Assessment of Confidence Impact on Pilot Training Performance

Gianna Avdic McIntire[1], Amy Dideriksen[2], Thomas Schnell[3], Colton Thompson[3], Katharine Woodruff[3], Jessica M. Greenwald[4]

[1]Collins Aerospace, Cedar Rapids, IA
[2]Collins Aerospace, Orlando, FL
[3]University of Iowa Operator Performance Laboratory, Iowa City, IA
[4]St. Ambrose University, Davenport, IA

gianna.avdic@collins.com, amy.dideriksen@collins.com, thomas-schnell@uiowa.edu, colton-thompson@uiowa.edu, katharine-woodruff@uiowa.edu, greenwaldjessica@sau.edu

ABSTRACT

One of the best predictors of student performance is an individual's belief or confidence that s/he possesses the necessary abilities and skills to complete a task within a context. Confidence is a key psychological variable and most frequently conceptualized as self-efficacy. A broader confidence construct is core confidence, conceptualized as a higher-order core construct that influences four manifestations, including hope, optimism, self-efficacy, and resilience (Stajkovic, 2006). Confidence is a key determinant to whether one will unleash existing potential, or hold it internally captive (Stajkovic, 2006). A confident individual is one who knows what to do, how to do it, holds positive future outcome expectations, and can bounce back from suboptimal outcomes.

Despite the vast body of research suggesting that one's confidence belief is the single best predictor of human performance, there is little evidence of its implementation in real world flight training environments. Instructional systems designers should design training solutions to maximize student confidence beliefs for optimal performance outcomes. In this study, we utilize subjective measures of state-based core confidence and objective measures of training effectiveness, measured through a combination of pilot task-performance and physiological measures of cognitive workload. The main focus of this research is to assess how core confidence relates to pilot task performance and cognitive workload, in a simulator and live flight environment. To understand the role of state-based core confidence in a training effectiveness framework, we associate it with our training effectiveness measures to determine how it changes through the training, how it varies between the simulator and live flight environments, and how its changes affect subsequent performance in both environments.

ABOUT THE AUTHORS

Dr. Gianna Avdic McIntire is a Principal Investigator and Systems Engineer in Collins Aerospace's Advanced Technologies Immersive and Autonomous Systems group. She has supported research in Training Effectiveness and Adaptive Learning, and is currently focused on manned and unmanned teaming (MUM-T) research. Gianna received her doctoral degree from St. Ambrose University's College of Business. Her doctoral dissertation research focused on exploration of theoretical domains and predictive power of specific and broad state-based confidence constructs including self-efficacy and core confidence.

Mrs. Amy Dideriksen, PMP is a Global Training Research manager in Mission Systems at Collins Aerospace with over thirty years of training experience and a background in Instructional Systems Design. She has a Master's of Science degree in Industrial Technology, specializing in Training and Development/ Human Resource Development. She is currently pursuing her doctorate in Industrial Engineering with a focus in Human Factors from the University of Iowa. She is the principal investigator in Advanced Technologies for research initiatives in cognitive State Assessment, Training Effectiveness, and Adaptive Learning. Her interests are in integrating digitally enabled technologies with simulation-based training solutions.

Dr. Thomas "Mach" Schnell is a Professor in Industrial and Mechanical Engineering with a specialization in Human Factors/Ergonomics at the University of Iowa. He is also the director and chief test pilot of the Operator Performance Laboratory (OPL). Tom has secondary appointments as professor in the departments of Electrical

Engineering and Neurology. He has an undergraduate degree in EE and an MS and Ph.D. in Industrial Engineering with a specialization in Human Factors. He has led over 282 human factors research projects with external funding of around $23 million and over 2,200 incident free flight test hours. He is a Commercial pilot with over 6,200 flight hours, research test pilot, and flight instructor with helicopter, jet, and glider ratings and has been actively involved in flight-testing using fixed wing and helicopter testbeds as a pilot and principal investigator since 2004. He is also an Unmanned Aircraft Systems (UAS) pilot with Beyond Visual Line of Sight (BVLOS) flight time. At OPL, he conducted research into the training effectiveness and skill development of UAS flight crew using cognitive, mission specific, and behavioral measures to assess the training readiness of remote pilots and sensor operators.

Mr. Colton Thompson is a Flight Test Engineer at the University of Iowa's Operator Performance Laboratory (OPL). He received his Bachelor of Science in Aerospace Engineering from Iowa State University in 2018 and began work at OPL shortly thereafter. He is currently working toward his Master of Science in the Industrial and Systems Engineering program at the University of Iowa, while working full time at OPL.

Miss. Katharine Woodruff is a Systems Engineer in Collins Aerospace's Advanced Technologies. She completed her Master of Science at the University of Iowa in the Industrial and Systems Engineering program with specialization in Human Factors and a minor in Psychology. She was a student-athlete for the University of Iowa on the Women's Varsity soccer team. While pursuing her Master's degree, Katharine worked as a graduate research assistant at the University of Iowa's Operator Performance Laboratory (OPL) where she supported research in pilot training effectiveness, physiological episode detection, and spatial disorientation prevention. Katharine's thesis focus was on electroencephalogram (EEG) signals. Her research will enhance the knowledge of pilot's brain waves during flight procedures.

