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

Topic choice: Simulation and Virtual Reality

Text/Image Count: Text 5398 words, References 881 words, 10 Tables, 6 Figures

Running Head/Short Title: Full VR vs. Integrated VR Training

Full Virtual Reality Training vs. Integrated Virtual Reality Training: A comparison of performance, feature usage, cognitive and physiological development

Richard T. Stone, **Elease J. McLaurin**, **Peihan Zhong**, and **Kristopher Watts**, Iowa State University, Ames, IA

50 WORD DESCRIPTION

This study evaluated two levels of virtual reality training in terms of feature usage, performance, physical and cognitive impact. Findings showed that the VR-100% training was comparable to VR-50% for low and medium level tasks and proper selection of critical features was a major indicator of success in both groups.

ABSTRACT

Objective: This study compared the impact of a full virtual training program vs. a mixed virtual reality and real-world training program on participants' pass-fail skill completion rate, physical and cognitive development. Additional this study explored how the presence of a visual, real-time feedback feature in a VR simulator training environment impacts participant performance.

Background: Two type of training programs have been developed for virtual reality simulators: full virtual training programs and mixed virtual reality and real-world training programs. There is a need to compare the programs in terms of their training effectiveness. **Methods**: Participants completed welding training either in a full or mixed virtual reality training program. Four weld types were taught which represented different levels of difficulty and required the development of specialized knowledge and skills. 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, and welding certification awards earned. **Results**: The two training programs were comparable in terms of performance for the simple tasks. For the more complex tasks, the mixed training program had better performance. **Conclusion**: Full virtual training is effective for learning simple techniques while more complex techniques require some real-world training. VR simulator feedback features can greatly improve performance, but their use must be balanced with respect to the amount of available attention. **Application**: The findings expand understanding of the impact of levels of VR training and the impact of feedback features.

INTRODUCTION

What does it take to become an expert? Gallagher, Ritter, Champion, Higgins, Fried, Moses, Smith & Satava (2005) proposed that the psycho-motor skills one needs to master in order to become an expert in a skill include knowing "what to do, what not to do, how to do what they need to do, and how to identify when they have made a mistake." When developing skill training programs, such as occupational training programs, it is important to take advantage of the technological advancements which are available in order to make training more efficient. One major area of technological advancement is the development of virtual reality (VR) simulator technology. The VR simulators used for occupational training range in capabilities and features (Choquet 2008; Mellet-d'Huart 2006; Satava 2001; White, Prachyabrued, Chambers, Borst & Reiners 2010). Some of the more advanced simulators, such as the VRTEX 360 welding simulator, provide the user with a fully immersive environment which includes manipulation of physical tools and high fidelity imaging (due in large part to advanced physic engines and graphics rendering capabilities) coupled with an ability to achieve realistic kinesthetic movements (due to magnetic displacement technologies allowing for 6 DOF movement) and both real-time and post-task feedback. The development of virtual reality (VR) simulator technology is largely fueled by a desire to reduce the material, time, and expert accessibility costs that are associated with traditional training methods. It has been suggested that VR simulator is effective at producing "pretraining novices" in that they can teach some learning aspects but not others (Gallagher, et al. 2005; Van, Ritter & Smith 2006; Di Giulio, Fregonese, Casetti, Cestari, Chilovi, D'Ambra, Di Matteo & Delle 2004). As a result, there is a need to understanding how to develop training programs which best capitalize on the technological advancements in the VR simulator designs.

Types of VR simulator training programs

The types of training programs based on VR simulators which are currently being researched can be classified into two categories. The most prevalent type of training program only involves using the simulator and thus is fully completed in the virtual world; for future reference we will refer to this type of training as "VR-100". Examples of this type of training programs are prevalent, particularly in the area of laparoscopy surgery (e.g. Gallagher & Satava 2002; Grantcharov, Kristiansen, Bendix, Bardram, Rosenberg & Funch-Jensen 2004; Ye et al., 1999; Wilson et al., 1997; Munz et al., 2004; Torkington et al., 2000; Gurusamy et al., 2008). The second type of training program involves integrating the virtual reality training technology with real-world training; for future reference we will refer to this type of training as "VR-50". Little work has focused on the VR-50 type of training program, and none has focused on comparing the VR-50 to the VR-100. The goal of the present study is to address this need by comparing a VR-50 training program to a VR-100 training program in terms of the post-training performance of the participants. The performance is defined in terms of pass-fail skill completion rate, physical development, and cognitive development. An additional and equally important goal of the study is to explore how the presence of a visual, real-time feedback feature in the VR simulator training environment impacts participant performance.

Previous study focus

In previous work completed by the authors, a VR-50 training program was compared with a traditional training program. For this study, the VRTEX 360 welding simulator was selected because it was capable of providing a level of realism and kinesthetic feedback appropriate for the study. The authors do not endorse this product over others that possess the aforementioned capabilities. Additionally, the domain of welding was selected due to the complex nature of physical movement and kinesthetic memory that are essential for execution of welding tasks. Jobs requiring regular performance of highly nuanced actions such welding necessitate the honing of highly specific physical movements. According to previous real-world research in the domain of welding (Beauchamp, Marchand & Galopin 1997; Herberts & Kadefors, 1976) and feedback gathered from experts, muscles that are of significant importance to welding performance include the deltoid, trapezius, extensor digitorum, and the flexor carpi ulnaris. Regarding physical development, 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 (Keir & MacDonell 2004). 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.

