Per-Anders Oskarsson

Training in the CV90 simulator

Titel	Träning i Befattningstränare Stridsfordon 90
Title	Training in the CV90 simulator
Rapportnr/Report no	FOI-R--3145--SE
Rapporttyp Report Type	Vetenskaplig rapport Scientific report
Sidor/Pages	26 p
Månad/Month	December
Utgivningsår/Year	2010
ISSN	ISSN 1650-1942
Kund/Customer	Försvarsmakten
Projektnr/Project no	E7151
Godkänd av/Approved by	Magnus Jändel

FOI, Totalförsvarets Forskningsinstitut
Avdelningen för Informationssystem
Box 1165
581 11 Linköping

FOI, Swedish Defence Research Agency
Information Systems
Box 1165
SE-581 11 Linköping

Sammanfattning

Två studier vid Befattningstränare Stridsfordon 90 (BFT 90) har genomförts. Den första studien genomfördes hösten 2009, under den tidigare delen av soldaternas utbildning. Den andra studien genomförde våren 2010 under den senare delen av soldaternas utbildning. 2009 deltog endast skyttar, medan även förare och vagnchefer inkluderades 2010. Studierna fokuserade på fidelity (överstämmelse med verkligheten), presence (närvarokänsla), motivation, om det är roligt, feedback, inlärning och transfer av träning.

Resultaten visar att skyttarna hade högre skattningar än övriga befattningar på inlärning i simulatorn och simulatorträningens effekt på prestation i verklig CV90. Skyttarna hade också tränat mer i simulatorn än övriga befattningar.

Under tidig utbildning 2009 tränade skyttarna mer på grundläggande handhavande och förberedelser inför träning i verklig CV90. Under senare utbildning 2010 tränade de däremot mer på teamträning. *Inlärning i simulatorn* skattades högre av skyttarna under tidig utbildning än senare, för övriga studerade begrepp av simulatorträning var det endast små skillnader. Det tycks rimligt att inlärning i simulatorn skattas högre under tidiga stadier med mycket procedurträning i simulator, jämfört med senare stadier med större fokus på teamträning i riktig CV90.

Trots att deltagarna hade många negativa synpunkter på simulatorns grafik och överensstämmelse med verkligheten så var deras skattningar av hur mycket de lärde sig under träningen och träningens effekt på prestation i verklig CV90 relativt höga. Detta tyder på möjligheten till träningseffekt i enkla simulatorer med låg fidelity.

Med komponentanalys och linjär strukturell ekvationsmodellering (LISREL) kunde deltagarnas skattningar användas för att skapa en modell av orsakssambanden mellan de studerade begreppen för simulatorträning. Modellen visar att träningsmiljön (mätt genom hur roligt träningen upplevdes, motivation, feedback, närvarokänsla och fidelity) hade en positiv effekt på träning i simulatorn (mätt genom inlärning i simulatorn och träningens effekt i riktig CV90). Träning i simulatorn kunde även uppdelas i två komponenter, vilket ger en modell med tre komponenter där träningsmiljön har en positiv effekt på träningen i simulatorn, vilket i sin tur har positiv inverkan på transfer av träning till verklig CV90.

Resultaten kan även ligga till grund för förbättring av det enkätinstrument som används vid undersökningarna.

För att kunna ge rekommendationer om förbättringar avseende träning och simulatorutveckling krävs studier med ett mer kvalitativt angreppssätt.

Nyckelord:

Simulatorträning, presence , närvarokänsla, fidelity, motivation, feedback, inlärning, träning, transfer av träning,

Summary

Two studies at the CV90 simulator (Sw. Befattningstränare Stridsfordon 90, BFT 90) have been performed. The first study was performed in the autumn 2009, during the early part of the soldiers' education. The second study was performed during the later part of the soldiers' education. 2009 only gunners participated, whereas drivers and commanders were included 2010. The focus of the studies were fidelity, presence, fun, motivation, feedback, learning, and transfer of training.

The results show that the gunners rated learning in the simulator and the trainings effect on performance in the real CV90 higher than the other positions. The gunners had also trained more in the simulator.

During early education 2009, the gunners trained more on basic handling and preparation before training in the real CV90, compared to later when they trained more on team training. Learning in the simulator was rated higher by the gunners during early education compared to later, for the other studied training concepts there were only small differences. It seems reasonable that learning in the simulator is rated higher in early stages with much procedural training in the simulator, compared to later stages with a larger focus on team training in the real CV90.

In spite that the participants had many negative views on the graphics and fidelity in the simulator, their ratings of how much they learnt during training and the effect of the training on performance in the real CV90 were relatively high. This indicates the possibilities of training effects in simple low fidelity simulators.

By component analysis and linear structural equation modelling (LISREL) a model of the casual relations between the studied training concepts could be created. The model shows that the training environment (measured by how fun the training was experienced, motivation, feedback, presence, and fidelity) had a positive effect on the training in the simulator (measured by learning in the simulator and its effect on performance in the real CV90). Training in the simulator could also be split into a two components, which led to a model with three components where the training environment has a positive effect on the simulator training which in turn has a positive influence on transfer of training to the real CV90.

The results can also be used for instrument development, concerning the questionnaire used at the studies.

To provide recommendations on improvements concerning training and simulator development, studies with a more qualitative approach are needed.

Keywords:

simulator training, presence, fidelity, motivation, feedback, learning, training, transfer of training

Table of contents

1 Introduction

Simulators can provide efficient training with reduced costs, reduced risks and focus can be put directly on the training task. Conversely, flying a real aircraft may require co-ordination with numerous other services (e.g. air traffic control and maintenance), and need of appropriate weather and visibility conditions (Lee, 2005). In principal, these types of circumstances also apply to training in other vehicles or systems, in this study combat vehicles. Furthermore, dangerous tasks can be safely trained in a controlled environment. The possibilities to have a controlled environment by holding extraneous variables constant, also enhances the instructor's capacity to getting an adequate picture of the trainees' skills, or learning curves, on specific procedures or manoeuvres. Some key concepts for simulator training are fidelity, presence, motivation, fun, feedback, learning, and transfer of training.

Fidelity is the degree of similarity between the simulator and the simulated system and its environment. The types of fidelity most often considered are the degree to which the simulator looks like (physical fidelity), acts like (functional fidelity), and creates similar psychological reactions (psychological fidelity) as the system being simulated (Stanton, 1996). There is no unitary view of the degree of fidelity that a simulator must have for achieving effective transfer of training. It seems however clear that perceptual motor tasks require a higher level of physical fidelity than that of a cognitive task (Stanton, 1996). The trainee's degree of expertise must also be considered. For example, a novice flight student may learn the most from low-fidelity instructions, whereas when faced with the highest possible level of fidelity (i.e. flying the real airplane) the student may be so confused and stressed that no learning occurs at all (Alessi, 1988). Furthermore, pedagogical requirements can be substantially different from those on physical and functional fidelity (Farmer, van Rooij, Riemersma, Jorna, & Moraal, 1999).

Presence is defined as the subjective experience of being in one place or environment, but physically situated in another (i.e., the simulator). Important factors for presence are degree of immersion in the virtual environment, involvement, and control over the task environment (Witmer & Singer, 1998).

