ERGONOMICS, 1993, VOL. 36, NO. 7, 777–784

Rapid Communication

Transfer of training from virtual reality

J. J. KOZAK, P. A. HANCOCK, E. J. ARTHUR and S. T. CHRYSLER

Human Factors Research Laboratory, University of Minnesota,
60 Norris Hall, 172 Pillsbury Drive, S.E., Minneapolis, MN 55455, USA

Keywords: Virtual reality; Training; Human factors; Human–computer interaction.

This experiment compared the value of real-world training, virtual reality training, and no training in the transfer of learning to the same task performed in real-world conditions. Results provide no evidence of transfer from a virtual reality training environment to a real-world task. There was no significant difference between the virtual reality training group and the group that received no training on the task. The group that received real-world training performed significantly better than both of the other two groups. The results question the utility of virtual training and suggest that in the present configuration, individuals learn performance characteristics specific only to the virtual reality context. Needed improvements to virtual reality for the purpose of enabling the transfer of training are indicated.

1. Introduction

Training is a major expense in system operation. The time and resources associated with the development of competence in controllers and maintenance personnel represent a substantial component of overall costs. Maintenance itself requires critical down-time periods in which productivity is frequently diminished. Consequently, alternative and improved methods of training, which also provide augmented support for operation, promise to provide great return on investment. Traditional computer-based training programs have provided such alternatives (Flexman and Stark 1987). The emerging properties of virtual reality provide an opportunity for immersion in three-dimensional, computer-generated worlds and hence a new window on the problem of operator training.

Building real-world (RW) simulation environments, such as flight simulators, is expensive and time consuming. However, modelling an environment in virtual reality (VR) is relatively less expensive and more flexible (Baum 1992). Permutations of a single world are easy to create and consume only computer memory space rather than physical space. As the principal operating cost of VR is software, the ability to re-design and re-configure is less expensive than traditional approaches such as those commonly used in aviation (see also Warren 1992). Due to the modularization of the software used, creation of each additional world becomes progressively less expensive.

Virtual reality systems have been portrayed as valuable training devices with both military and industrial applications (Baum 1992). VR systems represent new technologies in which initial effort has been directed toward existence proofs (Kruger 1991) and evaluation of input devices (Sturman 1992). Recently, Wells (1992) has discussed human factors aspects of virtual reality. However, to our knowledge, there are no empirical studies which have specifically focused on human factors and ergonomics issues relevant to training in virtual reality. Consequently, the aim of the present study is the evaluation of virtual reality as an environment for training for a real-world task. To examine this problem, the present work utilized a classic transfer of training paradigm in which the task primarily demands perceptual–motor abilities (Holding 1987). In industrial terms, the present task can be described as a pick-and-place (PPT) operation.

Because we modelled the real world as veridically as possible in the virtual reality system, we expected to find similar learning curves for real-world and virtual reality training. Further, due to this anticipated learning, we expected that training in the virtual environment would provide an advantage over no training but would not prove as useful as real-world practice in task completion time owing to limitation of the virtual system. These hypotheses are evaluated by examining both training time data and task performance data.

2. Method

2.1. *Subjects*

The participants in this study were twenty-one volunteer adults recruited from the University community. All subjects were right-handed. The ages of subjects ranged from 18 to 59, with an average age of 31·5 years. Six women and fifteen men participated.

2.2. *Apparatus and task*

The real-world task performed by all subjects was a pick-and-place sequence. The subject was seated in front of a row of five cans. Subjects were free to adjust their distance from the table. The subject was positioned so that the second can from the left (Can 2) was directly in front of his or her chest. The cans were empty aluminium cans (2·5 inches diameter and 5 inches tall) which had been painted a neutral colour. The cans were placed 6 inches apart on fixed paper discs. Another row of paper discs, representing the target locations, was positioned six inches beyond the row of starting locations. The overall set-up showing a subject in VR, with the real-world set-up adjacent to them is shown in figure 1.

