International Journal of Aviation Research

Volume 14 | Issue 01 Peer-Reviewed Article

Qualifying Eurofighter Aircrew Synthetic Training System as FFS-C/D Simulator

Daniel Moldrings
Embry-Riddle Aeronautical University

Hans-Joachim Ruff-Stahl
Embry-Riddle Aeronautical University

The German Air Force uses ASTA flight simulators for training and testing of Eurofighter pilots. This study aimed to determine whether the ASTA simulator meets the requirements for the subjective test of certification and whether the pilots consider ASTA a realistic training tool. 118 of a total of 130 Eurofighter pilots participated in the study and were asked to answer a questionnaire after flying a certification-type mission profile. Nine hypotheses were tested with inferential and descriptive statistics, the results led to four conclusions. First, the ASTA is perceived as a high-fidelity training simulator that confidently replicates the Eurofighter. Therefore second, the ASTA simulator's qualification process to an FFS C/D should be initiated. Third, motion systems do not seem indispensable for the transferability of simulation; thus finally, the rulemaker's qualification requirements to approve FFS C/D simulators should be updated to include ASTA type simulators.

Recommended Citation:
Moldrings, D., & Ruff-Stahl, H.-J. Qualifying Eurofighter aircrew synthetic training aids as full flight simulator level C/D. *Collegiate Aviation Review International*, *14*(1), 74-95.

In 2003, the Eurofighter 4th generation high-performance combat aircraft entered operational service in the German Air Force (GAF). To support aircrew training and testing, corresponding high-fidelity Aircrew Synthetic Training System (ASTA) flight simulators were procured (Timmermann, 1988) and until 2009, all four GAF Eurofighter Wings were equipped with this system.

In general, the quality of simulators varies from low to high fidelity, where the term “quality” may be understood as the extent to which the simulator enables a positive transfer of training to the real airplane. Current views argue that a mix of high physical, functional, and cognitive fidelity provides the best transfer of training effect (Allen, Hays, & Buffardi, 1986). Yet, only high fidelity simulators, which include a full motion feature, can be certified as a Full Flight Simulator (FFS) C/D by rulemakers such as the EASA, and only certified FFS C/D may be used for aircrew licensing.

Nevertheless, recent studies on motion cue fidelity effects on training transfer turned out to be inconclusive or even showed contradictory results (Myers, Starr, & Mullins, 2018). Thus, it is unknown whether motion cues are actually needed for the positive transfer of training and, in more general terms, whether a motion feature is required to simulate reality at all. This question is significant because if high fidelity simulators such as the ASTA could be certified as FFS C/D, these simulators without a full motion feature would be much more cost-effective and less maintenance intensive for the operators.

Certified FFS C/D simulators are considered Zero Flight Time (ZFT) simulators and hours flown in them fully count as regular flight hours. The GAF is already using the ASTA simulator as a ZFT simulator in certain training areas, yet without certification by the civilian rulemaker. While the GAF commander’s intent is to initiate the process of certification as soon as possible, it is yet unknown how well the ASTA simulator replicates the Eurofighter’s flight characteristics in the pilots’ eyes and thus, whether the ASTA enjoys a high enough fidelity. Consequently, this study aims to analyze pilots’ attitudes toward fidelity, reputation, and realism of the ASTA. It is aimed to find out whether the ASTA presents an adequate training system for type- and instrument rating flight training. Based on the results of this study, the decision will be made whether or not to initiate a civilian certification process.

Training Transfer

The purpose of any virtual flight training is to improve pilot performance in real-world tasks (Stanney, 2002). It can be achieved by providing the pilot with sufficient adequate simulation fidelity (Vincenzi, 2009), whereby fidelity describes the simulator’s conformity to the real aircraft and the resemblance of the aircraft’s behavior. Fidelity directly influences training transfer (Allen et al., 1986), which is understood as the process by which knowledge, skills, and attitude acquired during simulator flight training improve real-world performance (Hochmitz & Yuviler-Gavish, 2011). Yet, a negative transfer of training is also possible, which worsens real-world performance after simulator training (Myers et al., 2018; Swezey & Andrews, 2001) and may result in unsafe, even catastrophic situations (Lee, 2017).

Fidelity has cognitive, functional, and physical aspects, where physical fidelity describes a simulator's replication of the actual aircraft cockpit and its associated controls, switches, and indicators – including motion, visuals, and sounds (Allen et al., 1986). But there are limits to physical fidelity. Vaden and Hall (2005) point out that ideally, the simulator's motion cueing system would produce the same physiological cues as in actual flight, which however might not be achievable even with the most sophisticated motion systems. Therefore, stimulating motion perception rather than physical motion might be the better alternative (Adelstein, Lee, & Ellis, 2003). Recent studies indicate, the influence of mechanical versions of motion fidelity on training transfer is complex and sometimes even shows contradictory results (Myers et al., 2018; Vaden & Hall, 2005). Nevertheless, FFS C/D certification without motion feature is still impossible and all certifiable motion systems must produce similar cues as a six-degrees-of-freedom synergistic platform motion system (EASA Certification Specifications for Airplane Flight Simulation Training Devices CS-FSTD(A), 2018).

Apart from physical fidelity, functional fidelity defines the level to which the simulator reflects aircraft and human-machine interphase (HMI) functionalities and provides real-world cockpit environment stimuli (Allen et al., 1986). Cognitive fidelity, in turn, describes the degree to which a simulator replicates real-world psychological and cognitive demands and factors, to include anxiety, stress, situational awareness (SA), SA breakdown, and decision-making processes (Taber, 2014). It is widely agreed that a complementary mixture of high physical fidelity, functional fidelity, and cognitive fidelity training methods will result in a better training transfer (Hochmitz & Yuviler-Gavish, 2011).

Ground-Based Motion Simulation Technology

Due to a motion system being essential for certification as an FFS C/D, a closer look at these systems is necessary. Most simulators employ a six-degree of freedom (six-DOF) platform (hexapod), which is actuated by six hydraulic legs that produce motion cues. Yet, the actuators' linear displacement capacity is typically limited to 3 to 6ft, which defines the limits of all motion cues in duration and amplitude. Various washout algorithms have been implemented to generate motion cues within these constraints (Bürki-Cohen, Sparko, & Go, 2007). Other systems generate better motion cues, such as a vertical motion simulator (VMS), a centrifuge motion simulator (CMS), a Desdemona system, or a Facsimile Simulation Seat.

For example, a VMS located at NASA's Ames Research Center can travel up to 60ft vertically and 40ft laterally. The motion platform consists of a four-DOF platform mounted on a two-DOF large-amplitude platform (Bürki-Cohen et al., 2007). The acceleration is limited to approximately 22ft/s^2, or almost 0.75g (Garud-Barna, 2018). A CMS typically consists of a balanced cantilever arm, which is connected to a dual gimbal three-DOF gondola and can reach and sustain a high level of acceleration of up to 15g. However, such a centrifuge motion system causes a vestibular motion effect generated by pilots' head movements (Guedry & Benson, 1976), which is unrealistic.

