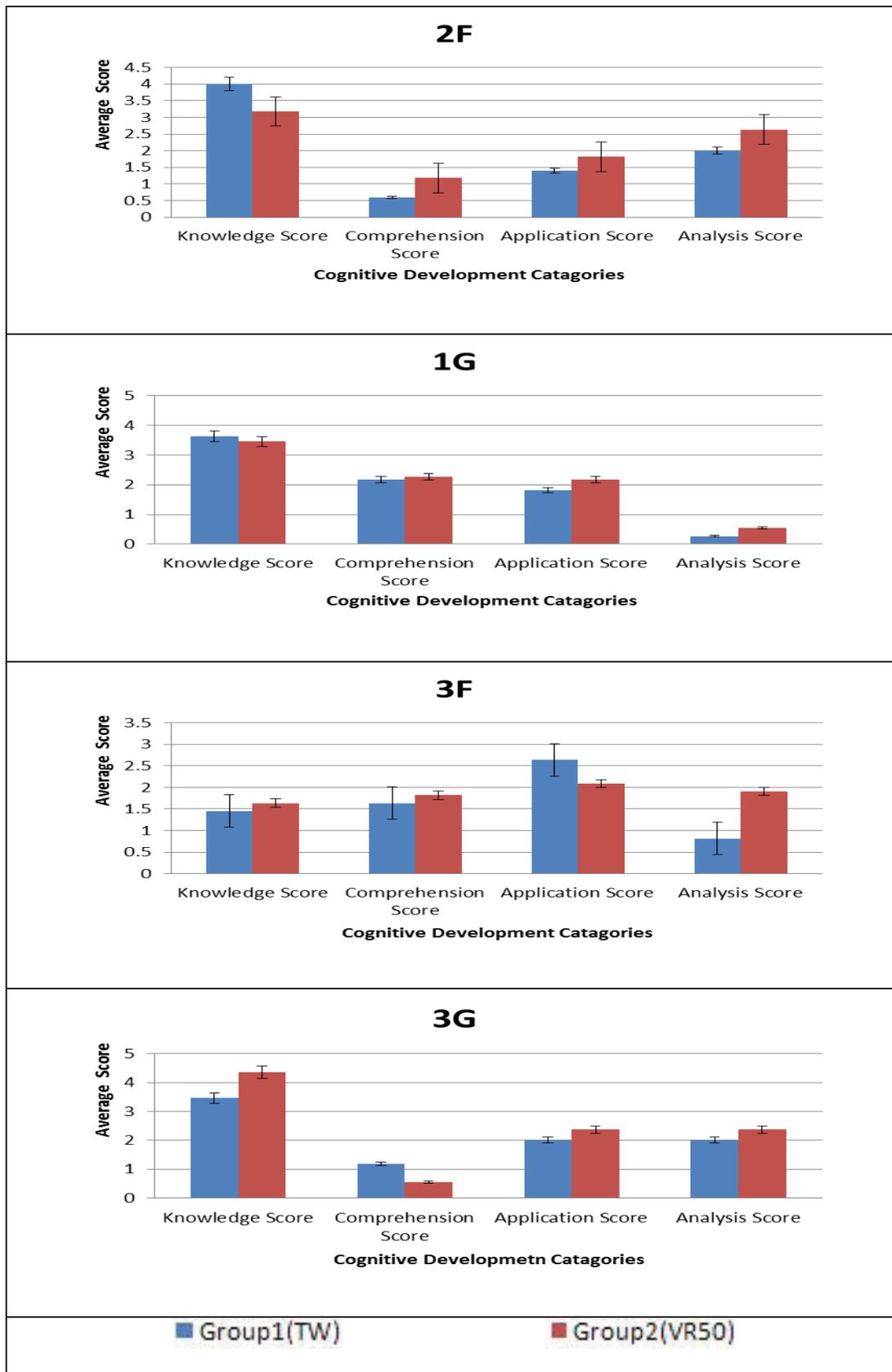*Figure 4. Mean training times by weld type (in order of increasing difficulty).*

Figure 5. Cognitive Scores by category across weld types for the VR50 and TW groups.