Motivation can be described as the willingness or desire to engage in a task. Motivated learners are thus more interested and involved, devote more time on a task, and are more committed to continued task activity. Motivated learners are enthusiastic, focused and engaged, and interested in what they are doing; and their behaviour is self-determined, and driven by their own volition rather than external forces (Garris, Ahlers, & Driskel, 2002).

To experience the training as fun has reasonably large importance for motivation and involvement. Here much can be learnt from computer gaming, where motivated players, voluntarily, with large commitment often devote huge amounts of time to gaming.

Feedback is essential to supporting performance, motivation and lerning. It provides the trainees with guidance on how they should adapt their behaviour to reach desired goals. Debriefing is a critical part of feedback and important for transfer of learning from the simulator to real world performance. During debriefing, the role of the instructor is a critical component (e.g. Garris et al., 2002). Feedback can also be provided by training tools implemented in the simulator (Wiese, Freeman, Salter, Stelzer, & Jackson, 2008; Nählinder, Borgvall, & Svensson, 2004). However, in this context it is a problem that a simulator may provide a very realistic training environment, but is often not built with the explicit purpose of being a pedagogical teaching device. Thus, simulators often replicate a real world situation without providing the proper pedagogical support and feedback that might facilitate learning (Nählinder, 2009).

Transfer of training refers to how previous learning influences behaviour in a later situation. In this context, how simulator training influences performance in the real system. Transfer of training is usually measured by changes in task performance time, and can thus be positive, nil, or negative (e.g., Roscoe & Williges, 1980; Liu, Blickensderfer, Macchiarella, & Vincenzi, 2008; Borgvall, Castor, Nählinder, Oskarsson, & Svensson, 2007). However, in this study the participants self ratings were used to assess transfer of training.

The casual relations between the concepts of training, discussed above, are not fully investigated. These types of relations can, however, be studied by linear structural equation modelling (LISREL). This is a powerful tool for analysis of causal relationships and the effects of different causal factors. The analyses are based on correlational statistics using the linear relationships between variables, and the common variance between the variables (Jöreskog & Sörbom, 1998). If this statistical modelling approach is used on empirical data, the resulting models represent the casual relationships between the empirical data; and the fit of the model is statistically tested.

In a previous study LISREL was used to study the casual relations between these concepts of training. That was a meta study of amalgamated data from four different simulator studies CV90 simulator, Tank 122 simulator, Sonar simulator, and MCM-wargaming. The model showed that feeling of involvement in the simulation influences the training in the simulator, which in turn has a positive influence on transfer of training to real world performance (Oskarsson, Nählinder, & Svensson, 2010).

1.1 Purpose

The training concepts, presented above, were investigated last year in the CV90 simulator, when the gunners were in the early phase of their CV90 education (Nählinder, Oskarsson, Lindahl, Hedström, & Berggren, 2009; Oskarsson et al., 2010). To learn more about the simulator training, it was decided to investigate if the gunners' view of the training concepts changed over time during their education. And for this reason, also investigate if the view of the different positions in the CV90, gunners, drivers and commanders, differed.

Thus, the main purpose of the present study was to investigate if the gunners' view of the training concepts in the early phase of training differed from the later phase. The second purpose was to investigate if the view of the three positions in the CV90 differed, gunners, drivers and commanders.

Concerning the casual analysis, the purpose was to narrowing the focus to the CV90 simulator. Thus, the main purpose was to investigate the causal relations between the concepts of training in the CV90 simulator and also to relate this to the previous model on amalgamated data from four simulators (Oskarsson et al., 2010).

2 Methods

The study consisted of two phases. Phase one was performed in autumn 2009 when the soldiers had recently begun their CV90 education and phase two was performed in spring 2010 when the soldiers were in the final stage of their CV90 education. The results from phase 1 were extensively presented in Nählinder et al. (2009) and in Oskarsson et al. (2010).

2.1 Participants

2.1.1 Phase one – autumn 2009

19 soldiers, all gunners, participated in the study. Their mean age was 18.9 ±0.2 (Mean ±SD) years. They had 16 hours simulator training, and practically no experience of the real CV90. 3 instructors also answered a similar questionnaire (only answers on open questions from the instructors are reported here).

2.1.2 Phase two – spring 2010

40 soldiers participated in the study, 14 Gunners, 13, Commanders, and 13 Drivers. Their mean age was 19.2 ±0.4). The gunners had also participated in phase one of the study. Originally, 43 soldiers participated, but due to incomplete answers the results from one soldier were cancelled, and the results from two because change of position in the CV90. One instructor also answered a similar questionnaire (only answers on open questions from the instructor are presented here)

The time of simulator training was approximately, gunners 40 hours, commanders 24 hours, and drivers 12 hours. The simulator training was performed in a number of training bouts of approximately four hours each. Each participant, of the tree positions, had several hundred hours experience in the real CV90.

The participants experience of computers was rated high 5.7 ±1.2 (Mean ±SD) (1 = no experience, 7 = large experience). They played computer games moderately often, 4.6 ±1.2 (1 = never play computer games, 7 = very often play computer games). Their view of simulator training as a complement to other training was rated moderately positive, 4.7 ±1.6 (1 = very negative, 7 = very positive).

2.2 Equipment

2.2.1 Combat Vehicle 90

The CV 90 (Combat Vehicle 9040) is a Swedish infantry caterpillar combat vehicle with rotating tower. It has three crew members, driver, gunner and commander. The commander and gunner are positioned in the turret and the driver in the hull.

2.2.2 CV90 Simulator

The CV90 simulator (Sw. Befattningstränare Stridsfordon 90, BFT 90) at the Swedish Army's Land Warfare Centre (LWC) (Sw. Markstridsskolan) can simultaneously train three CV90 crews. The three crew members: driver, gunner, and commander sit side by side in the simulator. They drive, shoot and interact with the simulator by PC based game controllers. In front of each crew member are three computer screens, which display instrumentation and simulated environment (Figure 1). The three CV90 crews are separated by walls in the simulator. The simulator is placed within a container, which can easily be moved by truck between training sites (Figure 2). In Figure 3, three crew members, driver gunner and commander can be seen sitting side by side in the CV90 simulator.

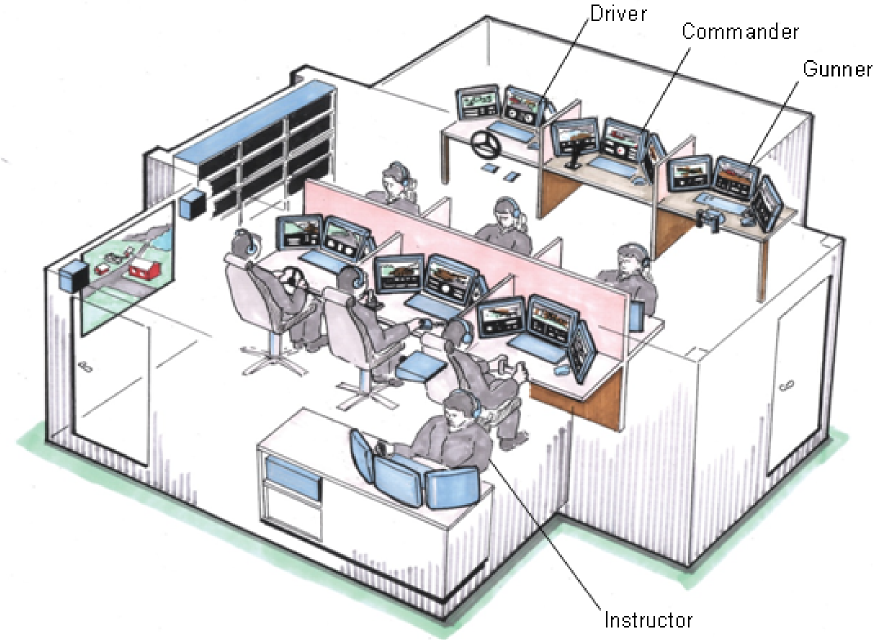

Figure 1. View of the CV90 simulator within the container. The three crew members in two of the three simulated CV90s sitting side by side.