The Desdemona simulator motion system combines a six-DOF platform motion system with the centrifuge motion concept. The cockpit is fully gimbaled, allowing unlimited rotation in all directions. A four-cascaded degree of freedom platform (360° of yaw, pitch and roll rotation, and 2m heave) is attached to a longitudinal track (8m) which rotates around the vertical axis, adding two extra synergistic degrees of freedom. A rotation of the track generates centrifugation up to 3g (Bles & Groen, 2009).

In contrast, a Facsimile Simulation Seat with G-cueing simulates g-force and vibration by changing the seat pad's shape, altering the tension on the seat harness straps, and raising or lowering the seat's height. Longitudinal acceleration is simulated by moving the back support cushion forward and aft, thereby giving the occupant a surge sensation. Lateral acceleration is simulated by moving the back cushion sideways for a sway sensation (Thales Training & Simulation, 2014). The G-cueing seat mimics the tactile feel of an accelerating aircraft through a varying contact between the pilot and the seat (Barrett-Schmidt, 1998). Additionally, G-cueing seat systems used in military fast jet fighter aircraft can connect inflatable anti-G suits for additional cutaneous tactile receptors stimulation.

In sum, the Hexapod, VMS, and Desdemona motion systems offer limited use for military fighter aircraft simulators due to their restricted physical performance characteristics. The CMS, in turn, may be able to accurately simulate g-load but would induce an unrealistic vestibular motion effect and is highly expensive. Additionally, CMS usually comes without a cockpit replica due to weight and cost concerns. Thus, it seems the Facsimile simulation seat with g-cueing currently offers the most efficient solution for high-performance aircraft simulators because it combines a high-fidelity cockpit with perceptual motion sensation.

Influence of Motion Systems as a Singular Variable on Training Transfer

Studies on the influence of motion systems as a singular variable for training transfer are inconclusive and show sometimes even contradictory effects. Due to their kinematic and dynamic constraints, motion systems offer a limited representation of the acceleration characteristics, which a real aircraft will produce (Myers et al., 2018; Winter, Dodou, & Mulder, 2012). The Eurofighter, for example, is capable of attaining -3gz to + 9gz (Rosenkranz, 2012), which cannot be achieved by any of the modern motion systems except by a CMS. Yet, if high fidelity motion simulation is impossible, how much motion fidelity is needed for a positive transfer of training from the simulator to the real aircraft?

Already more than a decade ago, Sparko and Bürki-Cohnen (2010) compared the transfer of training of an FFS with Hexapod motion system to a high-fidelity flight training device (FTD) that simulates motion with a Facsimile G-seat, including heave motion vibration cues and a high-level visual system. Results indicated that a seat-motion system, limited to only one DOF provides motion cues to the same qualitative level as an FFS (Sparko et al., 2010). Yet, in a comprehensive meta-analysis of transfer of training experiments using FFS motion as an independent variable, Winter et al. (2012) found a slightly more positive transfer effect in favor of a motion system, specifically in terms of true transfer and quasi transfer, where true transfer refers to training transfer from the simulator to the aircraft and quasi-transfer from an FTD to an FFS of the same aircraft. (Vaden & Hall, 2005; Winter et al., 2012; Zaal, 2019).

Overall, while motion systems may generally influence training transfer positively, FTDs using facsimile simulation seats with G-cueing and high fidelity visual systems provide motion cues almost as effective as FFSs using a six-DOF Hexapod (Sparko et al., 2010).

Multimodal Perception of Motion Cues

Earlier, Bürki-Cohen et al. (2007) have explained why the motion cues in an FTD with G-cueing might be almost as effective as a traditional motion simulation by hexapod. Typically, visual impressions are perceived by the eye, motion/sound by the vestibular system, and control input devices' resistance such as throttle and stick by mechanoreceptors located in the skin and muscles. Airplane motion is sensed by multiple receptors (Barlow & Mollon, 1982). The photosensitive receptors in the retina of the eye respond to changes in size, shape, texture, and changes in the position of objects, while the vestibular apparatus' mechanoreceptors in the inner ear send direct signals to the brain when they are deflected due to head accelerations.

Furthermore, pilots perceive airplane motion via additional mechanoreceptors such as tactile and proprioceptive sensors, which are distributed over the entire human body (Bürki-Cohen et al., 2007). Therefore, "simulator motion can be perceived without actual physical motion stimulating the vestibular and the proprioceptive systems, as long as the eye perceives the changes that would be expected from motion in an out-the-window view" (Bürki-Cohen et al., 2007, p. 3). This visually induced illusory motion is also known as vection (Hettinger, Berbaum, Kennedy, Dunlap, & Nolan, 1990). A well-coordinated mix of stimulation of photosensitive receptors and mechanoreceptors results in an adequate simulator motion perception.

Hence, physical and perceptual motion fidelity must be viewed as two different categories when considering motion-fidelity requirements. Physical motion fidelity is defined as the match between motion cues in the simulator and the aircraft. In contrast, perceptual motion fidelity describes the match between the pilots' subjective perception of motion between the simulator and the aircraft (Bürki-Cohen et al., 2007).

EASA Regulations Regarding FFS C/D

According to EASA rules, the respective FFS C/D shall be evaluated against applicable certification specifications for Airplane Flight Simulation Training Devices CS-FSTD (A) criteria for an initial evaluation in objective, subjective, and functional tests. (EASA Certification Specifications for Airplane Flight Simulation Training Devices CS-FSTD (A), 2018). While perceptual and functional fidelity can be measured objectively, the impression of motion fidelity in an ASTA can only be assessed subjectively. EASA regulation Easy Access Rules for Authority Requirements for Aircrew (Part-ARA) defines a mission profile for such a subjective test (Figure 1). The test consists of standard procedures, system malfunctions, and unusual attitude maneuvering (EASA Easy Access Rules for Authority Requirements for Aircrew (Part-ARA), 2019).