Dr. Jessica M. Greenwald is an associate professor of Management at St. Ambrose University. She earned her Ph.D. in Business from the University of Wisconsin, Madison and holds an MBA and an undergraduate degree in Marketing from Southern Illinois University, Carbondale. Her research focuses on work motivation including core confidence, attribution theory, and goal orientation. She is published in journals, such as *Organizational Behavior and Human Decision Processes*, *Human Resource Management*, *Personality and Individual Differences*, and the *European Journal of Work and Organizational Psychology*. Prior to obtaining her Ph.D., Greenwald was a Senior Account Executive for ASAP Software, now known as Dell Software – USA.

Assessment of Confidence Impact on Pilot Training Performance

Gianna Avdic McIntire[1], Amy Dideriksen[2], Thomas Schnell[3], Colton Thompson[3], Katharine Woodruff[3], Jessica M. Greenwald[4]

[1]Collins Aerospace, Cedar Rapids, IA
[2]Collins Aerospace, Orlando, FL
[3]University of Iowa Operator Performance Laboratory, Iowa City, IA
[4]St. Ambrose University, Davenport, IA

gianna.avdic@collins.com, amy.dideriksen@collins.com, thomas-schnell@uiowa.edu, colton-thompson@uiowa.edu, katharine-woodruff@uiowa.edu, greenwaldjessica@sau.edu

INTRODUCTION

Rapid technological advancements in modeling and simulation capabilities continue to foster substantial growth in the use of simulation-based training technologies in the United States Department of Defense (DoD). Several trends within the simulation and training industry continue to emphasize the need for effective training that will lead to optimal performance outcomes in live operational environments. The optimization of performance outcomes requires effective designs for simulation-based training devices.

Prior studies suggest one's confidence belief to be the best single predictor of human performance. While knowledge and experience each play a role in one's performance achievement outcomes, neither is likely to predict whether one will incite an action toward a goal. One's perception of ability or confidence is the central mediating construct that determines one's attempts toward achievements (e.g., Bandura, 1977; Bandura, 1986; Ericsson et al., 1993). In other words, given actual knowledge and prior experience, one's performance outcomes will be strongly dependent on the perception of ability and the decision to initiate an action, persist at that action, and complete the action. Should s/he lack self-confidence, s/he will likely not initiate an action that would lead to performance (Bandura, 1986; Stajkovic, 2006). However, there is little evidence of its implementation in real world training environments for warfighter pilots. To optimize training and performance outcomes, instructional systems designers must understand the important implications of changes in one's confidence beliefs. Research has shown that high core confidence individuals need less feedback than low core confidence individuals, and thrive under conditions of high task complexity (Linderman-Hill & Greenwald, 2019). Gilstrap and Greenwald (2016) report that individuals with high core confidence remain more engaged in the task despite low psychological availability. Individuals with high core confidence may have higher responsibility (Holdorf & Greenwald, 2018) compared to individuals with low core confidence. Confidence predicts persistence on the training task despite the presence of setbacks as well as facilitates learning transfer (Sitzmann & Ely, 2011; Dierdorff, Surface, & Brown, 2010). Core confidence also relates to effective self-regulation (Stajkovic, Lee, Greenwald, & Raffiee, 2015) and therefore may lead to optimal cognitive workload. The research on core confidence could also help to better select and train warfighter pilots. As a first step, we need to understand how confidence relates to outcomes in this context.

In this study, we utilize subjective measures of state-based core confidence, and objective measures of training effectiveness assessed through a combination of pilot task-specific performance metrics and physiological measures of cognitive workload (Dideriksen et al., 2018). The focus of this research is to assess how core confidence relates to pilot task performance and cognitive workload, in a simulator and live flight environment. We associate the construct of confidence with training effectiveness measures to determine how it changes throughout the training process, how it varies between the simulator and live flight environments, and how its changes affect subsequent performance and cognitive workload in both environments.

This paper makes some references to the importance of appropriate levels of fidelity. In the context of simulation-based training devices and systems, the term fidelity refers to the level of realism and immersion to which simulation-based training devices or systems replicate the real environment. A prior fidelity assessment research study shows that training effectiveness is impacted by complex factors such as fidelity or realism, and field of view (Bellows et al., 2020). The Bellows et al. (2020) fidelity study used the same testbed and methodology as this study.

Researchers compared training effectiveness measures between novice EPs who trained using MR versus EPs who trained in a traditional procedural trainer. The transfer of learning was assessed in a live capstone flight. Results of the referenced fidelity study showed that the immersive nature of the MR environment yielded better preparation for the dynamics of live flight and increased training transfer, whereas the traditional procedural trainer induced negative training. For example, in the live flight pilots who trained with Coalescence performed significantly better, maintained a significantly lower cognitive workload, and exhibited significantly higher situational awareness ratings.

Findings of this study will suggest if confidence should be considered when designing the fidelity of simulation-based training devices and systems. We theoretically propose and provide empirical evidence of the relationship between confidence and fidelity. However, because we do not manipulate fidelity, we suggest but cannot ascertain causality. Future research in this domain may lead to additional recommendations on fidelity and confidence, selection of warfighter pilots, orientation materials, how feedback is given, and more.