Participants were asked to grasp the cans one at a time and place them at their target locations. Participants were required to pick up each can using a full-handed grasp around the center portion of the cans using their right hand. Beginning with Can 1 at the far left, the subject moved respective cans to their target locations. After placing Can 5 on its target, the subject returned to Can 1 and moved all the cans back to their starting locations. A trial was defined as the time required to move all the cans to the distal targets and return them to their starting positions. Subjects were instructed to perform this task as quickly as possible and to continue with a trial even if a can was misplaced or knocked down.

2.3. *VR training*

Participants used a right-handed Dataglove™ and a head-mounted display (HMD) Eyephone(s)™ (for more detailed technical specifications see VPL Research Inc. 1990, 1991). While in the environment, a dynamic 3D model of a right hand was visible in addition to the representation of the task apparatus. Grasping a can consisted of intersecting the virtual hand with the virtual can while maintaining the appropriate hand posture. Pilot subjects were used to determine hand postures which would qualify as legal grasps in the virtual environment. This hand posture can be defined minimally

Figure 1. The photograph shows a subject in the virtual reality environment featuring the eyephones and the Dataglove.™ For the purpose of illustration, we have also included the real-world set-up for the pick and place task as shown on the table.

as the posture which exists when the thumb, index, and middle fingers are bent to form a closed circle. The range of allowable grasping postures varied from a closed fist to a position similar to that used to pick up a can in the real world. This wide range was necessary due to the variability in human hand size relative to the Dataglove.™

The VR training environment was coloured and scaled to represent the real-world task environment as veridically as possible. It appeared to the subjects that they were seated facing a virtual table. As virtual reality (VR) is a new technology that few people have used, participants assigned to Group 2 were each given 20 min of individual instruction and practice with the VR system. This familiarization phase occurred before attempting any measured trials for the VR training portion of the PPT experiment. During the familiarization phase, the participant was seated in a swivel chair and shown how to use the equipment. They were taught how to travel in the forward direction, reverse direction, stop, reset their initial position, and grasp objects. They were also shown how to achieve a desired position in the environment, such as facing an object head-on. If subjects felt comfortable with the VR environment in less than 20 min, they were allowed to proceed to the VR training trials. Due to individual differences in body size and preferred starting location in the virtual training environment, some subjects were unable to manipulate all the cans from a single position. Subjects were allowed to reposition themselves in the virtual or physical world in order to reach all cans during the training trials.

2.4. *Design and procedure*

Each subject performed 30 trials of the pick-and-place task (PPT) in the real world following training. The response times on this task served as our performance measure. The participants were randomly assigned to one of three training groups. Group 1 received no training prior to performing the target task. Group 2 received general VR

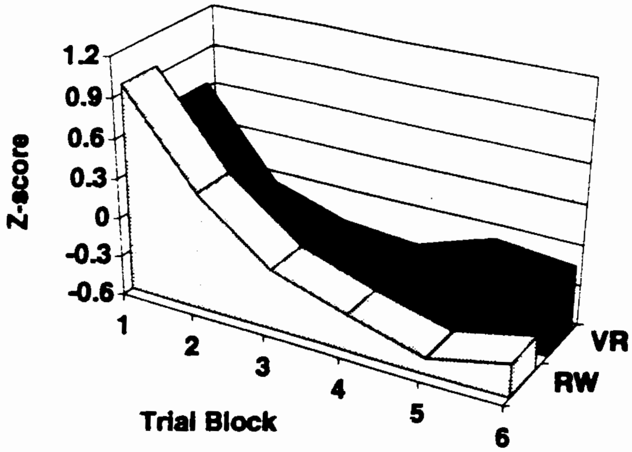

Figure 2. *Z*-scores of training times in real-world and virtual reality environments.

familiarization, followed by 30 trials of PPT training in the virtual environment which took approximately thirty minutes to complete. Group 3 received 30 trials of PPT training in the real world before doing the target task. The decision to use 30 training trials was based on pilot data which showed the point where participants' performance no longer varied by more than 10% per trial in the real-world task. Following the PPT task, subjects in Group 2 completed a questionnaire to give subjective feedback about the VR training. All subjects completed a general information sheet which asked for their impressions of the tasks.