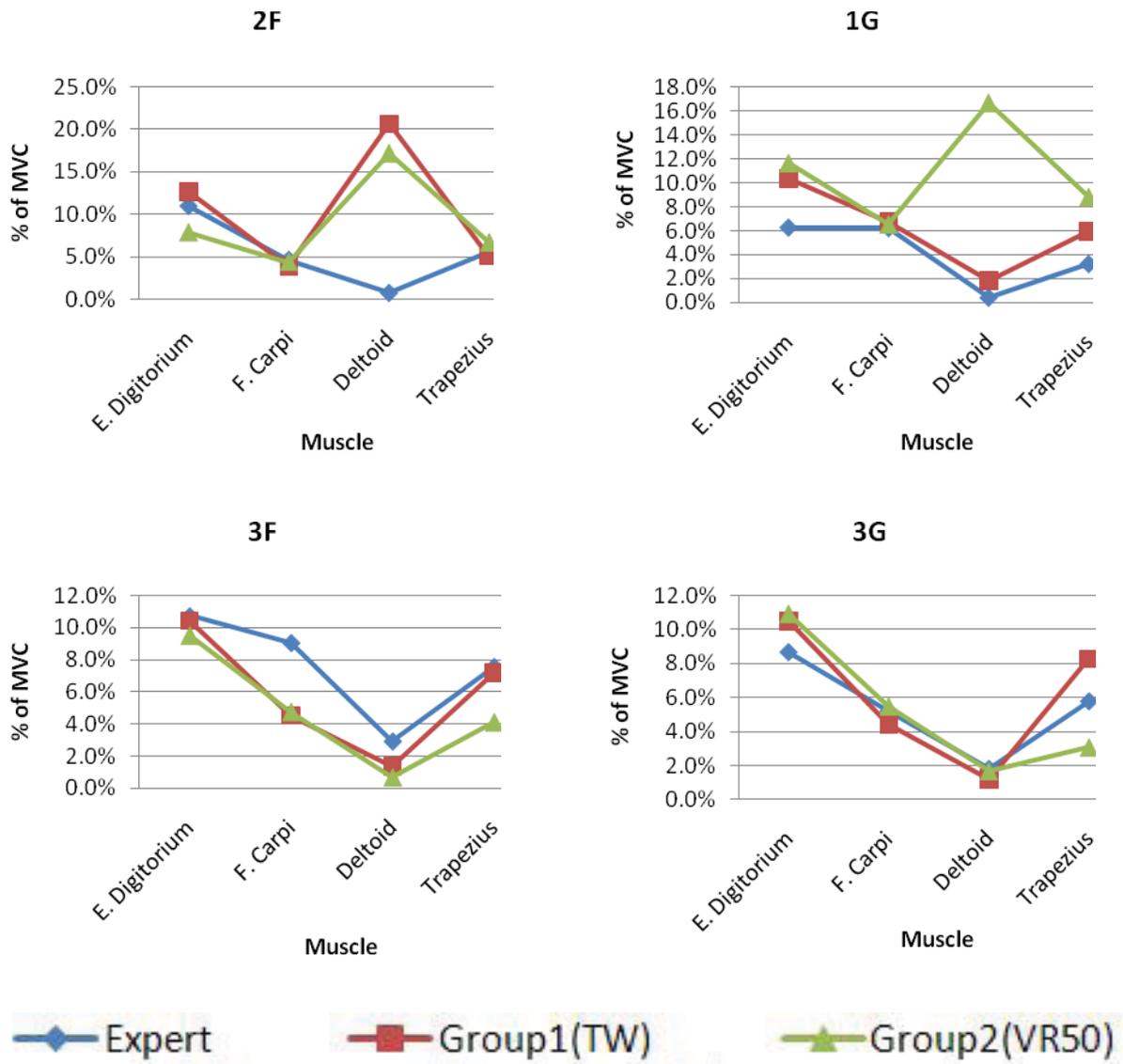
Figure 6. Muscle activity interaction profile by weld type for expert, TW and VR50 welders.