Figure 1

FSTD Flight Profile for FFS C/D Qualification

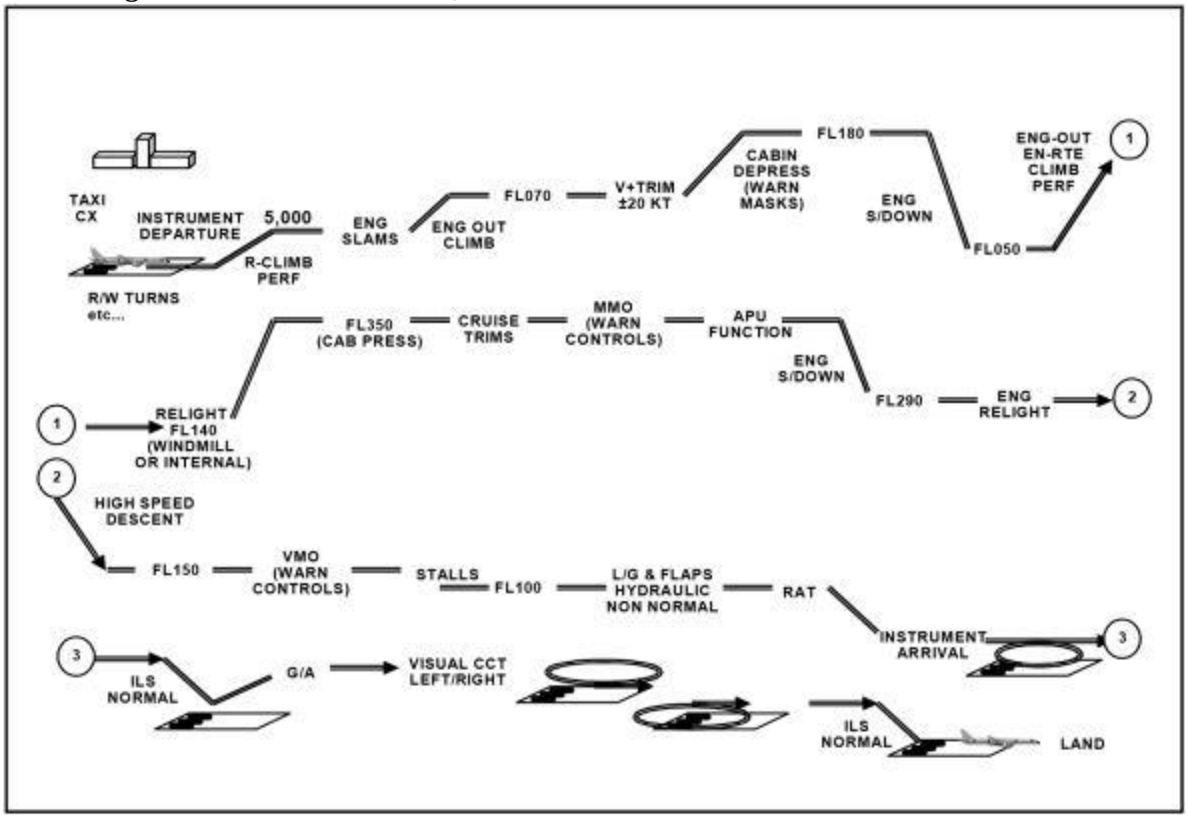


Figure 1. Typical flight profile of an FSTD qualification mission. This figure depicts the typical mission flight profile of an initial FFS C/D subjective test. The specific mission content is integrated into the mission profile. Adapted from “Easy Access Rules for Authority Requirements for Aircrew (Part-ARA),” by EASA, 2019.

German Air Force Regulations

In the German Airforce, simulators are used for the licensing and training of mission-specific and tactical procedures (LufABw, Qualifikation und Überwachung von Flugsimulatoren militärischer Luftfahrzeuge, A1-271/0-8901, 2019), as defined in the annual Tactical Combat Training-Program (TCTP). Combat wing pilots are supposed to fly 40 hours per year in the simulator (LwTrKdo, Tactical Combat Training Programm Strahlgetriebene Kampfflugzeuge, C2-271/0-2000-29, 2019; NATO ACO Forces Standards Vol III - Standards For Air Forces Partner Version, 2013).

Primary flight training includes type- and instrument rating missions. But to date, hours flown in the simulator cannot be logged as flying hours according to EASA rules, because the ASTA is not yet a qualified FFS C/D. On the other hand, the EASA acknowledges that not every civilian regulation may be transferable to military requirements. Specifically, “for non-commercial operations, the operational and licensing rules should be tailored to the complexity of the aircraft and a related definition should be set out” (EC Regulation No 216/2008 of the

European Parliament and of the Council, 2008). In short, EASA regulations can be tailored to operational needs.

Based on this exception, the German Air Force conducts annual instrument check rides in the ASTA simulators in accordance with a specified mission card that corresponds to EASA regulations and requirements and is therefore the basis for this study’s subjective test plan. (LufABw, Lizenzierung von Personal bemannter Luftfahrzeuge, A1-271/4-8901, 2018; LufABw, Prüfungen des Personals bemannter und unbemannter Luftfahrzeuge, A1-271/5-8901, 2019; (LwTrKdo, Überprüfung der Luftfahrzeugführer bzw. Luftfahrzeugführerinnen von strahlgetriebenen Kampfflugzeugen der Luftwaffe - Eurofighter, C2-271/5-2000-5, 2020).

ASTA System

The current Eurofighter simulator system ASTA consists of two different simulators, a Full Mission Simulator (FMS) and a Cockpit Trainer/Interactive Pilot Station – Enhanced (CT/IPS-E). As the FMS, the CT/IPS-E can provide training for Eurofighter aircrew in all possible training segments, from routine and abnormal procedures until advanced tactical multi-ship mission rehearsal. Both systems use “re-hosted aircraft software,” which corresponds to the real aircraft and cockpit functionalities of FMS and CT/IPS-E reflect the aircraft’s features, including original aircraft controls and switches. And all aural warnings and the ambient noise match the real world. In a few and less relevant cases, the simulator features high-fidelity facsimile parts.

The following main differences exist between the two simulators: ASTA FMS incorporates a G-seat motion system (EUROFIGHTER GmbH, 2015), a 360° visual system that fulfills the EASA qualification criteria, and an original Eurofighter head-up display (HUD). The CT/IPS-E does not incorporate G-seat functionality, presents a 300° field of view (FOW), and uses HUD symbology that is projected onto the visual dome´s surface (EUROFIGHTER GmbH, 2015, Project Syntropy, 2019; ZEISS Deutschland [ZEISS Germany], 2020). Both types of visual systems and HUDs fulfill EASA qualification criteria.

Hypotheses

Since ASTA has not yet been officially licensed as FFS C/D, the question is whether ASTA can generate a positive transfer of training for type- and instrument rating flying in the view of Eurofighter pilots. Overall, it is hypothesized that the ASTA enables positive transfer of training due to its high physical, functional, and cognitive fidelity. Specifically, the following sub-hypotheses will be tested:

H1A: There is a significant difference between the look and feel of ASTA and Eurofighter HMI.

H2A: There are significant differences between ASTA and Eurofighter HMI functionalities.

H3A: There is a significant difference between the ASTA re-hosted aircraft software and the Eurofighter aircraft software.