Figure 2. The CV90 simulator container during unloading and unloaded on ground.

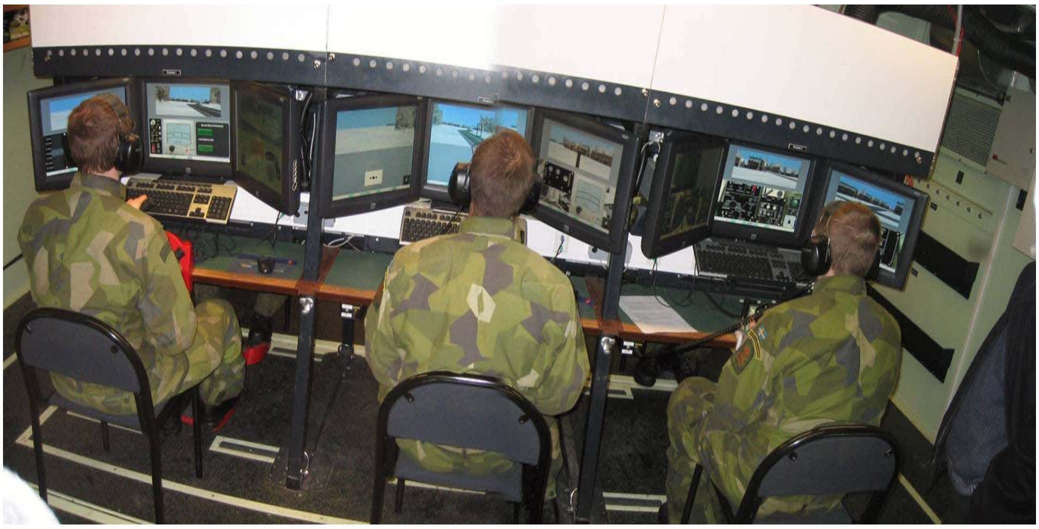

Figure 3. A CV90 crew seated in the CV90 simulator. From left to right: Driver, Commander, and Gunner.

2.2.3 Questionnaire

A questionnaire with 76 questions was used, with 7-point rating scales and open questions. The questionnaire included a number of questions about the seven training concepts (see Section 3.5 Reliability analysis of training concepts). It also included background questions and questions about amount of simulator use on different types of training. The instructors answered a similar questionnaire.

2.3 Procedure and design

Phase one was performed in autumn 2009, when the soldiers had recently begun their CV90 education. Phase two was performed in spring 2010, when the soldiers were in their final stage of their CV90 education.

In both phases, the questionnaire was sent by post, to the military base where the training was performed. On site the questionnaire was administered to the soldiers by an officer. Researchers at FOI provided the officer with instructions via telephone and e-mail, and each questionnaire was provided with a covering letter with instructions and a stamped envelope.

3 Results

In case of violation of the sphericity assumption in the analyses of variance (ANOVAs) the Greenhouse-Geisser corrected *p*-value is given (indicated: ^{G}).

Post Hoc testing was performed with Tukey HSD for the between groups effect, and with Sidak confidence interval correction for the within groups and interaction effects.

3.1 Comparison positions 2010

There were no significant differences between the three positions in the ratings of experience of computers, how much they played computer games, or their view of simulator training as a complement to real training.

The ratings of how much the simulator were used for different types of training were analysed with ANOVA with 3 (position in CV90, driver, gunner, commander) × 7 (type of training) mixed design, with position in CV90 as between groups variable and type of training as within groups variable.

The ratings of the training concepts were analysed with ANOVA with 3 (position in CV90, driver, gunner, commander) × 7 (concept of training) mixed design, with position in CV90 as between groups variable and concept of training as within groups variable.

3.1.1 Type of training

The ANOVA showed a significant main effects of position in CV90, $F(2, 37) = 3.87$, $p < 0.05$, and type of training, $F(6, 222) = 14.83$, $p < .001^{G}$, and a strong tendency of interaction effect of position in CV90 by type of training, $F(12, 222) = 1.86$, $p = 0.057^{G}$.

Post hoc testing of the main effect of position in CV90 showed that the drivers had used the simulator significantly less than the gunners for the defined types of training ($p < .05$) (Figure 4).

Post hoc testing of the main effect of type of training showed that the simulator was significantly most used training of *team cooperation* (*p*s < .05), and significantly least used for *preparation before training in reality* (*p*s < .01) (Figure 4). In detail, the simulator was significantly more used for *team cooperation* compared to the other types of training (*p*s < .05), except for *basic handling*. And that the simulator was significantly less used for preparation *before training in reality* compared to the other types of training (*p*s < .01), except for *emergencies* (however strong tendency, $p = .054$) and *gaming just for fun*.

Since the tendency of interaction of position by type of training was strong, and significant before Greenhouse Geisser correction, post hoc testing of the differences between the three positions was performed for each specific type of training. The test showed that the *drivers* used the simulator significantly less on *basic handling* compared to both gunners and commanders (*p*s < .05), and significantly less for *difficult actions* compared to the gunners ($p < .01$) (Figure 4).

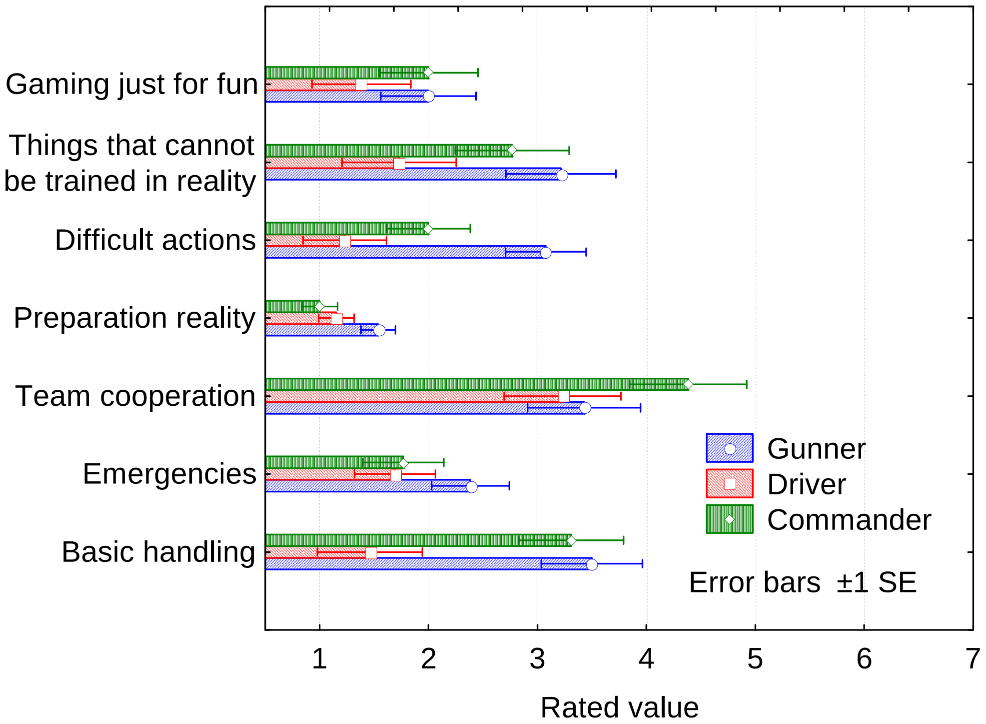

Figure 4. Ratings of how much the three positions used the simulator for different types of training (1 = very seldom, 7 = very often).