Confidence and Performance Outcomes

Confidence is a motivational construct and key factor in an overall human competence system (Bandura, 1997). It is “the gap between having dreams and believing one can achieve them” (Johnson, 2017, p. 16). Confidence is most frequently conceptualized as self-efficacy (Bandura, 1997), and most recently as core confidence (Stajkovic, 2006). It represents one’s judgments about their abilities to execute courses of action in an effort to attain desired outcomes (Bandura, 1986). Confidence beliefs play an important role in performance outcomes. For example, individuals with relatively similar skillsets and abilities differ in their performances based on their confidence beliefs (Bandura, 1997). In other words, controlling for ability, this simple can-do belief leads to performance differentials. In the last fifty years, tens of thousands of studies have explored confidence; yet, little evidence exists of its incorporation in this context.

Formation of confidence can be attributed to internal and external causes. Internal causes may include personality traits, such as conscientiousness and higher core self-evaluations. External causes include prior mastery experiences that provide cues for how confident one should be (Bandura, 1986). For example, competence judgments could be the result of the fidelity of a simulation-based training device. A mastery experience in a simulated training environment could lead to a formation of confidence toward one’s live flight performance. However, if the simulated training environment is not adequately representative of the realistic challenges associated with live flight environments, that newly found confidence may be qualified as overconfidence.

Overconfidence and Performance Outcomes

Overconfidence can be the result of internal motivational biases because confidence provides psychological benefits (Dunning, Leuenberger, & Sherman, 1995). When one is confident others tend to perceive them more positively and as more qualified (even if this is not the case), and they may achieve higher status as a result (Anderson, Brion, Moore, & Kennedy, 2012). Another internal cause may be due to personality traits, such as narcissism and hubris, which are associated with an inflated, grandiose, and overconfident sense of self (Brennan & Conroy, 2013; Gentile et al., 2013).

Perceptions of competence play an important role in an individual’s performance estimates as well as the errors in estimations (Ehrlinger & Dunning, 2003). An accurate perception of competence is predictive of performance outcomes (Bandura, 1997; Stajkovic et al., 2015), whereas an inaccurate perception may not be predictive. For example, overconfidence can result from inaccurate perceptions of competence that do not align with objective performance outcomes (Anderson et al., 2012).

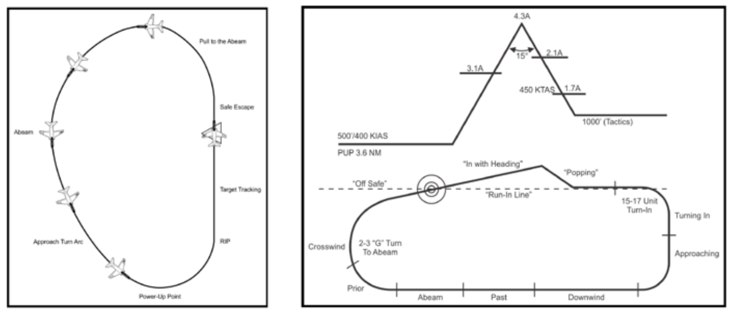
Collins Aerospace, in partnership with the Operator Performance Laboratory (OPL) at the University of Iowa, completed an exploratory, longitudinal study using a dependent sample of novice commercial pilots. The study associated self-reported confidence measures collected at three different times throughout the training process, with continuous objective measures of task-specific performance and electrocardiogram (ECG)-based cognitive workload measures collected during simulator training and in a live capstone flight. In collaboration with St. Ambrose University, the research team explored how confidence affected simulated training performance outcomes, how the simulated training performance environment may have contributed to overconfidence, and whether overconfidence is predictive of subsequent performance outcomes in a live flight environment.

Study Goals

This study explores the impact of confidence on simulation-based and live flight performance outcomes, examines if simulation-based training fidelity may associate with overconfidence toward live flight, and examines cognitive workload outcomes. We propose that if the simulation-based training device contributes to a formation of confidence that can be qualified as overconfidence toward live flight, that overconfidence will not predict subsequent live flight performance outcomes. Additionally, we explore the relationship between confidence and cognitive workload. Stajkovic et al. (2015) proposed that confident individuals are better able to self-regulate the attitudinal and behavioral psychosomatic processes involved in the pursuit of performance outcomes. As such, confidence may be predictive of participants in an optimal range of cognitive workload. However, with overconfidence such a relationship may not exist. This paper builds on work completed during the first three years of a multiyear research effort and reports the findings about how confidence assessment affects learning and performance outcomes.

METHOD

The research team created two simulated training bomb strike scenarios for this study, including Roll-In and Pop-Up patterns (Figure 1), using the CNATRA P-1209 Naval Air Training Command's Strike T-45 MPTS and IUT Flight Training Instruction publication. This document consists of a number of methods, strategies, and evaluation parameters appropriate for air-to-ground diving free-fall bombing missions. Given the length and complexity of the P-1209 document, it was not feasible to educate and train a novice pilot with no military experience in one day. Thus, the evaluation parameters were simplified to a level of appropriateness for novice pilots and test apparatus. A full analytical and experimental analysis was performed to substitute the L-29 flight parameters for the given T-45 parameters. After changing the given airspeeds and altitudes in the P-1209 diagrams, the necessary changes were made to ensure the geometry and kinematics of the bomb trajectory matched the new flight parameters. The Z-Diagrams used in the study reflected these changes. Before the study began, a full checkout flight was performed for all maneuvers to ensure they could be performed safely and accurately. The training scenarios included three distinct strike patterns (15-degree Roll-In, 30-degree Roll-In, and Pop-Up), and task performance was automatically scored. All training was conducted in the simulator, and the transfer of learning was assessed in the live capstone flight, where pilots performed the same scenarios practiced in the simulator.