H4A: There is a significant difference between the ASTA flight performance model and the flight characteristics of the Eurofighter.

H5A: There is a significant difference between the motion perception generated by the FMS and the motion perception generated by the CT.

H6A: There is a significant difference in rating the perceptual fidelity of ASTA concerning startup, taxi, and takeoff (STTO) between pilots with different Eurofighter specific experience levels.

H7A: There is a significant difference between the training of basic enroute elements conducted in the ASTA and the basic enroute elements conducted in real-world flying.

H8A: There is a significant difference in rating the perceptual fidelity of ASTA concerning landing, taxi, and shut-down between pilots with different Eurofighter specific experience levels.

H9A: There is a significant difference between the training of emergency procedure training elements conducted in the ASTA and the emergency procedure conducted in real-world flying.

Methodology

The missions were conducted in the ASTA simulators used by the German Air Force in its four wings. An anonymous survey with closed-ended questions was handed out to active-duty German Air Force Eurofighter pilots after a type- and instrument-rating mission. First, the pilots were briefed and questions answered before they flew the mission in only one of the two simulators. The mission content corresponded to a test procedure for the subjective qualification of an FFS C/D. After the flight, the pilots rated the perceived fidelity compared to the real aircraft. The questionnaire focused on HMI, the look and feel, and flight performance. The two types of simulators were not compared against each other, i.e. to find out which device would replicate the Eurofighter better than the other. Instead, the question was whether the flown simulator represented the aircraft in the eyes of the pilots and to what degree.

Each pilot filled out the questionnaire once. Besides basic information related to individual flight experience, specific questions had to be answered concerning the simulator’s perceptual fidelity. For this purpose, an even 6-Point-Likert-Scale was used, ranging from 1 (strongly disagree) until 6 (strongly agree) to avoid neutral answers. A written consent regarding the use of the data for the study is available for each participating pilot. The survey answers were analyzed quantitatively and sorted by experience levels. The data for the statistical tests originates in the statements about ASTA’s perceptual fidelity, level of experience, area of work, and teaching qualifications. Before the statistical analyses, the anonymous questionnaires were digitized and stored in a secure military network. The findings of the study were then forwarded to the German Air Force Command for further use.

The sample size of this study consists of 118 active-duty, Eurofighter qualified fighter pilots (59 novice, 47 intermediate, 12 expert). The total amount of active duty Eurofighter pilots within the German Air Force is 130. Hence, this study's sample represents 90.76% of the population of German Air Force Eurofighter pilots and thus conveys an unprecedented and accurate picture of ASTA's fidelity. The participating pilots' flying hours were coded into three experience levels: Novice (0 – 399 hours Eurofighter flying time), Intermediate (400 – 999 hours Eurofighter flying time), and Expert (> 1000 hours Eurofighter flying time). Pilots with a minimum of 400 hours in fighter aircraft become eligible for instructor pilot training; hence, they are not novices anymore. Likewise, a minimum of 1000 hours in fighters is generally required in the GAF to qualify for expert ratings such as 4-ship leader and mission commander. This differentiation is essential because a high-fidelity simulator should provide all pilots with the closest possible realism measured against the real aircraft – regardless of flying experience.

Perfect et al. (2014) developed a methodology to evaluate the fidelity of flight simulators, which includes an analysis of overall fidelity with quantitative metrics and the analysis of perceptual fidelity via subjective fidelity ratings. This procedure corresponds to EASA requirements for simulator qualification procedures (EASA Easy Access Rules for Authority Requirements for Aircrew (Part-ARA), 2019). Yet, EASA does neither specify a rating scale for the subjective test nor the sample size. Hence, the researcher followed the suggested method by Perfect et al. (2014) for a subjective fidelity rating. The experiment's consistency is given by the use of a standardized mission content and therefore, the results should be reproducible.

While hypotheses H1 until H4, H7, and H9 were analyzed with descriptive statistics, hypotheses H5, H6, and H8 were analyzed with a parametric statistical test. For H5, the Student´s T-test and H6 and H8, the One-Way-Analysis of Variance (ANOVA) test were selected and computed with StatCrunch.

Hypothesis H5 was intended to demonstrate a possible system-related difference in the two simulators' perceived motion fidelity. The respective means of the independent variable cockpit with the corresponding motion simulation were compared. The FMS provides motion stimulation via the visual system and a facsimile seat motion system, whereas the CT provides motion stimulation via the visual system only. The specific statements that the pilots had to rate were: "The facsimile g-seat significantly enhances subjective motion perception" (FMS) and "The missing g-seat negatively influences subjective motion perception significantly" (CT). In the opposite sense, this means that an available facsimile g-seat would improve the CT's perceived motion. In order to compare the means of the statements regarding motion perception FMS and motion perception CT, the scores for CT had to be recorded reversely.

For a comparison of the independent variable cockpit means, the Two-Tailed Independent Sample T-test was selected as the appropriate test. The following assumptions had to be fulfilled: (a) normal distribution and (b) homogeneity of variance. Normal distribution was tested with the Shapiro-Wilck test (Nornadiah Mohd Razali & Yap Bee Wah, 2011) and homogeneity of the population's variance with Levene's test. In case of a violation of one or more assumptions for the T-test, the Mann-Whitney U-test was the nonparametric alternative (Privitera, 2018).

An analysis of variance (ANOVA) was conducted to determine whether there were significant differences in landing perceptual fidelity correlating with experience level (hypotheses H6 and H8). Because the independent variable experience level consists of 3 groups (Novice, Intermediate, and Expert), the One-Way-Between-Subjects-ANOVA is the correct parametric test, if the following three assumptions can be met: (a) normal distribution, (b) homogeneity of variance, and (c) independence. Normal distribution (a) was evaluated using the Shapiro-Wilck test (Nornadiah Mohd Razali & Yap Bee Wah, 2011) and the Q-Q scatterplot, homogeneity (b) was tested using Levene's test, and independence (c) was assured by not repeating measures on the same subject. The randomized design of the survey and observations were not correlated in time and space and thus, pseudoreplication was avoided. The Kruskal-Wallis H-test was elected as the appropriate nonparametric test alternative if the assumptions are violated (Privitera, 2018).

For the T-test, and given a desired statistical power of .95, a significance level of .05, and a large effect size of .8, the required sample size was 84. For the ANOVA test and a desired statistical power of .95, a significance level of .05, and a large effect size of .4, the required sample size was 102. Hence, the actual sample size of 118 participating pilots resulted in a predictive high statistical power.

The tests were based on Likert-Scale questionnaires, which are robust when used for applied parametric and nonparametric tests (Schrum, Johnson, Ghuy, & Gombolay, 2020). The Likert-items were deliberately chosen to cover all aspects of flight simulator fidelity, such as the Human-Machine-Interface, simulator performance compared to aircraft performance, motion simulation, takeoff/landing performance, simulator flight resident software in different phases of flight including emergency procedure handling.

