Towards Predicting Completion for United States Air Force (USAF)

Remotely Piloted Aircraft (RPA) Training

by

Jade B. Driggs

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved February 2017 by the
Graduate Supervisory Committee:

Nancy Cooke, Chair
Mary Niemczyk
Rod Roscoe

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

Civilian and military use of remotely piloted aircraft (RPA) has significantly increased in recent years. Specifically, the United States Air Force (USAF) has an insatiable demand for RPA operations, that are responsible for fulfilling critical demands in every theater 24 hours a day, 365 days a year (United States Air Force, 2015). Around the clock operations have led to a manning shortage of RPA pilots in the USAF. The USAF MQ-9 "Reaper" Weapons School trains tactical experts and leaders of Airmen skilled in the art of integrated battle-space dominance (United States Air Force, 2015). Weapons Officers for the MQ-9 platform are also critically under-manned, with only 17% of allocated slots filled (B. Callahan, personal communication, January 28, 2016). Furthermore, the leading cause of training attrition has been attributed to lack of critical thinking and problem solving skills (B. Callahan, personal communication, January 28, 2016); skills not directly screened for prior to entering the RPA pilot career field. The proposed study seeks to discover patterns of student behaviors in the brief and debrief process in Weapons School, with the goal of identifying the competencies that distinguish the top students in Weapons School.

ACKNOWLEDGMENTS

I would like to recognize my academic advisor, Nancy Cooke, and my thesis committee, Rod Roscoe, and Mary Niemczyk for their patience, guidance, and expertise. I could not have completed my research in the compressed 18-month timeline without your help. I would also like to recognize Nathan McNeese for the time he spent working with me to refine my literature review. To the faculty of the Human Systems Engineering Department and my classmates, thank you for expanding my knowledge of human systems engineering. I have learned a great deal from each of one of you. Thank you for the enlightening, thought-provoking discussions, and for forcing me to think outside of my own domain-specific experiences.

To the United States Air Force (USAF) Weapons School, specifically the 26 WPS students and instructor pilots – thank you for taking the time to explain the mission and purpose of Weapons School to a non-flyer. I not only have a renewed pride in serving this great country, but I also have a profound respect and admiration for the rigorous process of becoming a USAF Weapons Officer. I can sleep soundly at night knowing that every six months, the Weapons School graduates a new set of Weapons Officer prepared to fight the wars of tomorrow. I would like to especially recognize Lt Col Brian Gyovai and Capt Nicholas Meyers for the extra time they took responding to my numerous questions and for facilitating my researcher visits – thank you for your time and patience.

Finally, to my husband Capt Christopher Driggs, for some unknown reason, you have continued to listen to countless rants on this extensive process and reviewed thesis drafts for me. Thank you for your constant support in all my endeavors.

TABLE OF CONTENTS

Page

APPENDIX

LIST OF TABLES

Table Page

LIST OF FIGURES

INTRODUCTION AND REAL-WORLD PROBLEM

The United States Air Force (USAF) is planning on sharply increasing the number of Remotely Piloted Aircraft (RPA) flights over the next four years. The vision is to carry out as many as 90 daily missions by 2019 (Walker, 2015); a significant increase from the 60 daily missions presently conducted (Schogol, 2015). The push to increase the number of RPA missions is met with a critical manning shortage of USAF RPA pilots. The Air Force has more than quadrupled the number of RPA pilots that it requires, from about 400 in 2008 to close to 1,650 in 2016 (United States Government Accounting Office, 2017). The high workload demands placed on RPA pilots has led to challenges in recruiting new pilots and difficulties in retaining existing RPA pilots, with the USAF offering retention bonuses of $125,000 for additional five-year service commitments (Losey, 2015). The USAF estimates that RPA pilots fly their aircraft more than any other pilots in the Air Force; fighter pilots fly an average of 200 hours annually, cargo and tanker pilots fly an average of 500 hours annually, while RPA pilots fly an average of 900 hours annually (United States Government Accounting Office, 2017).