CNATRA P-1209 Flight Training Instruction Strike T-45 MPTS and IUT

Figure 1. Roll-In (left) and Pop-Up (right) Patterns

Participants

Participants include twelve evaluation pilots (EPs) with low flight hours (less than 800 total) who volunteered to participate in the study. Eleven EPs were male, one was female, and all were between the ages of 20 and 36 holding a valid U.S. commercial pilot certificate. None had prior military experience.

Testbed

Like the prior two TE studies conducted at OPL, this study utilized the same L-29 aircraft-in-the-loop (AIL) during simulation and for live flight (Figure 2) (Dideriksen et al., 2018; Hoke et al., 2017), with two modifications aimed at increasing the realism for an out-the-window environment for our simulation-based training environment.

Figure 2. Aircraft-in-the-Loop Simulator (left), L-29 Jet (right)

Figure 3. MR Technology Coalescence

First, we integrated Collins Aerospace Coalescence™, a mixed reality (MR) technology, into the existing testbed (Figure 3). The mixed reality model used in this study included an Oculus Rift virtual reality headset with two USB cameras mounted on the front. With its 360° out-the window visuals, and hands-on visual access to real flight instruments, this technology allows for the pilots to see the virtual world around them while also being able to interact with the real flight controls and displays in front of them. The simulated visuals were generated by Collins Aerospace environmental EP2™ imagery and CORE™ simulation architecture software.

Second, we modified the hand on throttle-and-stick (HOTAS) to allow for a simulated bomb release. The simulator was configured to fly at Loom Lobby target range in Southern California. Loom Lobby is the range where the US Navy operates the T-45 Strike training program, on which this study was based. The simulator terrain included full elevation maps along with a detailed aerial target overlay imitating the real-world bullseye used by the Navy.

During the live flight, EPs used a Collins Aerospace F-35 helmet mounted display (HMD), which uses head-tracking and image projection to allow the EP to look around in the virtual world in live flight just as they did in the simulator. While an HMD system was used, the off-boresight symbology was deactivated. Instead, traditional F-18 HUD symbology was present and did not follow the head motion of the pilot. The HUD symbology was locked in place in the forward position where a traditional HUD would be found. This symbology was identical in the Coalescence trainer, procedure trainer, and live flight. Therefore, no modification of instruction was necessary.

Measures

We collected subjective (influenced by the observer's personal judgment) and objective (impartial, unbiased) measures during this study. Subjective measures were collected using self-reported surveys collected at three different times using well established and/or adapted psychometric scales. Objective measures were collected continuously during simulator training and live flight using aircraft state data and physiological-based assessment software.

Subjective Measures

Survey 1. This survey was used to collect state-based confidence measures and demographic information. It was completed at home prior to simulator training.

Survey 2. This survey was used to collect state-based confidence measures following simulator training and prior to live flight, thus was completed at home following simulator training and prior to the live flight.

Survey 3. This survey was used to collect state-based confidence measures like in survey 2, and it was completed at OPL following the live flight.

The descriptive statistics table illustrates confidence scores and changes over three points in time (Table 1).

Table 1

Descriptive Statistics

		State-based Confidence T1	State-based Confidence T2	State-based Confidence T3
N	Valid	12	12	9
	Missing	0	0	3
Mean		39.92	43.92	42.67
Median		39.50	43.00	44.00
Std. Deviation		4.64	2.50	3.94
Minimum		31.00	40.00	34.00
Maximum		48.00	48.00	46.00

Objective Measures

Objective measures included physiological and flight technical performance measures collected during simulator training and live flight performances.

Physiological Measures. The research team collected real-time cognitive workload generated from electrocardiogram (ECG) signals through Cognitive Assessment Tool Set (CATS) software (OPL, 2014) using a minimally intrusive system with body-worn electrodes and the NeXus-4 biofeedback device that did not interfere with performance tasks. CATS transforms the ECG waveform from its scalar space to an embedded phase space, where it is then coarse grained to provide a quantitative signature of cognitive workload (Engler, Schnell, &

Walwanis, 2013; Schnell & Engler, 2014). When using CATS, the baseline cognitive workload measures are not required or necessary. Cognitive workload measures are relative to each individual, so the results are normalized on a 10-point scale. This scale aligns with the definitions of the Bedford Workload Scale (Roscoe & Ellis, 1990), where scores ranging from 1-3 represent pilot under saturation, 4-6 are optimal levels of engagement and 7-10 indicate oversaturation.

Procedure
Prior to arriving at OPL, participants were given access to an online folder containing the Flight Training Instruction P-1209 document, a PowerPoint file outlining the key areas of the P-1209, and two videos with examples of the patterns the EPs will be flying. Following the provided instructions, EPs completed all three surveys from home.