3.1.2 Effects of training

The ANOVA showed a significant main effects of concept of training, $F(6, 222) = 11.90$, $p < 0.001$ [G]; a significant main effect of position in the CV90, $F(2, 37) = 3.58$, $p < .05$; and a significant interaction effect of concept of training by position in the CV90.

Post hoc testing showed that the main effect of concept of training was due to higher ratings of *feedback* compared to all other concepts (*ps* < .05); and lower ratings of *presence* compared to *fun*, *motivation*, *feedback*, and *fidelity* (*ps* < .05); and lower ratings of *learning in simulator* compared to *fun* ($p < .05$), fidelity ($p < .001$) and *effect in real CV90* (tendency, $p = .079$) (Figure 5).

Post hoc testing showed that the main effect of position in the CV90 was due to higher ratings of the gunners compared to the drivers ($p < .05$) (Figure 5).

Post hoc testing showed that the interaction effect of concept of training by position in the CV90 was due to higher ratings of *effect in real CV90* for the gunners compared to the drivers ($p < .005$); and higher ratings of *learning* in the simulator for the gunners compared to both the drivers and commanders ($p < .005$) (Figure 5).

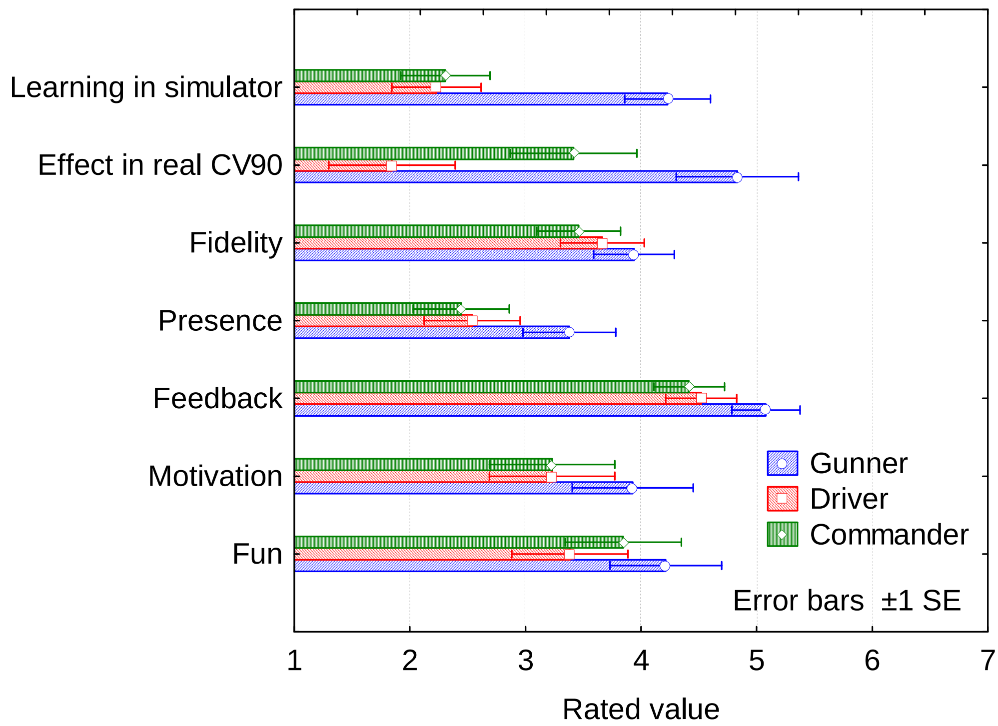

Figure 5. The ratings of the three positions of the studied concepts of simulator training. One column for each position in the CV90 (1 = very little, 7 = very much).

3.2 Comparison gunners 2009-2010

The ratings of how much the simulator were used for different types of training were analysed with ANOVA with 2 (phase of training, early or late) × 7 (type of training) mixed design, with phase of training as between groups variable and type of training as within groups variable.

The ratings of the training concepts were analysed with ANOVA with 2 (phase of training, early or late) × 7 (type of training) mixed design, with phase of training as between groups variable and concept of training as within groups variable.

The reason for treating phase of training as a between groups variable was that personal information was not collected at the two occasions; thus the individual results from the data collections (phases) could not be matched against each other. Furthermore, a number of the participants in the early phase 2009 did choose not participate 2010, and 2 were excluded 2010 because they had changed position in the vehicle. Thus, the answers from 19 gunners 2009 and 14 gunners 2010 were analysed.

3.2.1 Type of training

The ANOVA showed a significant main effect of type of training, $F(6, 186) = 12.87$, $p < .001$ [G], and a significant interaction effect of type of training by phase of training, $F(6, 186) = 8.94$, $p = 0.001$ [G]. There was no significant main effect of phase of training.

Post hoc testing of the main effect of type of training showed that the simulator was significantly most used for training of *basic handling* and *difficult actions*, and significantly least used for *gaming just for fun* (Figure 6).

In detail, *Basic handling* > *emergencies, team training, preparation before training in reality*, and *gaming just for fun* (*ps* < .01), and *things that cannot be trained in reality* (tendency *p* = .061) *Difficult actions* > *emergencies, team training, gaming just for fun* (*ps*

< .01), and *preparation before training in reality* ($p < .05$). *Gaming just for fun* < *basic handling*, *difficult actions* (*ps* < .01), and *things that cannot be trained in reality* ($p < .05$).

Post hoc testing of the interaction effect of type of training by phase of training showed that early in the education (2009) the simulator was significantly more often used for training of *basic handling*, *preparation before training in reality* (*ps* < .01)*, and difficult actions* ($p > .05$). On the contrary, in the later phase of the education, the simulator was significantly more used for *team cooperation* ($p < .01$) (Figure 6).