RPA crews in the USAF MQ-9 Reaper and the MQ-1 Predator consist of two individuals – a pilot and sensor operator. Despite only having crews of two individuals, it can take as many as 170 persons to launch, fly, and maintain RPAs, in addition to processing and disseminating intelligence, surveillance and reconnaissance information (Tirpak & Deptula, 2010). RPA operations are unique in that the aircrew and aircraft are not co-located (Tvaryanas, 2006). They are further delineated from their manned airframe counterparts because of the use of remote split operations. In remote split operations, the deployed crew, in environments such as Iraq and Afghanistan, are responsible for setting the RPA up, calibrating their systems and taking off and landing. This crew is referred to as the launch and recovery element, with a pilot and sensor operator in the deployed environment responsible for the take-off and landing phases of flight. After the RPA is airborne, however, a 2-man pilot and sensor operator team located stateside is responsible for flying the actual mission (Tirpak, et al., 2010). The hand-off portion of RPA flight is the most checklist intensive, with 140 items on the gaining crew checklist (B. Callahan, personal communication, January 28, 2016). A brief summary of some RPA Pilot duties, specific to the MQ-1 Predator and the MQ-9 Reaper is provided in the table below:

Table 1

MQ-1 Predator and MQ-9 Reaper Pilot Duties

1. Performing preflight and in-flight mission planning activities in accordance with unified combatant command and theater rules of engagement
2. Understanding tactics, techniques, and procedures for friendly and enemy air order of battle (AOB) assets
3. Ensuring airframe and supporting GCS systems for controlling the aircraft are operating efficiently and effectively
4. Performing checklists and monitoring systems controls during aircraft launch and recovery operations
5. Flying the aircraft en route to airspace of national interest while coordinating with air traffic control, as well as other aircraft and aircrew
6. Maneuvering the aircraft to gather surveillance and reconnaissance data over targets and areas of interest
7. Maneuvering the aircraft into strategic positions for the deployment of weapons (e.g., close air support of ground troops)
8. Receiving target briefs for weapons delivery and conducting battle damage assessments (BDAs)
9. Maintaining situational awareness to target imagery, friendly and enemy orders of battle, and offensive and defensive capabilities from various sources
10. Assembling target information, locating forces, and determining hostile intentions and possible tactics

Table reproduced from Chappelle, McDonald, and McMillan (2011)

The USAF MQ-9 Reaper Weapons School at Nellis Air Force Base (AFB) in Nevada, trains tactical experts and leaders of Airmen skilled in the art of integrated battle-space dominance across the land, air, space, and cyber domains (United States Air Force, 2015). Attendance at the Weapons School is a highly-competitive process, with only the top performing RPA pilots in units across the United States selected to attend. Even at this elite level of training, attrition levels are high. Lack of problem solving and critical thinking skills are thought to be the leading causes of attrition from Weapons School, not flying deficiencies (B. Callahan, personal communication, January 28, 2016). According to instructors, critical thinking skills, in particular, are most important in the debrief portion of the standard brief-fly-debrief cycles (B. Gyovai, personal communication, February 24, 2016). During the debrief portion, students are required to reconstruct the events of the mission just completed, and are cautioned not to make any assumptions. The debrief starts off as a fact gathering session before any concrete conclusions are drawn. Further, students need to prove that error(s) occurred and are expected to quantify the error(s) (B. Gyovai, personal communication, February 24, 2016).

The Chief of Staff of the Air Force, General Mark Welsh, has cited the USAF's "ability to continue to adapt and respond faster than our potential adversaries is the greatest challenge we face over the next 30 years" (Airman Magazine, 2014). To this end, methods and criteria used to select students for USAF MQ-9 Reaper Weapons School require further research and support by empirical evidence.

GAPS IN RPA PERSONNEL SELECTION

The literature points to gaps in formal assessment of critical thinking and problem solving skills in RPA pilots. From a real-world training perspective, attrition rate in the USAF MQ-9 program is attributed to deficiencies in critical thinking and problem solving skills. Howse (2011) cites the need for problem solving skills to improve in future years, which stands to reason given the insatiable demand for RPA operations in both military and commercial settings. Furthermore, the RPA personnel literature has pointed to gaps in critical thinking and problem solving skills, yet believes that there would be little added value in explicitly screening for these aptitudes (Paullin, et al., 2011; Williams, et al., 2011). Absence of critical thinking and problem solving skills as a leading reason for MQ-9 Weapons School attrition (B. Callahan, personal communication, January 28, 2016) provides a strong argument towards screening for these skills up front in ab initio, or pre-training, RPA pilots.

An additional gap in the literature is in identifying measures that predict long-term RPA outcomes in environments which require RPA flight (Barron, et al., 2016). Current predictor measures focus on predicting initial RPA training outcomes, specifically completion of Undergraduate RPA Training (URT). Validation of aptitude and traits predictive of RPA pilot success thus far has been limited to RPA pilot training outcomes that require manned flight (Barron, et al., 2016). Because of this limitation, there has been no validation of aptitude and traits predictive of performance in advanced RPA operator courses, such as MQ-9 Weapons School. Although there is value in predicting initial training outcomes, the Air Force is having difficulty not only recruiting

RPA pilots but more importantly, retaining RPA pilots. With retention in mind, screening for long-term RPA outcomes could prove to serve as a predictor for individuals with Air Force long-term career characteristics, such as the propensity to complete the prestigious Weapons School program.