Orientation. An orientation visit was conducted at the lab on the day of simulation-based training. The purpose of this visit was to review and sign the Institutional Review Board Informed Consent Document forms, receive an introductory briefing, and complete a 20-question quiz over the briefing materials they had reviewed from home. Following the quiz, a PowerPoint briefing was reviewed with the participants, specifically focused on the weakest performance areas of the quiz. Prior to advancing to simulation-based training, participants were briefed on the procedures and plan for the day.

Simulator Training and Live Capstone Flight. Participants visited OPL two times. Simulator training was completed during the first visit, and live flight was completed during the second visit. Following the orientation, researchers attached ECG leads. The pilots wore a Nomex® flight suit to help contain the wire leads and prevent any safety or performance issues. Flight maneuvers were flown in the AIL simulator, followed by a live flight in the L-29 jet. Pilots began with the 15-degree Roll-In patterns, and advanced to the 30-degree Roll-In and Pop-Up patterns. An experienced safety pilot (SP) was present during live flight, but not during simulator training. The SP sat in the front seat of the L-29 rather than the rear seat because the experimental equipment was only present in the rear seat. The SP served as pilot-in-command (PIC) and was therefore responsible for ATC communication and all critical phases of flight (engine start, taxi, takeoff, initial climb, landing, engine shut down). Aircraft control was positively transferred to the EP during the climb. While in route to the first setup point of the experiment, the SP allowed the EP to take control of the aircraft for a few minutes so he could become accustomed with the performance and handling qualities of the L-29. At the end of each maneuver the SP took the controls and set up the jet for the next maneuver. During the live flight, each pilot completed each of the patterns twice, for a total of six. The three surveys were administered upon completion of simulator training and live flight.

RESULTS

Changes in confidence beliefs. To assess if the confidence beliefs significantly changed over time, a paired samples t-test was performed using three confidence measures, where the first confidence measure was collected prior to simulator training (T1), second was collected following simulator training and prior to live flight (T2), and the third was collected following the live flight (T3). Additionally, a repeated measures one-way analysis of variance (ANOVA) was applied to the three measures of confidence. First, a Mauchly's test was performed and showed that the assumption of sphericity was not violated ($p > .05$). There was a significant effect on confidence measures over time, Wilks' Lambda = .37, $F(2,7) = 5.85$, $p = .03$.

A paired samples t-test was used to make post hoc comparisons between the three confidence variables. Confidence means are displayed in Table 1. A first paired samples t-test indicated that there was a significant statistical difference between the confidence means measured at times T1 and T2 ($M = -4.0$, $SD = 3.62$, $p = 0.003$), and T1 and T3 ($M = -3.89$, $SD = 4.54$, $p = 0.033$), but not between the times T2 and T3 ($M = 0.78$, $SD = 4.06$, $p = 0.581$). Note that comparisons with T3 use means with the reduced sample size of 9 due to data collection limitations during the pandemic.

The results suggest that confidence means significantly varied over time, so that confidence beliefs increased following the simulator performance and decreased following live flight (Figure 4).

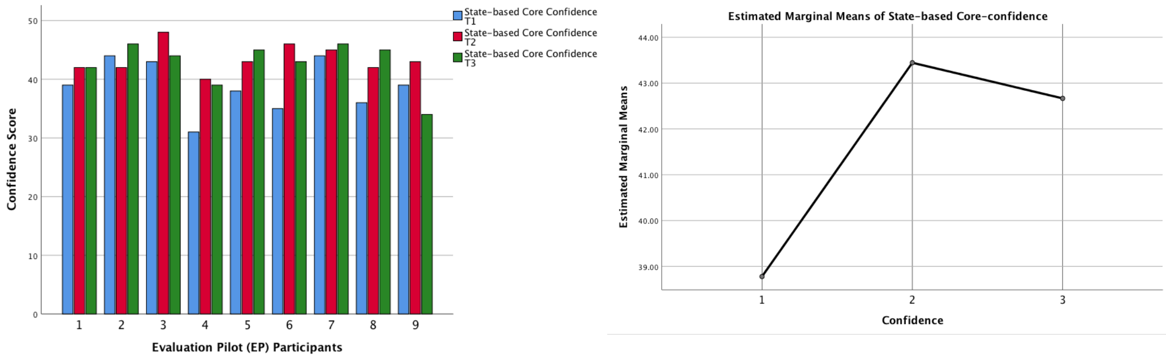


Figure 4. Changes in Confidence over time (T1, T2, and T3) for individual EPs (left) and marginal mean changes (right)

Confidence and performance outcomes. A hierarchical multiple linear regression test was used to determine if confidence beliefs predict subsequent performance outcomes. A calculated training grade was entered at stage one of the regression to control for prior performance outcomes, as some prior studies suggest that prior performance is a good predictor of future performance (Dideriksen et al., 2019). The second stage included confidence measures collected prior to simulator training (T1) and following the simulator training and prior to live flight (T2). The dependent variable was the live flight performance grade.