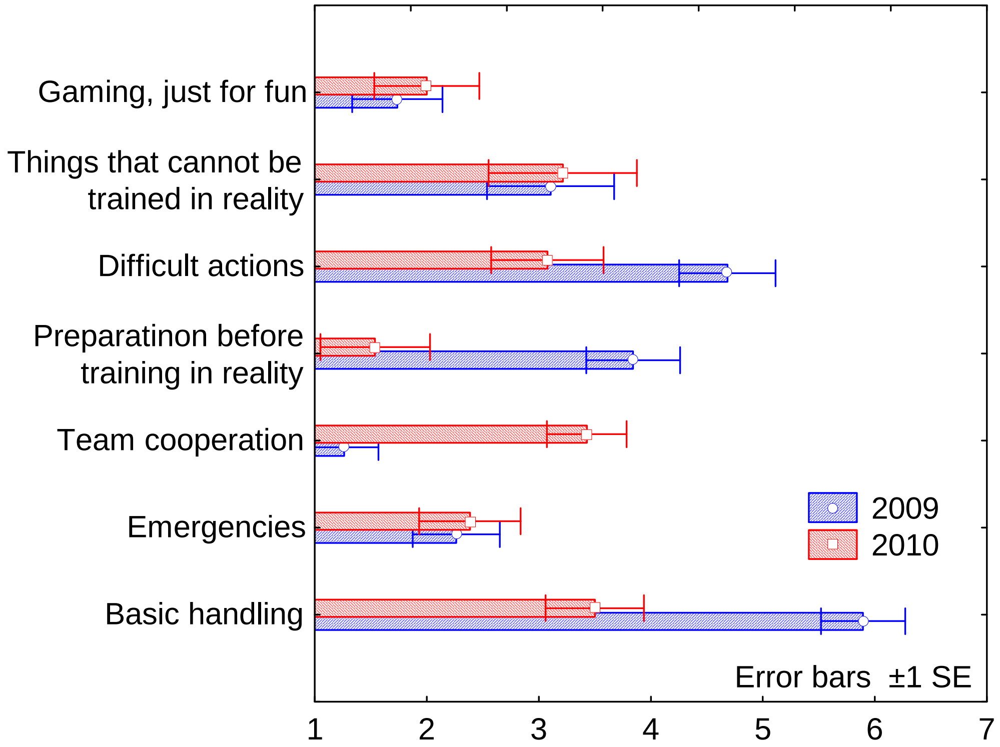

Figure 6. The ratings of the gunners of how much the simulator was used for the different types of training at 2009 and 2010 (1 = very seldom, 7 = very often).

3.2.2 Effects of training

The ANOVA showed a significant main effect of concept of training $F(6, 186) = 6.83$, $p < .001$ [G]. There was no significant main effect of phase of training, or interaction effect of concept of training by phase of training.

Post Hoc testing of the main effect of concept of training showed that *learning in the simulator*, *effect in real CV90*, and *feedback* were rated significantly higher than *presence* (*ps* < .05), and *feedback* was rated significantly higher than *motivation* ($p < .05$) and a tendency of higher ratings of motivation than of *presence* ($p = .67$) (Figure 7).

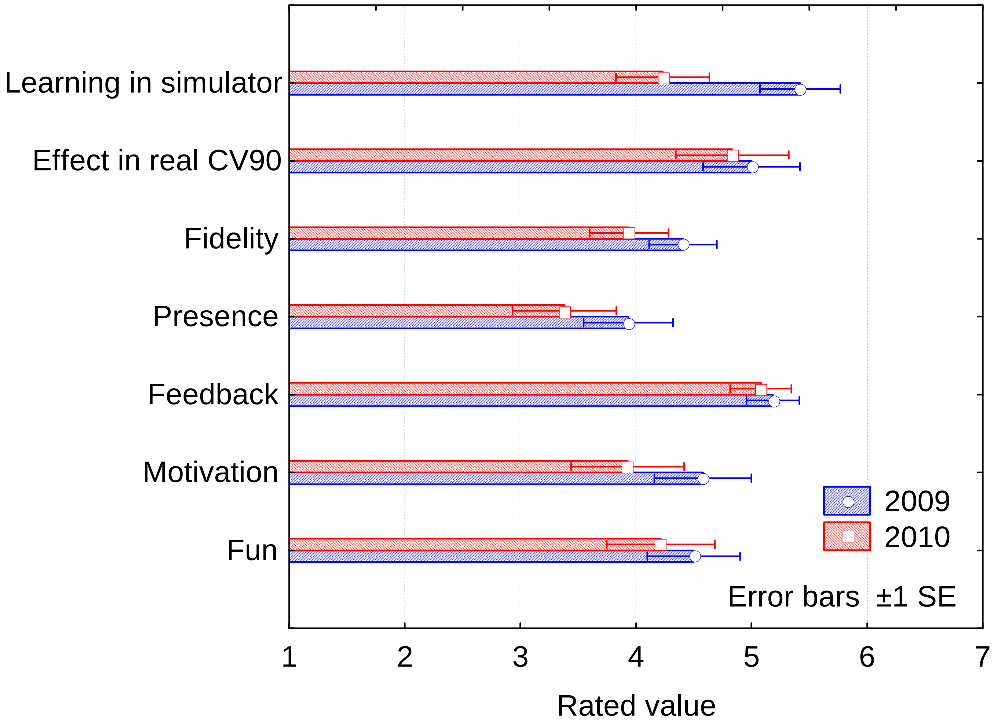

Figure 7. The ratings of the gunner of the studied training concepts at 2009 and 2010. (1 = not at all, 7 = very much).

As can be seen in Figure 7, the difference of *learning in the simulator* was rather large between the measurements at 2009 and 2010. The reason that the analysis of variance did not have enough power to indicate this difference by a significant interaction effect is most likely that this was the only deviance among similar ratings on the six other concepts of training (observed power of the interaction effect .35). Therefore a separate univariate analysis of variance of the ratings of *learning in the simulator* was performed. This ANOVA showed a significant effect of phase of training, $F(1, 31) = 4.99$, $p < .05$). That means that the ratings of *learning in the simulator* were significantly higher 2009 than 2010.

3.3 Open questions

Since the gunners answered the questionnaire at two occasions, both 2009 and 2010, they gave fewer comments on the open questions at the second occasion 2010. Thus, some of their comments from 2009 are also included. Instructor comments from 2009 are also included.

One instructor commented that the simulator is much used for mass training of gunners, which is especially valuable since live ammunition is expensive.

One commander commented that the simulator is usable for team training, but otherwise unrealistic for training of commanders. Two drivers commented that the simulator is too unrealistic and that they do not learn very much in it.

One driver commented liaison training is both fun and instructive and rather frequently trained. One gunner commented that training of "real war" with many enemies and much action is fun. One gunner meant that gaming just for fun just some time would have been fun.

One driver commented that the feedback in the simulator is very important, so that you learn from your mistakes. Another driver meant that it is very good to get feedback in the simulator, because when they train in the real CV90, the instructors have limited time to give feedback.

Five participants commented that team cooperation in the real CV90 was improved by the training in the simulator. Seven gunners (five in the early phase) meant that basic handling and aiming of weapon was improved by the simulator training. In the early phase, two gunners also commented the value of mass training.

Three participants commented that the feeling of presence is poor, two of the meant that that computer games generally provides a better feeling of presence. Some especially commented the presence for the driver as poor. Several participants commented that the graphics needs to be improved, some of them also wished improvement of the sound.

Four participants mentioned poor correspondence with reality for the driving task. Two commanders commented that the simulator is not really made for training of their tasks. One of them meant that it is mainly built for training of gunners.

The following comments were given on things that may have negative influence on the performance in the real CV90. Two drivers commented that the simulator is too easy to drive, and thus do not provide the real feeling of driving. Three gunners commented that the lack of feeling of reality in the simulator may have negative influence on performance in the real CV90. One commander mentioned handling of machine gun concerning stoppage. One gunner comment that the gun does not spread as in reality and that the simulated enemies do not make mistakes.