Results from this study showed that VR-50 welding training had numerous advantages over the traditional welding training including: increased weld quality, higher certification rates, reduced training time, improved kinesthetic learning and reduced costs for the simpler welds (Stone, R., Watts, K., Zhong, P. & Wei, C. 2011). In particular, before the VR-50 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 VR-50 trainees achieving the pre-set mastery level with the VR simulator for the simpler welds. This mastery level indicated that at that time they were sufficiently prepared to successfully complete these welds. Based on these trends, it was expected that if the VR training was isolated from the real world, the results of the VR-100 and the VR-50 would be similar. However, this expectation had not at that time been demonstrated experimentally and it remained to be understood how and why the features of the VR simulator contributing to the development of the trainee, specifically in terms of kinesthetic development and cognitive development. Nonetheless, the results of this previous work were promising, because they added to the growing body of evidence suggesting that there are aspects of skill learning that can be acquired without the mediation of real-world training (Ye, Banerjee, Banerjee, & Dech 1999; Wilson, Middlebrook, Sutton, Stone & McCloy 1997; Gurusamy, Aggarwal, Palanivelu, & Davidson 2008; Seymour, Gallagher, Roman, O'Brien, Bansal, Anderson & Satava 2002; Ganai, Donroe, Louis & Seymour 2007)

Follow-up study focus

Due to the observations and questions raised during the authors' previous work, the effect of the VR simulator features which allow real-world manipulation of tools and those which provide visual, real-time feedback became of particular interest. A number of studies have focused on how the fidelity of VR influences training efficiency (Kenyon & Afenya 1995; Rose, Attree, Brooks, Parslow & Penn 2000). However, little work had focused on assessing the impact of the real-world manipulation of accessory tools of a VR simulator on both the kinesthetic and cognitive development of the trainee. Prior studies have only focused on the cognitive development (Munz et al. 2004; Seymour et al 2002). Regarding the visual, real-time feedback feature, 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 (Calhoun, Draper, Abernathy, Patzek, & Delgado, 2005; Alexander, Stelzer, Kaber & Kim 2008; Ververs & Wickens 1998; Yeh, Merlo, Wickens & Brandenburg 2003).

The focus of this study is to measure the impact of the VR simulator training, in the form of VR-50 and VR-100 training, on the physical and cognitive development as well as the pass-fail skill completing rate. The authors hypothesized that: (1) a VR-100 training program that is comparable to the VR component of a VR-50 training program will produce comparable results in terms of the kinesthetic and cognitive learning that is developed, (2) for both the VR-100 and the VR-50 training program, the fidelity of the VR simulator in simulating welding conditions and the total training time will serve as limiting factors in how well a given weld type can be learned and (3) The selection of the type and number of real-time feedback indicators will be linked to the successful training of both the VR-50 and the VR-100 trainees.

METHODS

Pre-Study

As mentioned earlier, pervious work related to virtual integrated welding was conducted by the authors. During this time the principle investigator attended a formal weld school and became a certified welder. This experience was part of an ethnographic study that was used (along with follow up observational studies) to define the pedagogical, ergonomic and technological aspects of weld training. Next, eight expert welders were formally evaluated while conducting the four welds of interest in this study. This data was used to create expert models for comparison with the participants' kinesthetic development. Previous work involved the use of 22 participants divided among two experimental training groups (traditional and 50% virtual reality training). Each group was trained independently for two full weeks. Their progress was then compared. The outcomes of this study indicated that 100% VR training could both have substantial benefits for some welds and would allow researchers to better understand the functional aspects of VR that lead to enhanced skill acquisition.

Experimental Materials and Relevant system features

Two welding schools were constructed-a VR welding school and a real-world welding school. The real-world welding school was sanctioned by the American Welding Society (AWS). The materials stocked in each of the welding booths or generally available for the real-world school are listed in Table 1.

Lincoln Electric Power MIG 350MP welder	Welding table
with SMAW (stick metal arc welding) attachments	Quenching buckets
(2) auto adjusting welding helmets	Flat stock plates
Multiple sets of welding jackets and gloves	Groove plates
Power grinders	7018 electrodes
Slag hammer	Runoff tabs
Wire brushes	

Table 1. List of materials for real-world welding school

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 Welding Trainer with SMAW attachments and multiple sets of welding jackets and gloves. The VRTEX 360 trainer was chosen due to the fact that 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 virtual training systems created a very high fidelity situation, in which 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 could hold in place weld stock (the plates) as would a real-world welding table. In addition to preserving much of the kinesthetic and visual aspects of real-world welding, this VR system provided dynamic visual feedback, in the form of overlays, for known variables associated with welding. These overlays included: travel speed (the appropriate horizontal speed that the welder should move the stinger), work angle and travel angle (the appropriate horizontal and vertical angle the welder should keep between the electrode and the weld coupon), arc length (the appropriate distance the tip of the electrode should be from the weld coupon). It should be noted that the system allowed for travel and work angle as separate overlays, however our usage studies indicated that in all cases students 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 Table 2 and henceforth will be collectively referred to as overlay strategies and individually referred to by the assigned number.

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

Table 2: List of overlay strategies

Participants

There were 21 participants in total (21 males), with 11 in the VR-50 training and 10 in the VR-100 training. All participants had no practical welding exposure and no experience in stick metal arc welding (SMAW) prior to the beginning of the study. The VR-50 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) lbs. The VR-100 group had an average age of 32 (SD=12.64) years, average height of 71.2 (SD=1.1) inches, and average weight of 190.2 (SD=31.6) lbs.