A hierarchical multiple linear regression test revealed that at the first stage, simulator training performance grades did not significantly contribute to the regression model $F(1,7) = 1.05$, $p = .339$. At the second stage, simulator training performance grade, $\beta = 1.0$, $p = .009$ and confidence measured prior to simulator training, $\beta = -1.116$, $p = .006$, served as statistically significant predictors of subsequent live flight performance outcomes, whereas confidence measured following the simulator training was not, $\beta = 0.28$, $p = .238$. The causality appears to be bidirectional in the relationships between confidence beliefs and performance outcomes. Figure 5 shows variations in confidence and performance in this study. Together, simulator training performance and confidence measured prior to training accounted for 83% of variability in the dependent variable of the subsequent live flight.

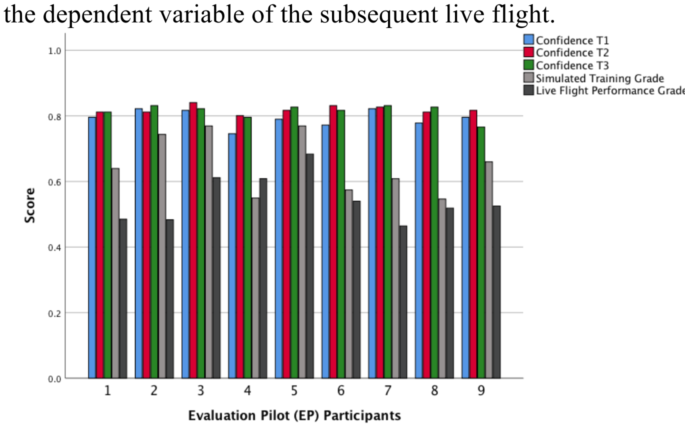


Figure 5. Confidence and Performance over time

Additionally, a repeated measures two-way analysis of variance (ANOVA) was applied to two measures of confidence (T1 and T2), and two measures of performance (simulator and live flight). Since the sphericity is always met for two levels of a repeated measure factor, the Mauchly's test was unnecessary. The results of the two-way repeated measures ANOVA revealed a significant main effect of both confidence, $F(1,8) = 11.79$, $p = .009$), and performance, $F(1,8) = 1675.02$, $p = .000$. Descriptive statistics revealed that confidence means significantly increased from time T1 ($M = 38.78$) to T2 ($M = 43.44$), and performance significantly decreased from simulated training ($M = .651$) to live flight ($M = .547$). There was also a significant interaction between confidence and performance such that performance was lowest following increase in confidence after the simulated training, Wilks' Lambda = .36, $F(1,8) = 14.01$, $p = .006$.

Confidence and cognitive workload. Given prior findings that confidence plays a role in self-regulation (Stajkovic, 2006) and accurate measures and classifications of cognitive workload play a key role in performance outcomes (Dideriksen et al., 2018; Hoke et al., 2017), two linear regression tests were used to assess if confidence predicted participants in an optimal range of cognitive workload. All participants' cognitive workload during simulator training were within the optimal range of 4-6 per the Bedford Workload Scale definitions, while only three participants' cognitive workload was in optimal range during live flight. This finding was consistent with prior studies that showed cognitive workload was typically higher in live flight compared to the simulator (Dideriksen et al., 2018; Hoke et al., 2017). A linear regression test showed that confidence measured at T1 served as a statistically significant predictor of participants in optimal cognitive workload range during simulator training performance, $\beta = -0.61$, $p = .035$. Another linear regression test showed that confidence measured following the simulator training (T2) did not serve as a significant predictor of participants in an optimal range of cognitive workload during the subsequent performance, $\beta = -0.328$, $p = .389$.

Additionally, a multiple linear regression model showed that confidence (T1) measure predicted participants in optimal cognitive workload during live flight ($t = -2.61$, $p = .048$), whereas subsequent confidence measure (T2) did not ($t = .32$, $p = .764$). Standardized beta of the confidence measure at T1 ($\beta = -.812$) indicated that confidence predicted the higher end of optimal cognitive workload range.

DISCUSSION AND CONCLUSION

The main goal of this study was to examine if a motivational construct of confidence should be a consideration for designing simulation-based training. We examined the following:

- Changes in state-based core confidence beliefs over time
- Predictive power and significance of the relationship between state-based core confidence and performance
- Relationship between state-based core confidence and optimal cognitive workload, given the important implications that cognitive workload has on learning and performance

Results show fluctuations in confidence such that confidence increased following simulation-based training, and both confidence and performance decreased in live flight. A statistically significant relationship between confidence and training effectiveness measures exists. This suggests that confidence could be used for warfighter pilot selection processes, orientation, and training purposes. The established correlation between state-based core confidence and performance and cognitive workload measures suggests that state-based core confidence may also be impacted by complex factors such as simulator level of fidelity.

Our findings show that the training environment contributed to a formation of overconfidence which may be due to the fidelity of the training environment inadequately replicating the task difficulty of the live environment. Overconfidence no longer had a significant relationship with performance and optimal cognitive workload, as opposed to an accurate confidence belief.