The experimenter observed the participant and timed each trial using a manually-operated digital stopwatch. The experimenter also verified that a full-handed grasp was used. Failure to use the suggested grasp resulted in a mistrial and the trial was excluded from data analysis. Participants were told the elapsed time after each trial. The cans were returned to their initial positions after each of the 30 trials. Participants were allowed to take a break at any time between trials.

3. Results

3.1. *Training data*

One subject was tested and then replaced as all of his response times were greater than two standard deviations above the overall group mean. Completion times for the two training groups (virtual vs. real-world) were compared. The overall mean completion time for one trial in the virtual training environment was 63·45 s. One trial in the real-world task took an average of 5·95 s to complete. Due to this discrepancy of scale, all response times were first converted to *z*-scores based on their respective group means and standard deviations. The 30 training trials were reduced to mean scores for six blocks of five trials each for purpose of analysis. The data are shown in figure 2.

A two-way repeated-measures analysis of variance, with trial block as a within subject factor and training group as a between subjects factor, showed no difference between training groups, $F(1, 12) < 1{\cdot}0$. There was a significant decrease in response times across blocks for both groups, $F(5, 60) = 14{\cdot}914$, MSe = 0·153, $p = 0{\cdot}001$. The training group by blocks interaction failed to reach significance, $F(5, 60) = 1{\cdot}583$, MSe = 0·153, $p = 0{\cdot}179$. Due to the non-significant interaction, follow-up tests on the block main effect were not conducted.

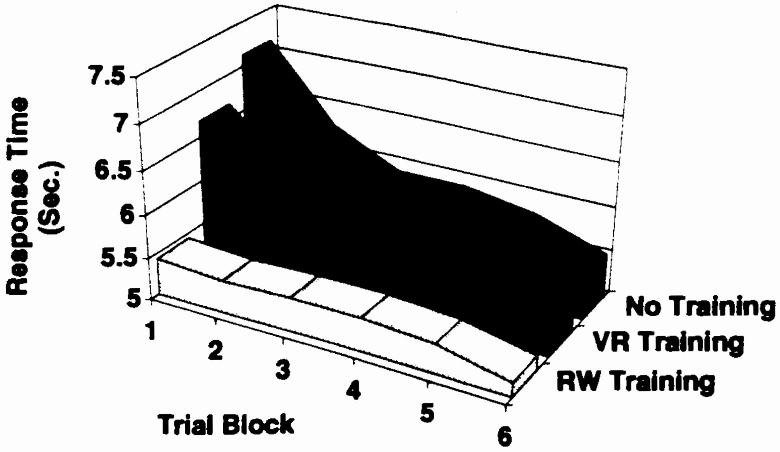

Figure 3. Mean response times on the target task for blocks of five trials by group.

3.2. *Task performance data*

All trials on which knock-downs or illegal grasps occurred were noted and discarded before the analysis.[1] This screening resulted in a loss of 13% of the completion time data. Due to the loss of data from knock-downs and illegal grasps, and the uneven number of cases it produced, the mean times for blocks of five trials were calculated for each subject. These block response times were used as the unit of analysis in a two-way repeated measures ANOVA. The data are shown in figure 3.

The main effect of group was marginally significant, $F(2, 18) = 2{\cdot}89$, MSe = 2·07, $p = 0{\cdot}08$. There was a significant main effect for blocks of trials, $F(5, 90) = 2{\cdot}587$, MSe = 0·067, $p = 0{\cdot}001$. The interaction of blocks by group was also significant, $F(10, 90) = 7{\cdot}37$, MSe = 0·497, $p = 0{\cdot}001$. Follow-up tests for simple main effects showed a significant difference between groups for Blocks 1 and 2 only (Block 1: $F(2, 18) = 8{\cdot}3$, MSe = 0·643, $p = 0{\cdot}003$; Block 2: $F(2, 18) = 4{\cdot}112$, MSe = 0·352, $p = 0{\cdot}034$). Within these simple main effects, a Tukey's HSD test showed a non-significant difference between the no training and virtual reality training groups for both blocks 1 and 2 (Block 1: $p = 0{\cdot}0554$; Block 2: $p = 0{\cdot}628$). Within Block 1 there was a significant difference between the real-world training and the virtual reality training groups ($p = 0{\cdot}025$) and between the real-world and no training groups ($p = 0{\cdot}003$). Within Block 2, only the difference between real-world and no training groups reached significance ($p = 0{\cdot}03$). There were no differences between groups for Blocks 3 through 6.