Independent and Dependent Variables

The primary independent variable in this experiment was training type with the two levels of 50% virtual reality training (VR-50) and 100% virtual reality training (VR-100). While the VR-50 group practiced real-world welding 50% of the time and virtual welding the other 50% of the time, the absolute amount of time that the VR-50 group spent in virtual welding training was equivalent to the total amount of time that the VR-100 group spent in virtual welding training. There were three dependent measures in this investigation: certification rate, kinesthetic development, and cognitive development. Also, the visual overlay strategies used for each practice run with the VR simulator was recorded. The performance measures were evaluated for each of four different weld types. The four weld types, in order of increasing difficulty, include the 2F (horizontal filet weld), 1G (flat groove weld), 3F (vertical filet weld), and 3G (vertical groove weld).

The primary cognitive development measure in this study was based on the Crook's consideration of Bloom's taxonomy (Crooks 1988). For this study, experimenters developed specific questions and tests to evaluate cognitive development in each weld type attempted by participants.

For the physical development, electromyography (EMG) feedback for the deltoid, trapezius, extensor digitorum, and the flexor carpi ulnaris muscles were used in this study. To collect this data, electrodes were placed on participants' skin in parallel with the fibers of the muscles of interest 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 to get a baseline for the maximum the participants were willing to exert their muscles. In order to get the MVC for the trapezius and deltoid, the 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 ularis, 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

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. Participants in the VR-50 group spent only 50% of their time training (lectures and practical lab training) under the direction of an AWS CWI for each weld type. The traditional training involved students working with real welding machines and repeatedly practicing the weld types. The remaining 50% of their time was spent training on the VRTEX 360 system and during this time the VR system served as the instructor by proving feedback after every weld and by providing optional use of visual overlay that would guide the user to improve key aspects of their weld including 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 the participants were able to earn a simulator generated quality score of 85% at least twice for a weld, they were permitted to discontinue their VR training time early. 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-100 group spent all of their time learning in the same VR situation as did the members of the VR-50 group but without any time spent in the commission of real world welds. For any given weld type (2F, 1G, 3F, 3G), participant in the VR-100 group were allowed to practice in the VR environment for approximately the same amount of time that there VR-50 counter parts did. For example, the average participant in the VR-50 group spent approximately 6.3 hours in their VR training for the 3G weld, so the training limit for those in the VR-100 group was set at 6.3 hours with a 15 percent tolerance. This method ensured that the way in which the participants utilized VR and VR instructional features would be directly comparable between groups. Both the VR-50 and VR-100 groups were given take home training materials and instructional videos on welding so as to supplement their study opportunities at home.

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. Once completed, the test pieces underwent a visual inspection by the CWI on site. 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 pretest practice and the final testing plate, participants were fitted with electrodes so that experimenters could record EMG data while participants conducted their welds. Immediately following each certification test for all four weld types, participants were given a written cognitive test related to the welding unit used and the weld type they had just performed.

RESULTS

For this study, two welding training methods were compared: VR-50 and VR-100. The performance measures observed were the weld certification rates, psychomotor skill development, and cognitive development. Also, the relationship between the strategic use of the visual overlay feedback and the welding performance was determined.

Certification Rate

As summarized in Table 3, 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 VR-50 group had significantly more 3G certifications than the VR-100 group. The certification rates achieved by participants in the two groups for each of the four weld types are shown in Figure 1.

	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: Data analysis summary for certification rates

In addition, regardless if the weld passed or not, each of the weld tests was accompanied by an overall weld 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 4. No significant difference was found in quality between the two groups for any of the weld types.

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	0.5916	88	10	0.9248	81	16	0.2670	61	24	0.2787
VR 100	89.7	3		89.5	2		71.8	19		53	19	

Table 4: Summary of quality data

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). The development was assessed for each of the weld types for expert welders and both of the experimental groups (VR-50 and VR-100). Figures 2-5 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 psychomotor skill development in this study. The results of the MANAOVA for the 2F weld type show a significant difference between the expert, VR-50, and VR-100 groups. However, post-hoc tests reveal no significant difference between any two conditions. The 1G weld type MANAOVA also revealed that there was a significant difference between the three conditions. Post-hoc tests revealed that the VR-50 and VR-100 did not differ from one another. However, both of the VR-50 and VR-100 groups were found to be significantly different from both the expert group. MANAOVA 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 5.

Weld Types	MANOVA results	Expert vs. VR-50	Expert vs. VR-100	VR-50 vs. VR-100
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: Analysis of muscle interactions

Cognitive Development

Cognitive development was measured across the four categories of the Crook's consideration of the Bloom's taxonomy. These categories are, in order of increasing understanding, knowledge, comprehension, application, and analysis. In order to determine if the VR-50 and VR-100 groups were significantly different it was necessary to first determine if the data was 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 6 below. Each test was conducted using α = 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 VR-50 group performed significantly better than the VR-100 group.