Changes in confidence beliefs. A repeated measures one-way ANOVA and dependent sample t-test both confirmed that confidence beliefs significantly varied between times T1, T2, and T3. Prior studies in the area of core confidence explored its trait-based properties, while this study specifically focused on exploration of its state-based properties under the assumption that confidence would change following training. An important finding of this study was that confidence significantly increased following simulator training, and significantly decreased following live flight. One explanation is that confidence measured prior to simulator training (T1) was a more accurate representation of confidence belief or a better 'calibrated' confidence belief to task performance. The subsequent confidence belief that followed simulator training (T2) was significantly higher. When associated with subsequent performance measures, the confidence beliefs at T2 appeared to be inaccurately calibrated to task performance, and inflated, which is consistent with overconfidence. Confidence beliefs following live flight (T3) significantly decreased, suggesting that the live flight 're-calibrated' confidence beliefs to task performance.

Confidence and performance. The second important finding of this study was the existence of a statistically significant bidirectional causal relationship between state-based confidence and performance. The results from a hierarchical multiple regression indicated that confidence beliefs measured prior to simulator training (T1) were predictive of subsequent performance outcomes, consistent with prior studies (Stajkovic, 2006; Stajkovic et al., 2015), whereas confidence beliefs measured following the simulator training (T2) represented overconfidence and were not predictive of subsequent performance outcomes. Simulation-based training environments can vary in levels of fidelity and realism. The L-29 simulator fidelity was estimated to be moderate. The results of this study showed that confidence significantly increased following simulated training, and significantly decreased following live flight. As well, the performance scores in the simulator were greater than the performance scores in live flight. Interestingly, high confidence scores that followed simulated training did not correlate with the subsequent performance scores in live flight. Although confidence following simulated training was not a significant predictor of live flight performance, we believe there is an explanation for this non-significant finding. One plausible explanation is that fidelity of the simulation-based training device using the L-29 with Coalescence contributed to overconfidence, and overconfidence appears not to be predictive of performance outcomes. Additional studies which manipulate fidelity should be conducted to ascertain the causality link between fidelity and overconfidence. It is likely that confidence measured at times T1 and T3 were better calibrated to task performance and as such, resulted in more accurate confidence beliefs.

Confidence and cognitive workload. Confidence has a cognitive dimension as explained through social-cognitive (Bandura, 1997) and self-regulation (Stajkovic, 2006) theories, which suggests that confidence and cognitive workload are related. A third important finding of this study was the existence of a statistically significant relationship between confidence beliefs measured at time T1 and optimal cognitive workload. Cognitive workload is a key component of learning (Sweller, 1988) and is representative of cognitive engagement. The relationship between confidence measured at time T1 and physiological cognitive workload measured during simulation training was statistically significant. This finding indicates the importance of an accurate measure of confidence. Confidence measured prior to simulator training was statistically significant in relation to the optimal cognitive workload, whereas confidence measured following simulator training was not. Although non-significant, we believe that a plausible explanation for this result is that confidence measured at time T2 represented overconfidence.

The Bedford Workload Scale (Roscoe & Ellis, 1990) definitions suggest the optimal cognitive workload range to be between 4 and 6 on a 10-point scale, but it does not indicate the optimal cognitive workload number within that range. Another important finding of this study is that accurately measured confidence is predictive of a higher end of the optimal cognitive workload range. As such, confidence could be used to help identify the optimal cognitive workload number within the optimal cognitive range.

Overall, the results of this study suggest that suboptimal or inadequate levels of fidelity in simulation-based training devices could contribute to the formation of overconfidence, which is not predictive of performance outcomes or optimal ranges of cognitive workload. Lower simulator fidelity may contribute to a creation of an inaccurate perception of one's competence, ultimately contributing to suboptimal training and lower subsequent performance outcomes. The findings in this study indicate that levels of immersion and realism could be important factors in the formation of accurate confidence beliefs, an important implication for designing the optimal level of fidelity of training devices. Confidence should be a consideration when designing appropriate levels of fidelity for simulation-based training devices, which could lead to an increase in training effectiveness, learning transfer, and ultimately, performance outcomes.

Instructional systems designers should carefully consider the implications of confidence on performance outcomes, and work closely with engineers when designing the fidelity requirements for simulation-based training devices. While this study offers an implication that simulator fidelity may contribute to a formation of overconfidence, it did not place research focus on understanding what the appropriate level of fidelity is for different training scenarios. The optimal level of fidelity may be different depending on the context, such as pilot competency and training content. While the level of fidelity is limited by financial constraints, instructional systems designers should collaborate with engineers to determine, design, and incorporate appropriate levels of fidelity based on the learning objectives for specific courses and temper levels of confidence formed based upon simulator fidelity. As well, confidence can inform pilot selection, training processes, and transfer of learning.

Limitations of the Study

Like most studies, this study had its limitations. Perhaps its greatest limitation is the small sample size that negatively impacts the statistical power. All pilots were commercially trained novices with no prior military experience, meaning that the two bombing scenarios they learned and performed contained a high degree of difficulty, which provided an opportunity to better measure the learning curve. While the results supported the theory, it is unclear whether the results would be validated should the sample consist of more experienced pilots, or pilots with military experience, or whether the results would generalize across other domains. Another limitation was access and ability to collect the same data using simulators of varying levels of fidelity that would enable a comparison and analysis of how these levels of fidelity affect the formation of confidence and provide evidence of causality.