3.4 Casual analysis of training concepts

In order to investigate the causal relations of the concepts of simulator training at the CV90 simulator the data collected in 2009 and 2010 was analysed with principal components analysis and linear structural modelling (LISREL).

In the correlation matrix of the seven concepts of training, analysed above the Kaiser-Meyer-Olkin Measure of sampling adequacy (KMO) = .86, and Bartlett's test of sphericity, χ^2= 252.37, df = 21, $p < .001$, showed that the correlation matrix is not an identity matrix. This indicates that factor analytical methods can be used. However, in the correlation matrix of the seven concepts, analysed above, the determinant was .01, which indicates a not positive definite correlation matrix, i.e. linearly dependences in the matrix. This was due to high correlations between fun and motivation, r = .83, and between motivation and presence, r = .84.

Since removing or merging variables was considered a drawback, principal components analysis with Oblimin rotation with Kaiser Normalisation was performed on the original matrix[1].

Extraction with eigenvalues above one resulted in a one component solution which explained 62 percent of the variance. Chronbach's alpha for this general component was .88.

Solutions for two and three components were also analyzed. The two component solution was considered making most conceptual sense, where *fun*, *motivation*, *feedback*, *presence*, and fidelity formed a component related to the *training environment. Effect in real CV90* and *learning in simulator* formed a component related to the *training effect* (Table 1). In the rotated solution 59 percent of the variance was explained by the *training environment* component (unrotated 62 percent) and 40 percent of the variance was explained by the

[1] Factor analysis, e.g. Principal Axis Factoring, on highly correlated items is generally problematic.

training effect component (unrotated 11 percent)[2]. Chronbach's alpha of the *training environment* component was .90, and of the *training effect* component .70.

Table 1. The two extracted components *Training Environment* and *Training effect* (component loadings from rotated solution).

	Components	
Concepts	Training Environment	Training Effect
Fun	.94	-.14
Motivation	.95	-.00
Feedback	.78	.00
Presence	.83	.15
Fidelity	.56	.22
Learning in simulator	.20	.71
Effect in real CV90	-.07	.95

The casual relationship of the two component solution was analysed with linear structural modelling (LISREL). The ovals in Figure 8 are latent factors (same as extracted components in the component analysis). The rectangles are the manifest variables (same as variables in the component analysis). The model shows that the *training environment*, measured by *fun*, *motivation*, *feedback*, *presence*, and *fidelity* has a positive influence on the *training effect* measured by *learning in simulator* and *effect in the real CV90* (Figure 8).

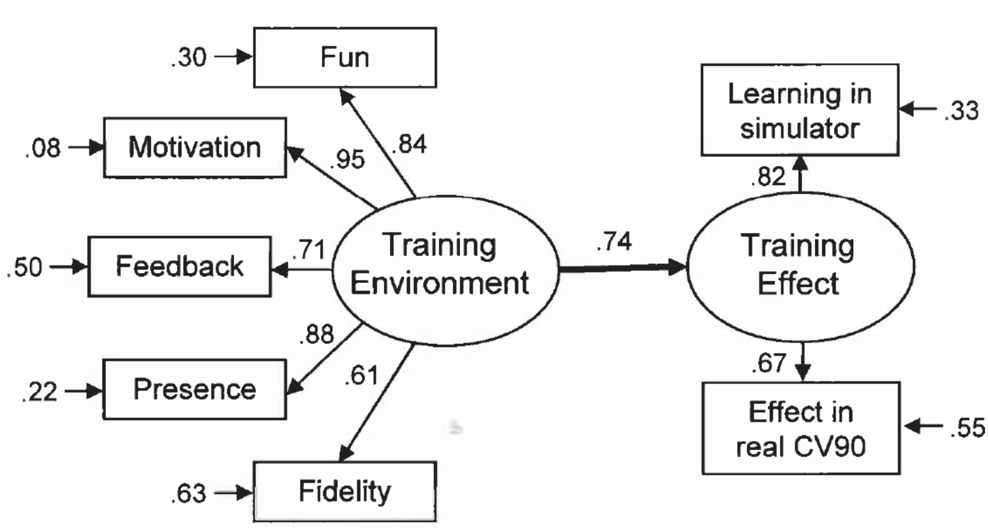

Figure 8. LISREL model showing that the training environment has a positive influence on the training effect. The rectangles represent the manifest variables, the seven measured concepts of training, and the ovals the two latent variables. Numbers indicate factor effects and error terms.

The LISREL model in Figure 8 does not significantly differ from the empirical data, $Chi^2 = 20.5$; df = 13; $p = .084$. Even though there is a tendency of a difference, the model

[2] With oblique rotation, the sum of explained variance is not relevant since it, because of correlated factors, may exceed 100 percent (even though in this case the correlated components only sum to 99 percent).

can be considered a relatively good representation of the empirical data. The Comparative Fit-Index (CFI) is .98 and the Goodness of Fit Index (GFI) is .91.

A model where the manifest variables *learning in the simulator* and *effect in real CV90* were assigned to one latent variable each, was also tested. In this model the latent variable *training environment* was the same, but the latent variable *training effect* was split into the two latent variables *training* and *transfer of training*. That is the *training environment* has a positive effect on *training* measured by *learning in the simulator*, which in turn has a positive effect on *transfer of training* measured by *effect in the real CV90* (Figure 9).

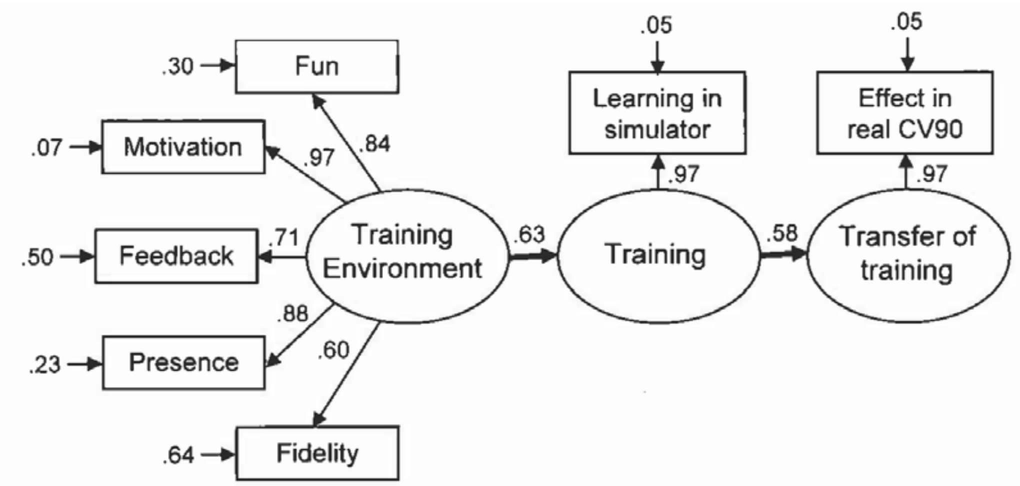

Figure 9. LISREL model showing that the training environment has a positive influence on the training in the simulator which in turn positively affects transfer of training to the real CV90. The rectangles represent the manifest variables, the seven measured concepts of training, and the ovals the three latent variables. Numbers indicate factor effects and error terms.