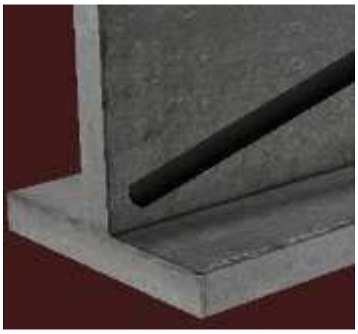
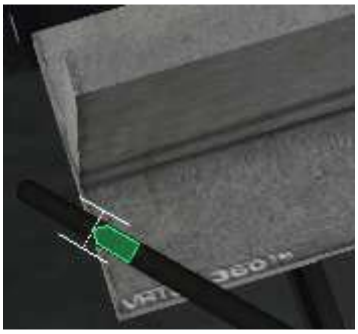
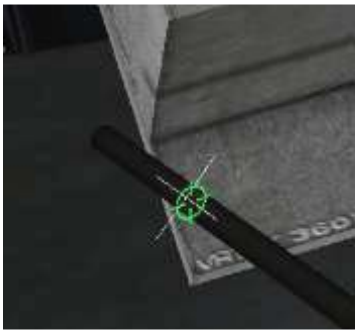
Crook's Consideration of Bloom's Taxonomy T-Test Values

Test Question Type	VR-100 Mean Score	VR-50 Mean Score	VR-100 SD	VR-50 SD	T Ratio	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***

Table 6: Portion of T-test results for Crook's test analysis

Overlays

Also of interest for this study was determining if the use of the visual feedback in the form of the overlays had any impact on the performance measures. There were eight possible strategies for the overlay use. These eight strategies are shown in Table 7 and will henceforth be referred to by the assigned number.

1	**2**	**3**
		
4	**5**	**6**

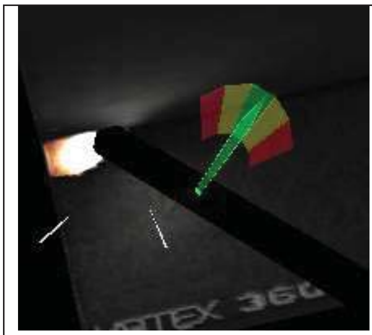
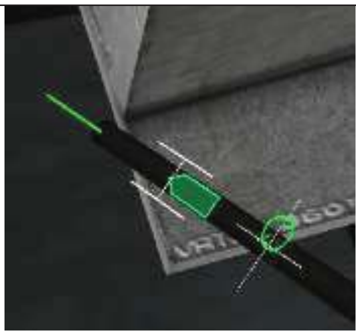
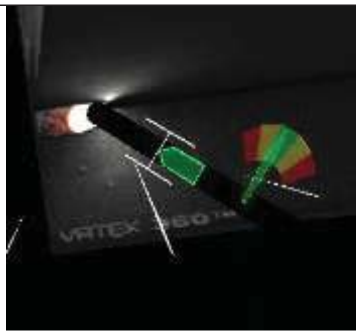
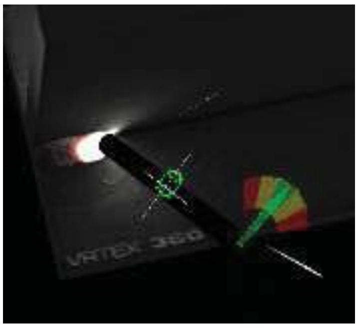
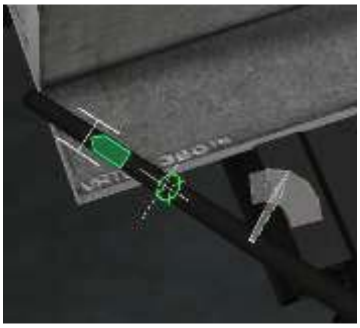
		
7 	**8** 	**1-No overlay** **2-Arc length overlay** **3-Work-travel angle overlay** **4-Travel speed overlay** **5-Arc length and work-travel angle overlays** **6-Arc length and travel speed overlays** **7-Work-travel angle and travel speed overlays** **8-Arc length, travel speed, and work-travel angle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

Table 7: Images of the possible overlay strategies

It was first determined if the VR-50 group used the overlays during their VR training differently than the VR-100 group. For each of the four weld types, a series of T-tests were performed to compare the percentage of usage for each of the overlay strategies for the two groups (VR-100 and VR-50). 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 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 8.

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 8: ANOVA results for significance in weld types

Next, a least squared difference (LSD) 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 9. 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).

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 9: Post-hoc test results for significance in overlay usage

However, looking at only the quality scores do not paint the whole picture. The dominant overlay strategy used for each participant for each weld type needs to be considered. Examining this 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 10 shows the dominate strategy by overlay type as well as the percentage of people who used that overlay type who passed their weld. By examining Table 10, and keeping in mind the previously determined information that overlay 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 widely used and successful strategy.

	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

Table 10: Dominant usage of overlays and associated participant pass percentages

Another trend observed was a decrease in the sampling of overlay strategies as the participants practiced more difficult weld types. Figure 6 shows the number of overlay strategies samples as the participants progressed through the VR welding training.

DISCUSSION

Regarding the comparison of performance measures for the VR-50 and VR-100 welding training, it was hypothesized that (1) a VR-100 training program that is comparable to the VR component of a VR-50 training program would produce comparable results in terms of the kinesthetic and cognitive learning that is developed and (2) for both the VR-100 and the VR-50 training program, the fidelity of the VR simulator in simulating welding conditions and the total training time would serve as limiting factors in how well a given weld type could be learned. A discussion of the results of this study in light of these hypotheses follows for each of the weld types.

2F Weld Type

Based on information gather from the welding experts, of the four welds examined in this study, the 2F weld type was the simplest type of weld to complete. The analysis of the certifications obtained by participants in both groups revealed no significant difference for the 2F welding type. This indicated that for the simplest weld, the VR-100 training was as effective as the real-world training. Also, since there was no difference in the muscle activities for the VR-100 and VR-50 as compared to the experts, the physical interaction with the VR simulator tools was sufficient to develop the same psychomotor skills as experts for the 2F weld type. Regarding the cognitive development for the 2F weld type, the results showed that the reference materials that the VR-100 group had access to (welding CD, welding texts, and information presented through the VR simulator interface) were sufficient to produce equivalent development to what was developed with the welding lectures and the real-world welding exposure.