Another dependent variable of interest is the degree of variability in response times among the different groups. A two-way ANOVA was performed on the variance scores for each trial using training group as a between subjects factor. This analysis showed a significant main effect of group, $F(2, 18) = 18{\cdot}95$, MSe = 0·123, $p = 0{\cdot}001$. *Post hoc* comparisons using Tukey's HSD statistic show that the virtual-reality group showed significantly greater variability in response times than the other two groups ($p = 0{\cdot}001$). The real-world and no training groups were not significantly different ($p = 0{\cdot}60$).

4. Discussion

As the learning curves show (see figure 2), the virtual reality training group obviously learned something during the training period. Despite the similarity in learning curves for those trained in VR and those trained in the real world, this learning did not transfer to the real-world task (see figure 3). These two results taken together seem to indicate that what the subjects learned during VR training was specific only to the context of virtual reality. A closer examination of aspects in the virtual reality environment reveals several skills which are important for performance in this environment but which are irrelevant to performance in the real-world task.

The first such demand is the subtlety required in performing a grasp. Subjects must learn to adjust the timing of their grasp to the time lag inherent in the system. In addition, they must discern the range of acceptable hand postures through a combination of guided instruction and personal experience. Second, subjects needed to learn the optimal position from which to view the task apparatus. Variability in visual perspective may have produced a lack of confidence that the cans were being placed directly on the targets which would have resulted in slower completion times. Third, some subjects, due to body size and reach, needed to reposition themselves in the virtual space to reach all the cans. This locomotion was accomplished by either 'flying' through the virtual environment using hand postures or by physically changing the position of the swivel chair. This former manner of locomotion obviously is specific to the virtual reality context.

In addition to these task demands, there were aspects of the virtual reality environment which were radically different from the real world and which may have contributed to the lack of transfer. There were occasions when subjects removed their virtual hand from view by moving the hand beneath the tabletop. They then had to spend time recovering from this lack of visual feedback by locating their hand in the virtual space. Another factor is the lack of tactile and acoustic feedback. To the subjects, the cans feel weightless and if dropped from a grasp they simply remain in the same position regardless of the relation of that position to either gravity or supporting surfaces. These context-specific features most probably contributed to the subjective difference between the virtual and real environments.

During VR training, subjects may have been continuing to learn further aspects of grasping, locomotion, and Dataglove™ sensitivity that may not have been critical in the familiarization phase. Although subjects picked up and moved objects during VR familiarization, these objects were larger than the cans used during the VR training and the movement of these objects did not require accurate positioning. In subsequent experimentation, we have considered a familiarization procedure which more closely resembles the actual training task. By providing familiarization with skills which are necessary to perform the training task before entering the actual training environment, we would hope to minimize any effects of additional VR learning on the actual training of a specific task.

Despite geometric similarity and task similarity, the VR training environment did not provide adequate training for a motor skill task in the real world. One way to improve transfer of training is to improve the similarity between the training context and task context. We can always try to make the virtual reality environment more like the real world. System limitations make this prospect unlikely for the near future. Current graphics capability of most VR display devices, especially head-mounted displays, limit the field of view, shading, and texture effects which are possible in real time. Processing speed (rendering, communications) and lag time also limit the potential veridicality of the VR interface. A more promising approach to improving transfer may be to make the real world, in essence, more like the virtual reality environment. We can begin by determining the key aspects of the real world task context which are necessary for adequate performance. We need to define which non-application-specific aspects of three-dimensional interactions are necessary for convincing, functional three-dimensional interfaces. Extracting a minimalistic representation for three dimensions is similar to the desktop metaphor for two-dimensional interaction. The desktop of modern PC interfaces looks little like an actual desktop *per se*, but the functionality associated with a desktop remains. By carefully modelling only the necessary parts of the real world, we can begin to bring the virtual and real contexts closer together. This is of particular importance when considering real world designs that might significantly benefit from the convenience of using VR training.