While an objective test will compare simulator performance data against aircraft performance data, the Likert-Scale ratings provide subjective test results. Yet, the two tests complement each other and are equally important, because a positive transfer of training can be determined only if both objective and subjective data achieve conclusive results. The results, in turn, will serve as an essential data set for the acceptance analysis for ASTA to qualify as an FFS C/D simulator.

Results

H1 Hypothesis

It was hypothesized that there is no significant difference between ASTA and Eurofighter HMI's look and feel. Unless otherwise stated, all of the following hypotheses were tested with descriptive statistics, considering pilots' experience levels. Results show, 85.5% of the Eurofighter pilots disagree with the statement that the look and feel of ASTA HMI differs significantly from the Eurofighter HMI. The span of the result is from 89.8% to 81.2%.

For Novice pilots, the rating look and feel had an average of 2.39 (SD = 1.02, SEM = 0.13). For Intermediate, the rating of look and feel had an average of 2.17 (SD = 1.05, SEM = 0.15), whereas, for Expert, the rating of look and feel had an average of 2.08 (SD = 1.24, SEM = 0.36). Hence, we fail to reject the null hypothesis of no significant difference between ASTA HMI look and feel and Eurofighter HMI look and feel.

H2 Hypothesis

It was hypothesized that there is no significant difference between ASTA HMI functionalities and Eurofighter HMI functionalities. 89.8% of the Eurofighter pilots disagree with the statement that the ASTA HMI functionalities differ significantly from the Eurofighter HMI functionalities. The span of the result is from 94.3% to 85.3%.

For Novice, the rating of HMI functionality had an average of 2.07 (SD = 1.06, SEM = 0.14). For Intermediate, the rating of HMI functionality had an average of 2.04 (SD = 1.00, SEM = 0.15), and for Expert, the rating of HMI functionality had an average of 1.67 (SD = 0.49, SEM = 0.14). Therefore, we fail to reject the null hypothesis. This means that there is no significant difference between ASTA HMI functionalities and Eurofighter HMI functionalities.

H3 Hypothesis

It was hypothesized that there is no significant difference between the ASTA re-hosted aircraft software and the Eurofighter aircraft software. 88.2% of the Eurofighter pilots agree with the statement that the ASTA HMI functionalities do not differ significantly from the Eurofighter HMI functionalities. The span of the result is from 92.6% to 83.8%.

For Novice, the re-hosted aircraft software rating had an average of 4.59 (SD = 0.85, SEM = 0.11). For Intermediate, the re-hosted aircraft software rating had an average of 4.60 (SD = 1.01, SEM = 0.15). For Expert, the re-hosted aircraft software rating had an average of 4.92 (SD = 0.90, SEM = 0.26). Hence, we fail to reject the null hypothesis, indicating no significant difference between ASTA re-hosted aircraft software and the Eurofighter aircraft software.

H4 Hypothesis

It was hypothesized that there is no significant difference between the ASTA flight performance model and the Eurofighter flight performance. 80% of the Eurofighter pilots agree with the statement that the ASTA flight performance model does not differ significantly from the Eurofighter flight performance characteristics. The span of the result is from 84.0% to 76.0%.

For Novice, the ASTA flight performance model's observations had an average of 2.73 (SD = 1.23, SEM = 0.16). Intermediate pilots averaged at 2.49 (SD = 0.91, SEM = 0.13). Expert pilots rated the ASTA flight performance model with an average of 2.42 (SD = 1.31, SEM = 0.38). Hence, we fail to reject the null hypothesis, indicating that there is no significant difference between perceived ASTA flight performance and the performance characteristics of the Eurofighter.

H5 Hypothesis

It was hypothesized that there is no significant difference between the FMS's motion perception and the CT's motion perception. The hypothesis was tested with the Two-Tailed Independent Sample T-test. Shapiro-Wilk tests were conducted to determine the assumption of a normal distribution (Nornadiah Mohd Razali & Yap Bee Wah, 2011) and turned out significant results based on an alpha value of .05. For the FMS, $W = 0.89$, $p < .0001$ and for the CT, $W = 0.90$, $p < .0006$. Hence, the assumption of a normal distribution was violated and Levene's test was conducted. It showed non significant results based on an alpha value of .05, $F(1, 113) = 1.77$, $p = .1852$. Therefore, the assumption of homogeneity of variance was met.

Next, a two-tailed Mann-Whitney two-sample rank-sum test was conducted to examine whether there were significant differences in motion perception rating between the levels of cockpit. The result was not significant based on an alpha value of .05, $U = 3721.5$, $p = .538$, which suggests that the distribution of motion perception rating for the FMS was not significantly different from the distribution of motion perception rating for the CT. Hence, we fail to reject the null hypothesis, indicating that there is no significant difference between motion stimulation perception in the FMS and motion stimulation perception in the CT.

H6 Hypothesis

It was hypothesized that there is no significant difference in rating the perceptual fidelity of ASTA concerning startup, taxi, and takeoff (STTO) between pilots with different Eurofighter experience levels. The hypothesis was tested with the One-Way-Between-Subjects-ANOVA. A Shapiro-Wilck, after a Q-Q scatterplot showed an abnormal distribution based on an alpha value of .05: Novice $W = 0.86$, $p < .0001$, Intermediate $W = 0.75$, $p < .0001$, and Expert $W = 0.77$, $p < .0042$. A Levene's test demonstrated that the assumption of homogeneity of variance was met .05, $F(2, 114) = 0.47$, $p = .6286$. The ANOVA results were not significant, $F(2, 113) = 1.726$, $p = .183$, indicating the differences in STTO perceptual fidelity rating among the different experience levels were similar.

In addition, a Kruskal-Wallis H-test was conducted because it does not share the ANOVA's assumption of a normal distribution (Conover & Iman, 1981). The results were not significant either based on an alpha value of .05, $\chi 2(2) = 4.08$, $p = .13$, indicating that the mean rank of STTO fidelity rating was similar for each level of experience (Table 11). Because no significant difference was detected, a post-hoc analysis was not performed and we fail to reject the null hypothesis.

H7 Hypothesis

It was hypothesized that there is no significant difference between the training of basic enroute elements conducted in the ASTA and the basic enroute elements conducted in real-world flying. 100% of the Eurofighter pilots agreed with the statement that the training conducted in ASTA concerning instrument enroute navigation elements does not differ significantly from real-world aircraft handling and aircraft system behavior. The span of the result is from 100.0% to 95.0%.