Next Steps

This study confirmed an important role of a key psychological construct of state-based core confidence in pilot training effectiveness and performance, and its relationship with cognitive workload. Replicating the study using a larger and more diverse sample size and study design would help validate the results and increase the statistical power.

This study also emphasized the importance of the confidence belief as a significant predictor of performance and optimal cognitive workload, indicating that confidence is only useful when it is assessed accurately. Only a subjective scale for assessing state-based core confidence exists and was recently introduced (Sergent et al., 2020). It would be beneficial to examine additional ways in which confidence could be accurately measured, including the collection of objective measures. Appropriately calibrated and effectively built confidence could be used to enhance pilot training and performance outcomes for both individuals and teams (e.g., human-human and human-machine).

REFERENCES

Anderson, C., Brion, S., Moore, D. A., &Kennedy, J. A. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology, 103*(4), 718-735.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Bellows, P., Dideriksen, A., Williams, J., Schnell, T., Woodruff, K, Thompson, C., & Cover, M. (2020). A Quantitative assessment of flight training effectiveness in mixed reality. Manuscript submitted for publication.

Brennan, N. M., & Conroy, J. P. (2013). Executive hubris: The case of a bank CEO. *Accounting, Auditing & Accountability Journal*, *26*(2), 172-195.

CNATRA P-764 (04-19). Naval Air Training Command Flight Training Instruction.

Dideriksen, A., Reuter, C., Patry, T., Schnell, T., Hoke, J., & Faubert, J. (2018). *Define "expert": Characterizing proficiency for physiological measures of cognitive workload.* Paper presented at the Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Dierdorff, E. C., Surface, E. A., & Brown, K. G. (2010). Frame-of-reference training effectiveness: Effects of goal orientation and self-efficacy of affective, cognitive, skill-based, and transfer outcomes. *Journal of Applied Psychology, 95*(6), 1181-1191.

Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology, 69*(1), 58–68.

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology, 84*(1), 5-17.

Engler, J., Schnell, T., & Walwanis, M. (2013). Deterministically Nonlinear Dynamical Classification of Cognitive Workload. Proceedings of the Interservice/Industry Training, Simulation and Education Conference. Arlington, VA: National Training and Simulation Association.

Geil, M.A. and Greenwald, J. (2020), "An investigation of characteristics of effective global leaders", *Journal of Business & Industrial Marketing, Vol. ahead-of-print No. ahead-of-print.*

Gentile, B., Miller, J. D., Hoffman, B. J., Reidy, D. E., Zeichner, A., & Campbell, W. K. (2013). A test of two brief measures of grandiose narcissism: The Narcissistic Personality Inventory–13 and the Narcissistic Personality Inventory-16. *Psychological Assessment*, *25*(4), 1120-1136.

Gilstrap, R., & Greenwald, J.M. (2016). Creating an inimitable resource: An analysis of engagement, core confidence, and performance. Paper presented at Midwest Academy of Management, Fargo, N.D.

Hoke, J., Reuter, C., Romeas, T., Montariol, M., Schnell, T., & Faubert, J. (2017). *Perceptual-Cognitive & Physiological Assessment of Training Effectiveness.* Paper presented at the Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Holdorf, W.E., & Greenwald, J.M. (2018). Toward a taxonomy and unified construct of responsibility. *Personality and Individual Differences, 132*, 115–125.

Johnson, J. (2017). 3 ways to build confidence. *Print + Promo, 55*(4), 16-18.

Linderman-Hill, K. and Greenwald, J.M. (2019). The effect of trait core confidence on the relationship between job characteristics and engagement. Academy of Management Proceedings (No. 1, p. 12926). Briarcliff Manor, NY 10510: Academy of Management.

OPL. (2014). *Cognitive Assessment Tool Set (CATS) User Manual* Retrieved from Iowa City, Iowa:

Roscoe, A., & Ellis, G. (1990). A subjective rating scale assessing pilot workload in flight. A decade of practical use. *Royal Aerospace Establishment*. Farnborough, UK.

Schnell, T., & Engler, J. (2014). Entropic skill assessment of unmanned aerial systems (UAS) operators. Journal of Unmanned Vehicle Systems, 2(02), 53-68.

Sergent, K., Lee, D., Stajkovic, A. D., Greenwald, J. M., Younger, S., & Raffiee, J. (2020). Mitigating role of trait core confidence on psychological distress in entreprenurship. *Applied Psychology: An International Review, 0*(0), 1-26.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*(3), 421-442.

Stajkovic, A. D. (2006). Development of a core confidence-higher order construct. *Journal of Applied Psychology, 91*, 1208-1224.

Stajkovic, A. D., & Luthans, F. (2003). Social cognitive theory and self-efficacy: Implications for motivation theory and practice. In L. W. Porter, G. Bigley, & R. M. Steers (Eds.), *Motivation and work behavior* (pp. 126-140). New York: McGraw-Hill/Irwin.

Stajkovic, A. D., Lee, D. S., Greenwald, J. M., & Raffiee, J. (2015). The role of trait core confidence higher-order construct in self-regulation of performance and attitudes: Evidence from four studies. *Organizational Behavior and Human Decision Processes, 128,* 29-48.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12*(2), 257-285.