1G Weld Type

Based on information gather from the welding experts, the 1G weld type was a medium difficulty weld. The analysis of the certifications obtained by participants in both groups revealed no significant difference for the 1G welding type. This indicated that for this medium difficulty weld, the VR-100 training was as effective as the real-world training. Regarding the cognitive development for the 1G weld type, the results showed that the reference materials that the VR-100 group had access to (welding CD, welding texts, and information presented through the VR simulator interface) were sufficient to produce equivalent development to what was developed with the welding lectures and the real-world welding exposure.

There was a significant difference in the muscle activities for the VR-100 and VR-50 groups as compared to the experts, indicating that the physical interaction with the VR simulator tools did not develop the same psychomotor skills as experts for the 1G weld type. However, these differences exist because both the VR-100 and VR-50 groups adopted the same, more beneficial posture for completing the weld. Note that, although the VR-50 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

Based on information gather from the welding experts, 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 for the 3F welding type. This indicated that for this medium difficulty weld, the VR-100 training was as effective as the real-world training. Also, since there was no difference in the muscle activities for the VR-100 and VR-50 as compared to the experts, the physical interaction with the VR simulator tools was sufficient to develop the same psychomotor skills as experts for the 3F weld type. Regarding the cognitive development for the 3F weld type, the results showed that the reference materials that the VR-100 group had access to (welding CD, welding texts, and information presented through the VR simulator interface) were sufficient to produce equivalent development to what was developed with the welding lectures and the real-world welding exposure.

3G Weld Type

Based on information gather from the welding experts, of the four welds examined in this study, the 3G weld type was the most complex weld to complete. The analysis of the certifications obtained by participants in both groups did reveal a significant difference for the 3G welding type, with fewer certifications for the VR-100 group. This information indicates that the VR-100 training was not sufficient for training the participants in how to correctly complete this complex weld. The VR-50 group had a higher certification rate than the VR-100 group, indicating that the fidelity of the VR simulator played a role in the decreased efficiency of the training. Also, although the VR-50 group had a higher certification rate than the VR-100 group, the number of certifications earned by the VR-50 group for the 3G weld was significantly less than the number earns for the other three weld types, indicating that the time spent training on the welds also played a role in the effectiveness of the training.

Also, there was a significant difference in the cognitive development for the 3G weld type, with the VR-50 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-100 and VR-50 as compared to the experts, the physical interaction with the VR simulator tools was sufficient to develop the same psychomotor skills as experts for the 3G weld type.

Overlays

Regarding the relationship between the strategic use of the visual overlay feedback and the welding performance, it was hypothesized that the selection of the type and number of real-time feedback indicators would be linked to the successful training of both the VR-50 and the VR-100 trainees. This study has supported this hypothesis. It was observed that some overlay strategies led to better quality scores than others. The overlay strategies which consistently led to higher quality scores were 4 and 5. The overlay strategies which 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 increases, 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 (Kaber, Kim, Kaufmann, Alexander, Steltzer & Hsiang 2009), up to a point. As participants continued to increase complexity, it was expected that they would reach a “tipping point” where performance no longer increase, but rather decreased. This phenomenon can be considered analogous to clutter in visual displays. Previous research has shown that clutter in visual displays tends to decrease performance (Yeh et al., 2003; Ververs and Wickens, 1998; Alexander, Wickens & Hardy 2005; Alexander et al. 2008). This effect was observed in this study and the tipping point was three overlays.

It should also be noted that 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 becoming accustomed to the VR simulator 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.

CONCLUSION

The results of this study have shown that VR-100 and VR-50 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-100 and VR-50 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 thus required supplementation 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 arch 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 full and integrated virtual reality training in terms of usage, performance, cognitive and physical development.

KEY POINTS

- 100% VR training is comparable to 50% VR training at the low and medium difficulty levels.
- At the high levels of task difficulty, there are significant differences in cognitive development, but not in sensory motor development between the 100% VR and 50% VR training groups.
- The usage of visual overlays is similar among participants regardless of VR training group.

- The least successful welders tended utilized to many visual overlays at one time, thus overtaking their visual attention capabilities and robing focus from transferable cues.

REFERENCES

Alexander, A., Stelzer, E., Kaber, D. B. & Kim, S-H (2008). Bottom-up and top-down contributors to pilot perceptions of display clutter in advanced flight deck technologies. In *Proc. of the 52nd Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

Alexander, A. L., Wickens, C. D., Hardy, T. J. (2005). Synthetic vision systems: The effects of guidance symbology, display size, and field of view. *Human Factors*, 47, 693-707.

Calhoun, G. L., Draper, M. H., Abernathy, M. F., Patzek, M., & Delgado, F. (January 01, 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.

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.

Di, G. E., Fregonese, D., Casetti, T., Cestari, R., Chilovi, F., D'Ambra, G., Di, M. G., ... Delle, F. G. (January 01, 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.

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.

Gallagher, A. G., Ritter, E. M., Champion, H., Higgins, G., Fried, M. P., Moses, G., Smith, C. D., ... Satava, R. M. (January 01, 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.

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.