For Novice, the rating of enroute elements showed an average of 1.72 (SD = 0.59, SEM = 0.08). Intermediates rated enroute elements at an average of 1.66 (SD = 0.56, SEM = 0.08), whereas Experts rated at an average of 1.58 (SD = 0.51, SEM = 0.15). Hence, no significant difference between the training of basic enroute elements conducted in the ASTA and the basic enroute elements conducted in real-world flying was diagnosed and we fail to reject the null hypothesis.

H8 Hypothesis

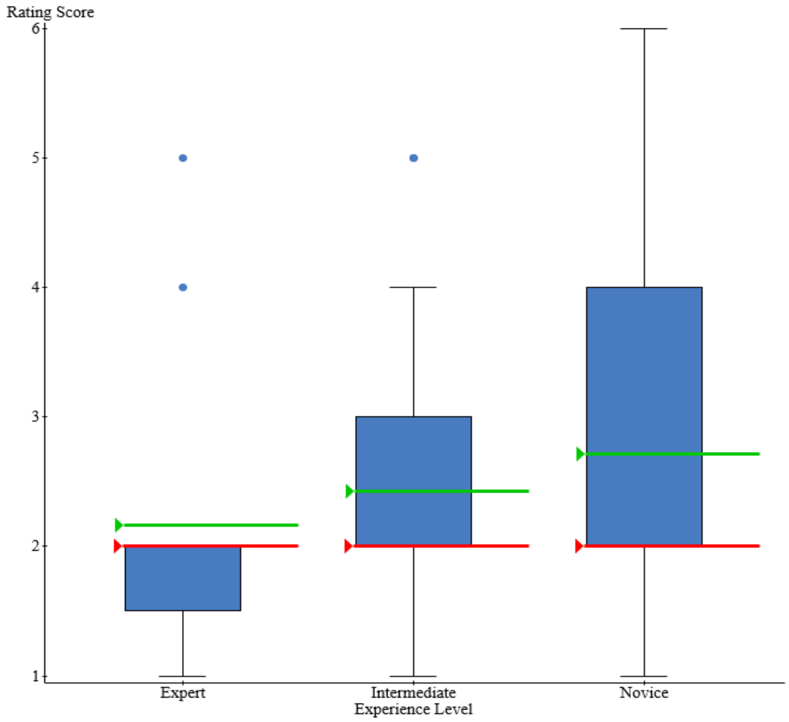
It was hypothesized that there is no significant difference in rating the perceptual fidelity of ASTA concerning landing between pilots with different Eurofighter experience levels. The hypothesis was tested with the One-Way-Between-Subjects-ANOVA. Similar to the treatment of H6, a Shapiro-Wilck test was performed to confirm a potential abnormal distribution indicated by a Q-Q scatterplot. Results showed a violation of normality assumption, which required a Levene's test to assess the assumption of homogeneity of variance. Based on an alpha value of .05, $F(2, 115) = 3.58$, $p = .0311$, the results were significant and showed a violation of variance homogeneity.

Nevertheless, an ANOVA test was performed and compared to the Kruskal-Wallis results. Based on an alpha value of 0.05, the ANOVA results were not significant, $F(2, 115) = 1.338$, $p = .267$, indicating the differences in STTO perceptual fidelity rating among the different experience levels were similar. The Kruskal-Wallis H-test, in turn, does not share the ANOVA's assumption of a normal distribution (Conover & Iman, 1981) and turned out non significant results (alpha .05, $\chi2(2) = 2.35$, $p = .31$) equally indicating that the mean rank of landing fidelity rating was similar for each level of experience. Since both tests, Kruskal-Wallis H-Test and ANOVA did not show significantly different ratings of ASTA's the perceptual fidelity between pilots with different Eurofighter experience levels concerning the landing phase, we fail to reject the null hypothesis.

The boxplot in Figure 2 displays the score distribution among the different experience levels and shows a difference especially between Novice and Expert. Thus, a two-tailed Mann-Whitney U-test test was performed post-hoc to compare the means of Novice and Expert. The result turned out to be non significant based on an alpha value of .05, $U = 347.0$, $p = .18$.

Figure 2

Boxplot of Landing Phase of Flight Rating by Experience Level



Note. The boxplot shows three outliers for the ratings of the pilots in the Intermediate group. The green line represents the Mean, and the red line indicates the Median.

H9 Hypothesis

It was hypothesized that there is no significant difference between the training of emergency procedure elements conducted in the ASTA and the emergency procedure conducted in real-world flying. 98.4% of the Eurofighter pilots agree with the statement that the training conducted in ASTA concerning instrument enroute navigation elements does not differ significantly from real-world aircraft handling and aircraft system behavior. The span of the result is from 100.0% to 93.5%.

For Novice, the rating of enroute elements had an average of 1.91 (SD = 0.63, SEM = 0.08). Intermediate rated enroute elements at an average of 2.23 (SD = 0.67, SEM = 0.10), whereas Expert rated at an average of 1.75 (SD = 0.45, SEM = 0.13); hence, we fail to reject the null hypothesis.

Other Results and Summary

H6 and H8 show a differing performance assessment of the ASTA for the takeoff and landing phases depending on the experience level of the pilots. Higher perceptual fidelity scores correlate with increasing flying experience levels. However, even the difference in scores between the two extremes Novice and Expert was statistically insignificant, as shown by the post hoc test of hypothesis 8.

The group of participating pilots was not only made up of squadron pilots. Pilots with management and staff functions up to the highest organizational levels also took part in the study. Thus, the two groups of student pilots and squadron pilots were merged post-hoc into one group, labeled line pilots. Furthermore, pilots with command and staff functions were merged into a group labeled leaders. The reason for this post-hoc analysis was to check for a potential management bias. These two groups were then compared using perceptual fidelity score means for the landing phase, because firstly, this phase is considered critical and must be replicated as realistically as possible and secondly, a potential management bias would have been noticeable in this specific flight phase precisely due to its criticality and the associated flying training effectiveness and flight safety considerations.

Overall, the results indicate no significant differences between perceptual fidelity in the simulator and the Eurofighter aircraft. Although motion systems are regarded as an indispensable component of physical fidelity, pilots of all experience levels rated the ASTA's physical fidelity good to very good in all areas. Moreover, the study found no significant difference in perceptual fidelity between a facsimile G-seat motion system in the FMS and no motion system at all in the CT. Specifically, a facsimile G-seat in the ASTA was viewed as only a slight improvement, if any, in motion perceptual fidelity. Likewise, pilots of all experience levels attested to the ASTA's high perceptual fidelity, too. Unexpectedly, the study showed a correlation of higher experience levels and higher perceptual fidelity ratings – particularly in the critical phases of take-off and landing.