Nor the model in Figure 9 differed significantly from the empirical data, Chi ^{2} = 22.4; df = 14; p = .070. Even though, also in this case there was a tendency of difference. The Comparative Fit-Index (CFI) is .98 and the Goodness of Fit Index (GFI) is .90. However, a drawback with this model is that the two latent variables *training* and *transfer of training* only have one manifest measurement variable each, which means that the error terms for these variables had to be given a fixed value (.05 was chosen).

Both LISREL models were also tested with the manifest variable *feedback* moved to the latent variable *training effect* in model 1 (Figure 8) and to the latent variable *training* in model 2 (Figure 9). However, both these models differed significantly from empirical data.

3.5 Reliability analysis of training concepts

The previously reported training concepts of the gunners 2009 (Nählinder et al., 2009; Oskarsson et al., 2010) were calculated by the arithmetic mean value of a number of questions. Before analysis of the training concepts in this study the reliability of the concepts were investigated, and in some cases underlying questions were excluded. Thus, for those concepts where questions were excluded, minor differences compared the previously presented results occur.

Since the number of questions in the questionnaire was large in proportion to participants in the study factor analysis was not performed to isolate the questions related to each questions. Instead reliability analysis and correlation analysis was used.

The questions about learning and fun in the simulator were asked for each type of training (see Figure 4. Ratings of how much the three positions used the simulator for different types of training (1 = very seldom, 7 = very often). The reliability of these questions was high. For learning Chronbach’s alpha was .74, and for fun .89. However, since the amount of training on the different types of training varied much between the different positions in the CV90 2010, and between the gunners 2009 and 2010, it was still considered more reliable, for comparisons between the groups, to use the single general question for each of these concepts. That is, how much they generally learned in the simulator, and how fun they generally found the training in the simulator.

Four questions were asked about feedback. It was how much feedback the participants received from the simulator and how much they received from the simulator and how much they learned from these two types of feedback respectively. The correlation between how much feedback they received from the simulator and how much they learned from the feedback from the instructors was relatively low (r = .22). Except from this the correlations were high and Chronbach’s alpha was relatively high, .76. Thus, all four questions were used for calculation of feedback.

Training effect in the real CV90 was investigated by two questions, if anything became easier in the real CV90 because of the simulator training, and if anything became more difficult. The correlation between these two questions was positive. That is, participants who considered that things became easier in the real CV90 because of the simulator training also tended to think that other things became more difficult. This was however, even if correlated, on a much lower level. The ratings on if anything became more difficult in the CV90 (1 = not at all, 7 = very much) were very low, 2.1 ±1.4 (Mean ±SD), and there were no differences between the groups of participants. Because of the positive correlation, it was not considered suitable to reverse the scale of the question about if anything became more difficult, and use the arithmetic mean of the two variables. Thus, only the question about if anything became easier was used for the concept of training effect in the real CV90.

Presence was assessed by five questions. Reality of feeling of movement, convincing feeling of presence, real feeling during driving, and involvement during the simulation. Chronbach’s alpha (.93) for these questions was very high. Thus, all five questions were used for calculation of presence.

Fidelity was assessed by four questions. Correspondence with reality of the physical parts, as instrumentation; quality of presentation, as graphics and sound; realism of simulation, e.g. CGFs; ease of acting (performing tasks) compared to in real CV90. Chronbach’s alpha was low, .48, and the last question about acting had low correlations with the other questions. Therefore, this question was excluded from the calculation of fidelity, which resulted in a relatively high Cronbach’s alpha, .74.

4 Discussion

The gunners used the simulator most, and the drivers least. This was especially true concerning basic handling and difficult actions.

Feedback was rated highest of the training concepts and presence lowest. The gunners generally rated highest, and especially on learning and the effect in the real CV90. If only the drivers and commanders are considered learning was rated very low and if only the drivers are considered effect in real CV90 was rated very low.

The fact that the simulator was most used most by the gunners and least by drivers and that learning was rated lower by the drivers was also reflected by the open questions. It was, for example, commented that the simulator is too unrealistic and too easy to drive to provide good learning for the drivers. That the gunners used the simulator for training of basic handling is reflected by the comment that the simulator is very usable for mass training, which is especially valuable since live ammunition is expensive, and by the gunners who mentioned that the training lead to improved basic handling and aiming of weapons. That the simulator was much used for team cooperation is reflected by the comments that team cooperation in the real CV90 is improved by the training, and that liaison training is fun, instructive and frequently trained.

Concerning the comparison between the gunners early in education 2009 and late 2010; there was a shift from a higher amount of training of basic handling and preparation before training in reality in the early phase towards more training of team cooperation in the later phase. This seems reasonable, since in early phases there is a larger amount of training on basic handling of weapons etc. In the early phase they had also very little experience of the real CV90, thus all this training, could in some sense be regarded as a preparation for their future tasks in the real CV90. Since the interaction between the crew members becomes more important at later phases, when the crew members must work together in the real CV90, a larger amount of crew training becomes natural.

It seems reasonable that learning in the simulator is rated higher in the early phase when the major part of the training is performed in the simulator, than in the later phase when a large part of the training had been transferred to the real CV90.

Learning in the simulator and effect in the real CV90 is rated relatively high, especially in the early phase. This is interesting, since the open questions indicate a rather negative view of the fidelity of the simulator. This indicates the possibility of getting relatively good training results in a simple low fidelity simulator. It must however be considered that the importance of fidelity, and type of fidelity, is different for different tasks.

Even if it is likely that the training results would be improved by increased fidelity, it is necessary to consider the costs in relation to the results. Since improving fidelity as a rule means higher costs. That the fidelity of the driver’s task needs to be improved seems obvious from the answers on the open questions. It is however notable that the drivers did not rate fidelity lower than the gunners and commanders. To sort out the need of general improvement of fidelity, e.g. graphics and sound, requires further investigation.

Even if the open questions suggest that there are things that may have a negative influence on the performance in the real CV90, the ratings of if anything became more difficult in the real CV90 because of the training in the real CV90 were low. Thus, these comments should rather be regarded as suggestions of improvement than indicators of real problems.

The low ratings of how much the simulator was used for gaming just for fun are in principal self-explanatory, since the participants were not allowed to use the simulator for that purpose. Thus, this may have been an irrelevant question at this simulator. The reason for this question is that the questionnaire was originally developed for use at the Flightbook simulator at the Swedish Air Combat Training School, where a free pedagogic is used, where the participants are free to choose their training tasks. It was however considered of value for comparisons between simulators, to use the same questions.

Also, more efforts ought to be devoted to exploring the possibilities of utilizing the experiences from computer gaming; where skillful game developers create games that are considered so fun, amusing, and exciting that gamers with large involvement voluntarily devote huge amounts of time to playing.

The first LISREL model indicates that the training environment has a positive effect on the training effect in the simulator. The alternative LISREL model indicates that the training environment has a positive effect on the training in the simulator which in turn has a positive effect on transfer of training to the real CV90. This model, in principal, replicates the previous model on amalgamated data from four simulators (Oskarsson et al., 2020). The only differences are that in the previous model fidelity belonged to the latent variable simulator training, and that the first latent variable was considered to reflect involvement in the training. Partly, because fidelity was now assigned to the first latent variable, it was instead considered reflecting the training environment. This model probably also makes most conceptual sense. However, a problem is that in this model the latent variables training in the simulator and transfer of training only have one manifest measuring variable each. If further studies are performed, the questionnaire should thus be complemented with more questions on these concepts.