The present results might appear somewhat disappointing for the application of VR technology. However, we believe that many of the present barriers to transfer are due to the technological state-of-the-art, rather than VR *per se*. Consequently, we remain enthusiastic about the potential for VR and human–computer interaction in general. Where two-dimensional interfaces are the constraint on presenting three-dimensional information, for example in air traffic control and flight applications, the promise of VR training and operation is, we believe, especially significant. Indeed VR represents the next generation of the computer interface and therefore will have a vast impact on how anyone operates with such systems. Therefore, as indicated by Wells (1992) there are a plethora of questions for the ergonomist to address in this realm. In addition to the immersive quality it is the capability for immediate relocation and the ability to interact with objects of sizes that do not accord with 'humanscale' (Hancock 1990) that represent the most exciting possibilities.

Current research in our laboratory is directed toward two goals. First, we are attempting to identify the minimal set of real world features necessary for a veridical virtual space. In these experiments we are evaluating spatial cognitive capabilities as a function of real world versus virtual environmental cues. It is possible that such a task which emphasizes cognitive skills, such as memory, may benefit more from virtual training than the presently described experiment. However, it is also possible that the two-dimensional task is limiting the potential for transfer from the immersive interface as provided by VR. Therefore, we are also evaluating tasks involving three-dimensional movement. As a follow-up to the present experiment, a study using a three-dimensional stacking task similar to the present two-dimensional pick and place task is in progress. In this study, half of the subjects first train in the real world and the other half train in VR first in an attempt to separate VR-context-specific learning from task-specific learning. We hope the comparison of results of the two-dimensional and three-dimensional tasks will show VR to be a useful tool for three-dimensional task training.

Acknowledgements

This research was supported by grants from NASA (NAG-1-1118), the Minnesota Department of Transportation, and the University of Minnesota Graduate School to the second author. The Center for Research in Learning, Perception, and Cognition provided clerical and financial support. The first author is currently at 3M Corporation. This work was presented at the 36th annual meeting of the Human Factors Society in Atlanta, October 1992. Correspondence concerning the present article can be addressed to: P.A. Hancock, Human Factors Research Laboratory, 60 Norris Hall, 172 Pillsbury Drive, S.E., University of Minnesota, Minneapolis, MN 55455, USA.

Note

1. There was no significant difference of number of knock-downs or illegal grasps between groups, $F(2, 18) = 2{\cdot}48$, $p = 0{\cdot}11$. The full ANOVA was also performed including trials on which a knock-down or illegal grasp occurred. This inclusion did not alter the significance values for the results reported.

References

BAUM, D. R. 1992, Virtual training devices: illusion or reality? unpublished manuscript, Hughes Training Inc., Minneapolis, MN.

FLEXMAN, R. E. and STARK, E. A. 1987, Training simulators, in: G. Salvendy (ed.) *Handbook of Human Factors*. (John Wiley & Sons, New York), 1012–1038.

HANCOCK, P. A. 1991, The aims of human factors and their application to issues in automation and air traffic control, in J. A. Wise, V. D. Hopkin, and M. Smith (eds), *Automation and Systems Issues in Air Traffic Control* (NATO, Springer, New York), 187–199.

HOLDING, D. H. 1987, Concepts of training, in G. Salvendy (ed.), *Handbook of Human Factors* (John Wiley & Sons, New York), 939–962.

KRUGER, M. W. 1991, *Artificial Reality II* (Addison-Wesley, Reading, MA.).

STURMAN, D. J. 1992, Whole-hand input, unpublished doctoral dissertation, MIT, Department of Media Arts and Sciences.

VPL RESEARCH, INC. 1990, *RB2 Swivel User Interface Documentation* (VPL, Foster City, CA).

VPL RESEARCH, INC. 1991, *Body Electric Manual v3.0* (VPL, Foster City, CA).

WARREN, R. 1992, Fidelity issues in flight simulation, self-motion, and virtual reality displays, *Insight: The Visual Performance Technical Newsletter of the Human Factors Society*, **14,** 1–3.

WELLS, M. J. 1992, Virtual reality: technology, experience, assumptions, *Human Factors Society Bulletin*, **35,** 1–3.

Received 9 September 1992.
Revision accepted 20 October 1992.