Discussion

This study's research question was whether ASTA is an adequate training system for type rating and instrument rating missions and therefore, whether ASTA would meet the requirements as an FFS C/D simulator. According to EASA regulations, the qualification process must include a subjective test, which is typically performed by a single person. But a single, subjective test does not reflect the heterogeneous group of fighter pilots in reality and thus must be considered as unreliable. In contrast, this study provides an unprecedented comprehensive test across all levels of pilot experience.

The study's most significant finding beyond the immediate results is that the requirements for a qualification as an FFS C/D may not be valid anymore. In motion cueing, a facsimile G-motion system improves motion perception well enough to be highly transferable to a fighter aircraft in the real world. Regarding multimodal motion perception, both the ASTA and the CT were found to be adequate and confirm previous studies such as and Bürki-Cohen et al. (2007).

Overall, these findings suggest a new view on the cost-benefit ratios of simulators. While motion cueing systems are highly expensive in development and maintenance, these systems only offer a minimal improvement in the transfer of training compared to systems such as the ASTA and the CT. In short, if the quality of simulation and transferability of training are the main objectives, a motion-cued simulator may be an investment, where an at best questionable improvement in training transfer stands against high procurement and maintenance costs.

Conclusion

The ASTA simulator is an excellent training device that enjoys a very high reputation among Eurofighter pilots in terms of type rating and instrument flying training. Problem areas associated with the technology were not identified in this study. Inexperienced, experienced, and command & staff pilots agree that there are no significant differences in fidelity between the subsystems FMS and CT. Moreover, no significant benefit of tactile g-motion stimulation in the FMS – which is not available in the CT – could be detected. In short and in the view of the pilots, the ASTA FMS and the CT fulfill the requirements for a FFS C/D simulator. Thus, it should be certified as such. The impressive sample size of 118 participants out of a population of 130 licensed and qualified Eurofighter pilots in the GAF convincingly make this case.

These results are particularly eminent because outdated EASA regulations still dictate the use of highly expensive moving motion platforms, although more cost-effective alternatives are available, as shown in this study. The EASA regulations in question seem to ignore recent advancements in realistic perceptual fidelity and the insight that perception shapes motion cueing considerably more than functional fidelity – even in highly maneuverable fighter jets. Clearly, EASA rules are designed for commercial air traffic. However, it may be concluded that a simulator with motion fidelity is an expensive and unnecessary investment particularly for applications with limited maneuvering. Hence, these rules should change accordingly and regard simulators with high perceptual fidelity certifiable as FFS C/D simulators.

Finally, qualification regulations need to be revised. System-specific certification specifications for qualification as an FFS C/D simulator must be developed, and their focus must be placed exclusively on the transfer of training. Budgets to develop, buy, and maintain costly motion platforms should be redirected to develop and buy better visual systems, which significantly influence a flight simulator’s multimodal perceptual fidelity.

References

Adelstein, B. D., Lee, T. G., & Ellis, S. R. (2003). Head Tracking Latency in Virtual Environments: Psychophysics and a Model. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 47(20), 2083–2087. https://doi.org/10.1177/154193120304702001

Allen, J. A., Hays, R. T., & Buffardi, L. C. (1986). Maintenance Training Simulator Fidelity and Individual Differences in Transfer of Training. Human Factors: The Journal of the Human Factors and Ergonomics Society, 28(5), 497–509. https://doi.org/10.1177/001872088602800501

Barlow, H. B., & Mollon, J. D. (1982). The senses. Cambridge texts in the pysiological sciences: Vol. 3. Cambridge, New York: Cambridge University Press.

Barrett-Schmidt, P. (1998). The effect of G-Seat Tactile Cuing on Linear Motion Perception: Submitted to the Department of Aeronautics and Astronautics in partial fulfillment of the requirements for the degree of Master of Science in Aeronautics and Astronautics at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY; February 1998.

Bles, W., & Groen, E. (2009). The DESDEMONA Motion Facility: Applications for Space Research. Microgravity Science and Technology, 21(4), 281–286. https://doi.org/10.1007/s12217-009-9120-1

Bongers, A., & Torres, J. L. (2014). Technological change in U.S. jet fighter aircraft. Research Policy, 43(9), 1570–1581. https://doi.org/10.1016/j.respol.2014.03.009

Bürki-Cohen, J., Sparko, A., & Go, T. (2007). Training Value of a Fixed-Base Flight Simulator with a Dynamic Seat. In AIAA Modeling and Simulation Technologies Conference and Exhibit (p. 271). Reston, Virigina: American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/6.2007-6564

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. Medical Education, 42(12), 1150–1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Conover, W. J., & Iman, R. L. (1981). Rank Transformations as a Bridge between Parametric and Nonparametric Statistics. The American Statistician, 35(3), 124–129. https://doi.org/10.1080/00031305.1981.10479327

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. Psychological Methods, 2(3), 292–307. https://doi.org/10.1037/1082-989X.2.3.292

EUROFIGHTER GmbH (2015). EF ASTA Device System/Subsystem Design Description for T2C. (SP-J-171-M-0132).

EASA Certification Specifications for Aeroplane Flight Simulation Training Devices CS-FSTD(A) (2018).

EASA Easy Access Rules for Authority Requirements for Aircrew (Part-ARA) (2019).

(EC) Regulation No 216/2008 of the European Parliament and of the Council (2008).

Garud-Barna, S. (2018). Vertical Motion Simulator (VMS) Technical Details. Retrieved from https://www.nasa.gov/simlabs/vms/technical-details

Ghauri, P. N., Grønhaug, K., & Strange, R. (2020). Research methods in business studies (Fifth edition). Cambridge: Cambridge University Press.

Guedry, F. E., & Benson, A. J. (1976). Coriolis Cross-Coupling Effects: Disorienting and Nauseogenic or Not. Retrieved from https://apps.dtic.mil/docs/citations/ADA036899

Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., & Nolan, M. D. (1990). Vection and simulator sickness. Military Psychology: The Official Journal of the Division of Military Psychology, American Psychological Association, 2(3), 171–181. https://doi.org/10.1207/s15327876mp0203_4

Hochmitz, I., & Yuviler-Gavish, N. (2011). Physical Fidelity Versus Cognitive Fidelity Training in Procedural Skills Acquisition - Ilanit Hochmitz, Nirit Yuviler-Gavish, 2011. Retrieved from https://journals-sagepub-com.ezproxy.libproxy.db.erau.edu/doi/10.1177/0018720811412777

Holbrook, H. A., & Cennamo, K. S. (2014). Effects of High-Fidelity Virtual Training Simulators on Learners' Self-Efficacy. International Journal of Gaming and Computer-Mediated Simulations, 6(2), 38–52. https://doi.org/10.4018/ijgcms.2014040104

Kirkpatrick, D. L. I. (2004). Trends in the costs of weapon systems and the consequences. Defence and Peace Economics, 15(3), 259–273. https://doi.org/10.1080/1024269032000123203

Lee, A. T. (2017). Flight Simulation: Routledge. https://doi.org/10.4324/9781315255217

Likert, R. (1932). A Technique for the Measurementof Attitudes. Retrieved from https://legacy.voteview.com/pdf/Likert_1932.pdf

LufABw, Lizenzierung von Personal bemannter Luftfahrzeuge [Licensing of Personnel of Manned Aircraft], A1-271/4-8901 (2018).