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

Kaber, D.; Kim, S-H., Kaufmann, K.; Alexander, A.; Steltzer, E.; Hsiang, S. (2009). Modeling the effects of HUD visual properties, pilot experience and flight scenario on a multidimensional measure of clutter. Hampton, VA: NASA Langley Research Center.

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.

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

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.

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

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

Satava, M. (2001). Surgical education and surgical simulation. *World Journal of Surgery*, 25(11), 1484-1489.

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.

Stone, R. T., Watts, K., Zhong, P., & Wei, C. (2011). Physical and Cognitive Effects of Virtual Reality Integrated Training. *Human Factors*. Accepted.

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.

Van, S. K. R., Ritter, E. M., & Smith, C. D. (January 01, 2006). The pretrained novice: using simulation-based training to improve learning in the operating room. *Surgical Innovation*, *13*, 3, 198-204.

Ververs, P. M., & Wickens, C. D. (1998). Head-up displays: Effects of clutter, display intensity, and display location on pilot performance. *International Journal of Aviation Psychology*, 8, 377–403.

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.

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.

Yeh, M., Merlo, J. L., Wickens, C. D., & 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.

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.

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.

Elease J. McLaurin

Elease McLaurin is a first year M.S. student studying under Dr. Stone in the department of Industrial and Manufacturing Systems Engineering at Iowa State University. Her research interests include biomechanics and human performance 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).

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.

List of Figures

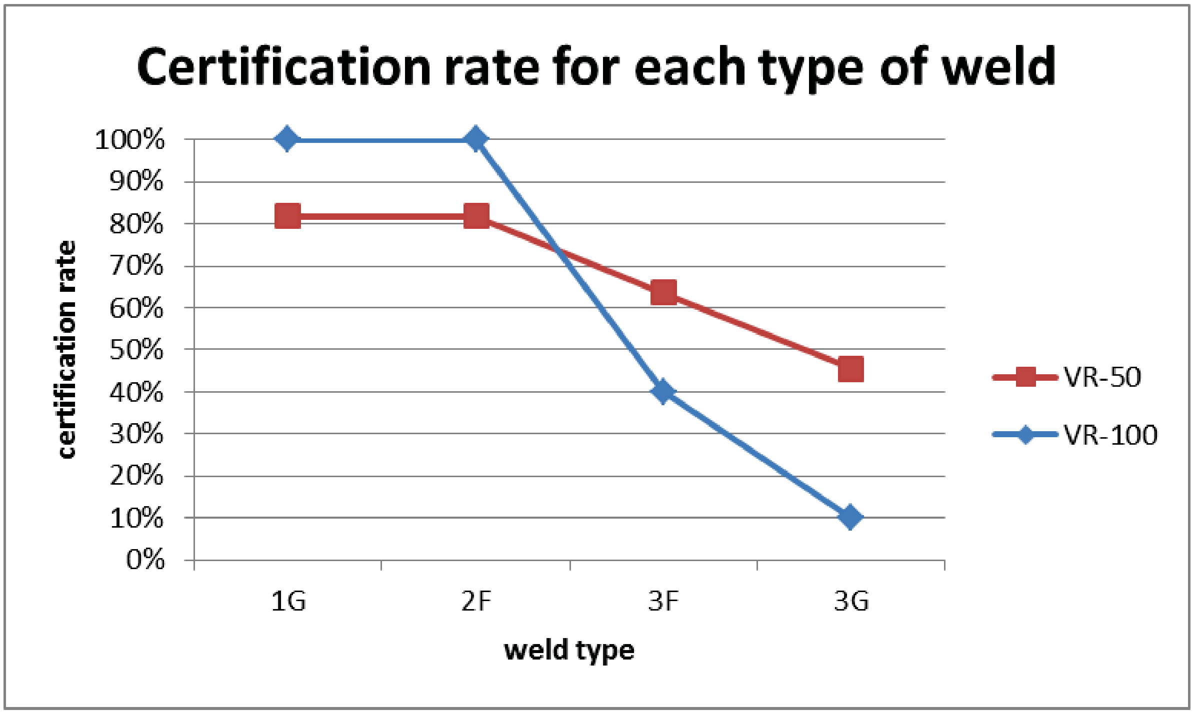


Figure 1. Certifications of each group from each type of weld

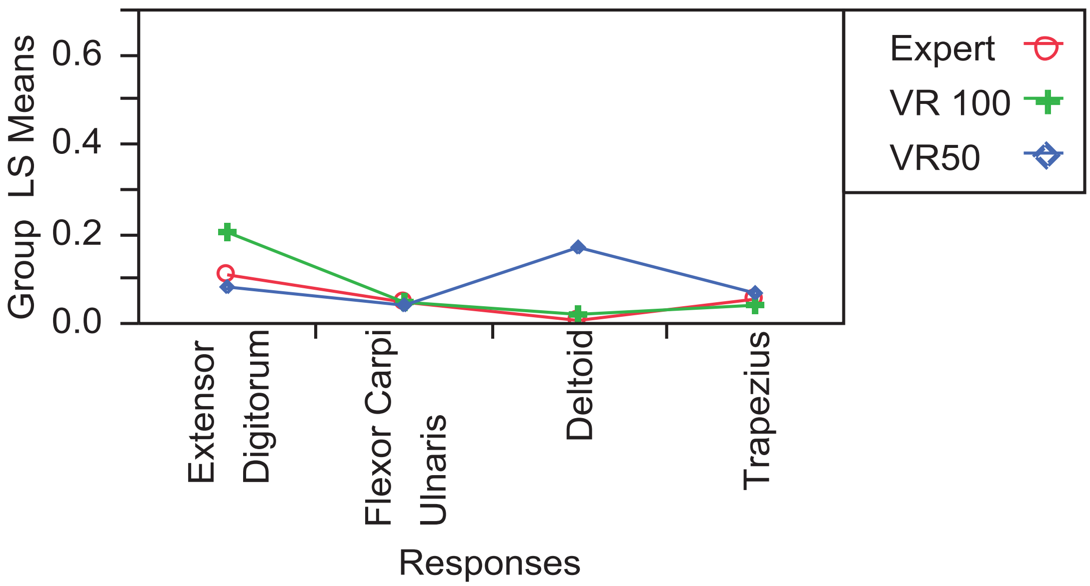


Figure 2. Muscle responses for 2F weld type for expert, VR-50, and VR-100 welders