That feedback would belong to the second latent variable in the models, training in the simulator, may seem reasonable. This was also tested, but this model did significantly differ from empirical data. It is not that fidelity is uncorrelated to learning in the simulator (r = 42), and the factor effect between feedback and training in the simulator was high (.74). The reason that this model differed from empirical data was that the relation between feedback and the variables reflecting the training environment was stronger than its relation with learning. Viewing feedback as a part of the training environment is however not unreasonable.

Compared to the pervious more general model on amalgamated data from four simulators (Oskarsson et al., 2010), this model was created on data from one single simulator (CV90), and can thus be expected to be more valid and reliable. However, the training tasks differ between the early and late phases of the education, and the three positions in the CV90 perform different tasks. With more data from each position, and each phase of education, it would thus be possible to test if the model fits each position and phase of education.

Some may think that the LISREL models are obvious, and not provide anything new. In that case the models have the advantage of confirming the relations with statistical methods. That conceptually sensible LISREL models could be created also provide a certain amount of reliability to the results, including those of the AVOVAs. If the questionnaire had bee too poor, or if the participants had answered too carelessly or arbitrarily no sensible LISREL models could have been created.

Data reduction, in this study component analysis, has the advantage of reducing a larger number of manifest variables into a smaller number of latent dimensions or factors. These underlying factors often provide a truer picture and are thus more reliable than the individual variables from which they were formed (e.g. Svensson, 2003; Pett, Lackey & Sullivan, 2003).

The reliability analysis of the questions of the training concepts will be considered for further development of the questionnaire. More questions will also be added, especially concerning training and the effect of training in the real CV90.

This study builds on subjective rating data and answers on open questions. Inferences made on rating data imply certain limitations. Task performance in the real CV90 could of course have been measured quantitatively, but since training was performed in both the simulator and in the real CV90, it had had been very difficult to know how much of the training effect that could be ascribed to the simulator training. Furthermore, with quantitative performance measures, a large amount of measurements for a range of tasks had been necessary, which had made it impossible to perform the study within budget limits. In addition, feeling of presence, fun and motivation are abstract concepts and thus not suited for quantitative performance measures. Previous research has however shown high correlations between subjective self ratings of abstract concepts as mental workload and situation awareness with psychophysiological measures and task performance (e.g. Castro et al., 2003). This indicates reliability of subjective assessments of abstract concepts. Nevertheless, a disadvantage with subjective rating data is that they often only show changes, and for example if something is good or bad. To isolate why changes occur and why something is good or bad, a more qualitative approach is generally needed. This can, for example, be performed by structured interviews, focus groups and participatory on sight observations.

5 Conclusions

The gunners used the simulator most, and the drivers least.

There was a shift in the gunners' training from a higher amount of training of basic handling and preparation before training in reality in the early phase towards more training of team cooperation in the later phase.

The gunners were more positive to the simulator than the other positions, especially concerning how much they learned in the simulator and the effect of the simulator training in the real CV90.

The drivers rated learning and effect in the real CV90 low. They also commented the lack of realism of the driving task.

The gunners rated learning in the simulator higher in the early than in the later phase of their education, otherwise there were no differences between the two training phases.

There were many complaints on the graphics and the low correspondence with reality. Still, the ratings of learning and its effect on performance in the real CV90 were relatively high. This indicates the possibility of receiving training effects with simple low fidelity simulators.

The open questions indicate the need of improving certain parts of the simulator, especially concerning the driver's task.

The LISREL model statistically confirms that the empirical data shows that the training environment has a positive effect on the training situation.

6 References

Alessi, S. M. (1988). Fidelity in the design of instructional simulation. *Journal of computer-based instruction, 15*(2), 40-47.

Borgvall, J., Castor, M., Nählinder, S., Oskarsson, P.-A., & Svensson, E. (2007). *Transfer of training in military aviation* (FOI-R--2378--SE). Linköping: Swedish Defence Research Agency.

Castor, M., Hanson, E., Svensson, E., Nählinder, S., LeBlaye, P., MacLeod, I., et al. (2003). *GARTEUR Handbook of mental workload measurement* (No. GARTEUR FM AG13 Final Report - GARTEUR TP 145).

Farmer, E., van Rooij, J., Riemersma, J., Jorna, P., & Moraal, J. (1999). *Handbook of Simulator-Based Training.* Ashgate Publishing Limited: Aldershot, England.

Garris, R., Ahlers, R., & Driskel, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441-467.

Jöreskog, K., & Sörbom, D. (1998). *LISREL 8:Structural Equation Modeling with the SIMPLIS Command Language.* Hove: SSI Scientific Software International.

Lee, A. T. (2005). *Flight Simulation Virtual Environments in Aviation*. Ashgate Publishing Limited, Aldershot, England.

Liu, D., Blickensderfer, E. L., Macchiarella, N. D., & Vincenzi, D. A. (2008a). Transfer of training. In D. A. Wincenzi & J. A. Wise & M. Moulana & P. A. Hancock (Eds.), *Human factors in simulation and training* (pp. 49-60). Boca Raton, FL: CRC Press.

Nählinder S. (2009). *Flight Simulator Training: Assessing the Potential*. [Thesis]. Linköping: Linköping University Electronic Press. Linköping Studies in Science and Technology. Dissertations, 1250.

Nählinder, S., Borgvall, J., & Svensson, E. (2004). Using embedded pedagogical tools for optimizing within visual range combat training - the initial experience. In *Proceedings of the second Swedish-American workshop on modelling and simulation*, Cocoa Beach, FL (pp. 158-163).

Nählinder, S., Oskarsson, P.-A., Lindahl, B., Hedström, J., & Berggren, P. (2009). *Effects of simulator training - motivating factors* (FOI-R--2926--SE). Linköping, Sweden: FOI, Swedish Defence Research Agency.

Oskarsson, P.-A., Nählinder, S., & Svensson, E. (2010). A meta study of transfer of training. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, San Francisco, CA* (pp. 2422-2426). Santa Monica, CA: Human Factors and Ergonomics Society.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis - The use of factor analysis for instrument development in health care research*. London: SAGE.

Roscoe, S. N., & Williges, B. H. (1980). Measurement of transfer of training. In S. N. Roscoe (Ed.), *Aviation Psychology* (pp. 182-193). Ames: Iowa State University Press.

Stanton, N. (1996). Simulators: a review of research and practice. In N. Stanton (Ed.), *Human factors in nuclear safety* (pp. 117-140). London: Taylor & Frances.

Svensson, E. (2003). Statistical techniques for data reduction and modelling. In M. Castor & E. Hanson & E. Svensson & S. Nählinder & P. LeBlaye & I. MacLeod & N. Wright & J. Alfredson & L. Ågren & P. Berggren & V. Juppet & B. Hilburn & K. Ohlsson (Eds.), *GARTEUR Handbook of mental workload measurement* (pp. 122-132): GARTEUR FM AG13 Final Report - GARTEUR TP 145.

Wiese, E. E., Freeman, J., Salter, W. J., Stelzer, E. M., & Jackson, C. (2008). Distributed after-action review for simulation-based training. In D. A. Vincenzi & J. A. Wise & M. Moulana & P. A. Hancock (Eds.), *Human factors in simulation and training* (pp. 287-299). Boca Raton, FL: CRC Press.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*.