LufABw, Prüfungen des Personals bemannter und unbemannter Luftfahrzeuge [Evaluation of Personnel of Manned and Unmanned Aerial Vehicles], A1-271/5-8901 (2019).

LufABw, Qualifikation und Überwachung von Flugsimulatoren militärischer Luftfahrzeuge [Qualification and monitoring of flight simulators of military aircraft], A1-271/0-8901 (2019).

LwTrKdo, Tactical Combat Training Programm Strahlgetriebene Kampfflugzeuge [Tactical Combat Training Program Jet Propulsed Fighter Aircraft], C2-271/0-2000-29 (2019).

LwTrKdo, Überprüfung der Luftfahrzeugführer bzw. Luftfahrzeugführerinnen von strahlgetriebenen Kampfflugzeugen der Luftwaffe - Eurofighter [Checking of pilots of jet-powered combat aircraft of the German Air Force - Eurofighter], C2-271/5-2000-5 (2020).

Marken, R. S. (2007). Absorbing and developing qualified fighter pilots: The role of the advanced simulator. RAND Project Air Force. Santa Monica, CA: Rand Corp. Retrieved from http://www.jstor.org/stable/10.7249/mg597af https://doi.org/10.7249/mg597af

Mircioiu, C., & Atkinson, J. (2017). A Comparison of Parametric and Nonparametric Methods Applied to a Likert Scale. Pharmacy (Basel, Switzerland), 5(2). https://doi.org/10.3390/pharmacy5020026

Myers, P., Starr, A., & Mullins, K. (2018). Flight Simulator Fidelity, Training Transfer, and the Role of Instructors in Optimizing Learning. International Journal of Aviation, Aeronautics, and Aerospace. Advance online publication. https://doi.org/10.15394/ijaaa.2018.1203

Nieuwenhuizen, F. M., Mulder, M. [M.], van Paassen, M. M., & Bülthoff, H. H. (2013). Influences of Simulator Motion System Characteristics on Pilot Control Behavior. Journal of Guidance, Control, and Dynamics, 36(3), 667–676. https://doi.org/10.2514/1.59257

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. Advances in Health Sciences Education: Theory and Practice, 15(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Nornadiah Mohd Razali, & Yap Bee Wah (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests (Vol. 2). Retrieved from https://www.researchgate.net/publication/267205556_Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov_Lilliefors_and_Anderson-Darling_Tests

NATO ACO Forces Standards Vol III - Standards For Air Forces Partner Version (2013).

Perfect, P., Timson, E., White, M. D., Padfield, G. D., Erdos, R., & Gubbels, A. W. (2014). A rating scale for the subjective assessment of simulation fidelity. The Aeronautical Journal, 118(1206), 953–974. https://doi.org/10.1017/S0001924000009635

Privitera, G. J. (2018). Statistics for the behavioral sciences (Third Edition). Los Angeles: SAGE.

Privitera, G. J. (2020). Research methods for the behavioral sciences (Third edition). Los Angeles, London: SAGE.

Project Syntropy (2019). EUROFIGHTER Full Mission Simulatoren (Luftwaffe) [EUROFIGHTER Full Mission Simulators (Airforce)]. Retrieved from https://www.project-syntropy.de/portfolio-item/eurofighter-full-mission-simulatoren-luftwaffe/

Reiser Simulation and Training (2020). Simulation Seats. © 2020 reiser-st.com. Retrieved from https://www.reiser-st.com/components/simulation-seats/

Rosenkranz, M. (2012). Eurofighter Typhoon. Retrieved from http://eurofighter.airpower.at/technik-daten.htm

Schrum, M. L., Johnson, M., Ghuy, M., & Gombolay, M. C. (2020). Four Years in Review: Statistical Practices of Likert Scales in Human-Robot Interaction Studies. Advance online publication. https://doi.org/10.1145/3319502.3378178

Sparko, A., Bürki-Cohen, J., & Go, T. (2010). Transfer of Training from a Full-Flight Simulator Vs. a High-Level Flight-Training Device with a Dynamic Seat. In AIAA Modeling and Simulation Technologies Conference. Reston, Virigina: American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/6.2010-8218

Stanney, K. M. (2002). Handbook of virtual environments: Design, implementation, and applications. Human factors and ergonomics. Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10084539

Swezey, R. W., & Andrews, D. H. (2001). Readings in training and simulation: A 30-year perspective. Santa Monica, CA: Human Factors and Ergonomics Society.

Taber, M. J. (2014). Simulation fidelity and contextual interference in helicopter underwater egress training: An analysis of training and retention of egress skills. Safety Science, 62, 271–278. https://doi.org/10.1016/j.ssci.2013.08.019

Taherdoost, H. (2016). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. SSRN Electronic Journal. Advance online publication. https://doi.org/10.2139/ssrn.3205040

Thales Training & Simulation (2014). Prime Item Development Specification (PIDS) for the G Cueing Subsystem of the Motion ‘G’ Cueing & Life Support System (MQLS).

Timmermann, M. (1988). Jagdflugzeug 90 (JF90/EFA) [Interceptor Aircraft 90 (JF90/EFA)], Der Bundesminister der Verteidigung 110144/2.

Vaden, E. A., & Hall, S. (2005). The Effect of Simulator Platform Motion on Pilot Training Transfer: A Meta-Analysis. The International Journal of Aviation Psychology, 15(4), 375–393. https://doi.org/10.1207/s15327108ijap1504_5

Vincenzi, D. A. (2009). Human factors in simulation and training. Boca Raton: CRC Press. Retrieved from http://gbv.eblib.com/patron/FullRecord.aspx?p=381355

Winter, J. C. F. de, Dodou, D., & Mulder, M. [Max] (2012). Training Effectiveness of Whole Body Flight Simulator Motion: A Comprehensive Meta-Analysis. The International Journal of Aviation Psychology, 22(2), 164–183. https://doi.org/10.1080/10508414.2012.663247

Zaal, P. (2019). Motion Cueing for Stall Recovery Training in Commercial Transport Simulators. Advance online publication. https://doi.org/10.2514/6.2019-0979

ZEISS Deutschland (2020). ZEISS VELVET 1600 SIM VISIR. Retrieved from https://www.zeiss.de/simulation-projection-solutions/produkte-loesungen/zeiss-velvet-1600-sim-visir.html