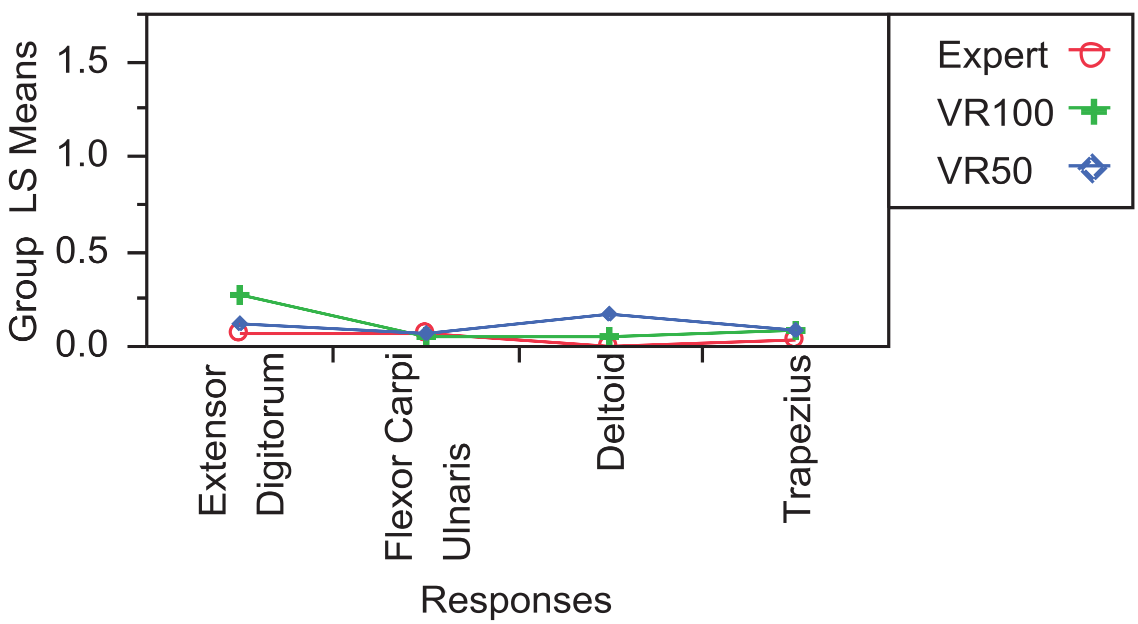


Figure 3. Muscle responses for 1G weld type for expert, VR-50, and VR-100 welders

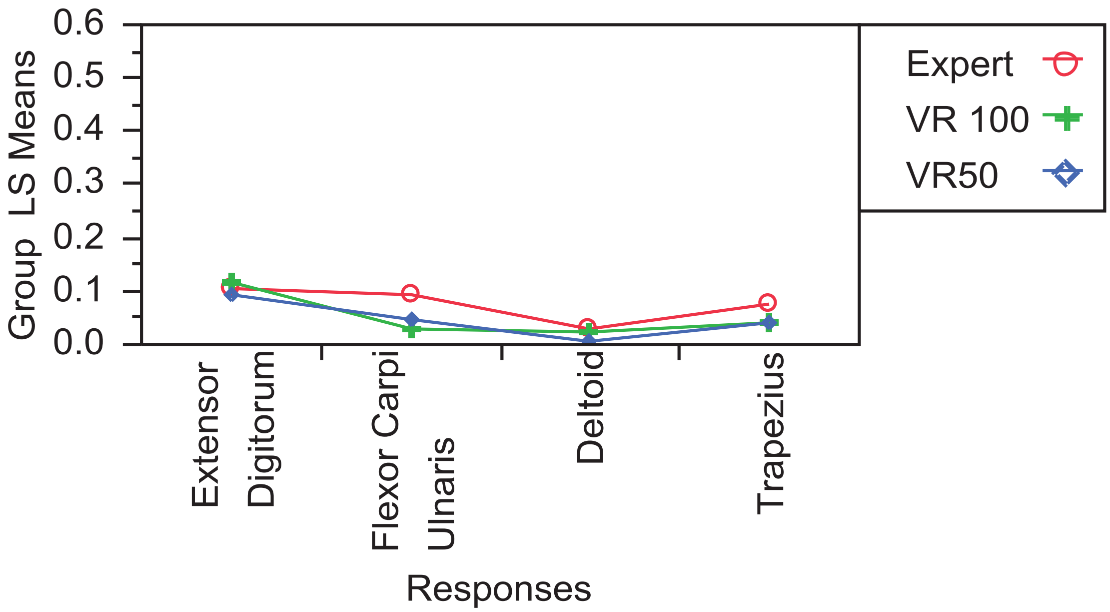


Figure 4. Muscle responses for 3F weld type for expert, VR-50, and VR-100 welders

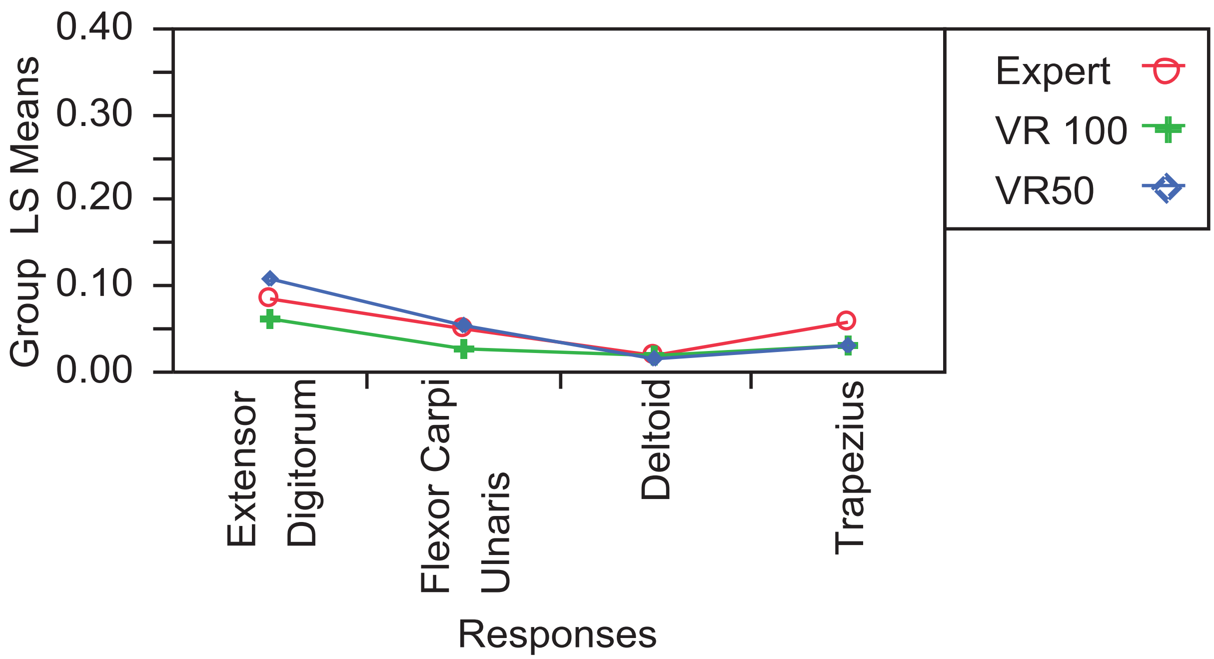


Figure 5. Muscle responses for 3G weld type for expert, VR-50, and VR-100 welders

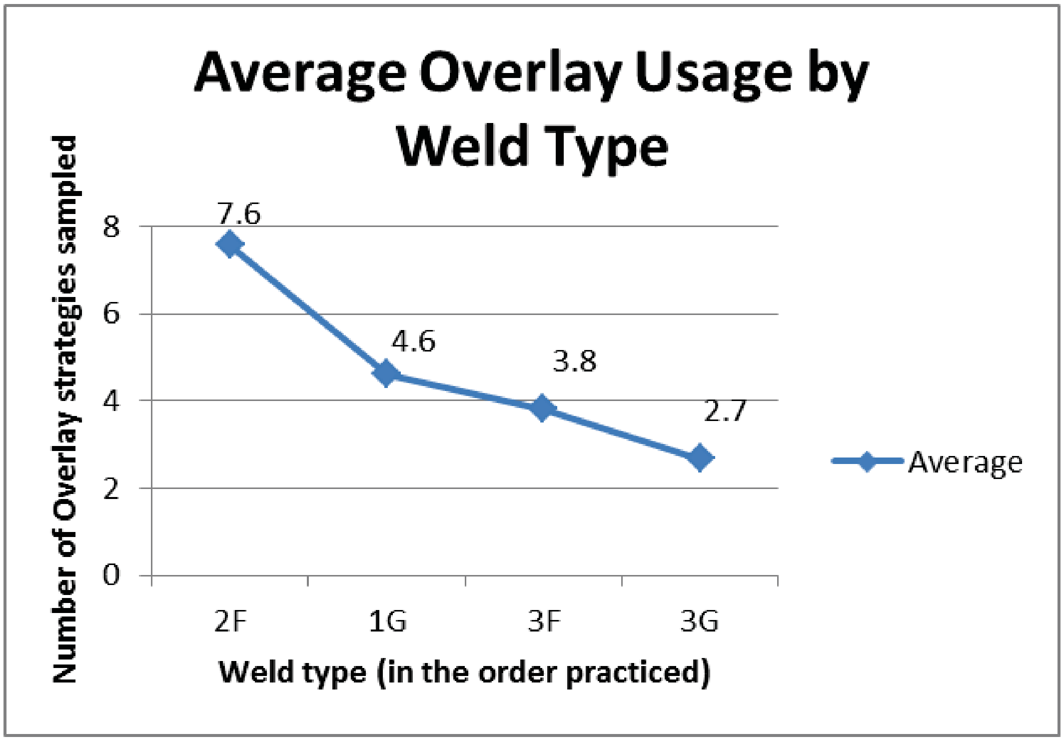


Figure 6. Trend of sampling as weld training progressed

.