The importance of screening for long-term outcomes has been examined in the medical domain, specifically in nursing. Wong and Cummings (2007) examined the relationship between nursing leadership and patient outcomes. Evidence of significant associations between positive leadership behaviors and increased patient satisfaction and reduced adverse events were found. Further, Davis, Flett, and Besser (2002), examined using a measure of problematic internet use, the Online Cognition Scale (OCS), combined with measures of diminished impulse control, loneliness/depression, social comfort, and distraction, for pre-employment screening. As hypothesized, the OCS predicted being reprimanded at school or work for inappropriate internet use.

The goal of the proposed study is to determine the competencies that distinguish successful from unsuccessful Weapons School students. In this context, competency is defined as “an observable, measurable pattern of knowledge, abilities, skills, and other characteristics that individuals need to perform work roles or occupational functions successfully” (United States Government Accounting Office, 2017). Are critical thinking skills the distinguishing factor in student performance? Another goal of the study is to further tailor critical thinking to the Weapons School context, as well as determine representative competencies and behaviors. Because Weapons School has a distinguished graduate, who is recognized as the top all-around student in the class, it is hypothesized that the competencies that the distinguished graduate exhibits differ from those of their classmates. It is these competencies that we seek to quantify. The table below outlines the problem motivation, research gaps, research objective and research question. The research gaps will be expanded upon in the literature review.

Table 2

Alignment of Problem Statement/Research Gaps/Objective/Questions

Problem Motivation		
Inadequate critical thinking and problem solving skills cited by instructors as the leading causes of attrition from USAF MQ-9 Weapons School (B. Callahan, personal communication, January 28, 2016).		
Research Gaps	**Research Objective**	**Research Questions**
1. Gap in measure of critical thinking/problem solving skills in USAF RPA pilots (Paullin, et al., 2011; Williams, et al., 2014). 2. Gap in validation of screening measure for broader job performance criteria; most research has focused on identifying measures that predict initial RPA training outcomes for *ab initio* (pre-training) pilots (Carretta, 2013; Rose, Barron, Carretta, Arnold & Howse, 2014; Barron, Carretta, & Rose, 2016).	1. Develop a profile outlining the competencies, in an academic sense, of a successful Weapons School student. 2. Determine if critical thinking/problem solving skills is truly the reason behind poor Weapons School performance and ultimately, attrition.	1. Are deficiencies in critical thinking/problem solving skills the reason behind unsatisfactory Weapons School performance? 2. If so, what would be a suitable critical thinking/problem solving assessment for the USAF to use in screening potential candidates? If not, what would be a suitable additional screening tool based on the identified distinguishing factors between successful and unsuccessful students?

LITERATURE REVIEW

Historically, research efforts in identifying and screening individual differences that could separate high and low flying aptitudes has dated back to World War I and the initial development of apparatus-based pilot selection tests (Damos, 2011). According to Carretta (2013), the military has since refined pilot selection techniques by measuring aviation job knowledge/experience, cognitive, and psychomotor abilities in pilot candidates by using the Air Force Officer Qualifying Test (AFOQT), Test of Basic Aviation Skills (TBAS), and the Pilot Candidate Selection Method (PCSM), respectively. In recent years, RPA research has focused on identifying requisite knowledge, skills, abilities, and other characteristics (KSAOs) of potential RPA pilots and sensor operators. The RPA literature as a whole, has arrived at a consensus regarding the desirable knowledge, skills, abilities, and other characteristics (such as personality traits), for operators and has transitioned efforts towards matching KSAOs to measures in existing Air Force owned assessments. For KSAOs that cannot be matched to existing measures, research efforts have focused on developing new measures to tap into the required KSAOs.

RPA Operator KSAOs

Pavlas, et al., (2009) developed a general taxonomy of desirable knowledge, skills, and attitudes (KSA) for Unmanned Aerial System (UAS) operators. There was no overarching baseline construct which they referenced in compiling the KSAs. The authors gathered the KSAs by searching both psychology and military research databases with the terms “UAS teams,” “unmanned aerial vehicle teams,” “UAS training,” and related terms. Further, they provided methods to develop these KSAs, with the ultimate goal of combining practice with science (see APPENDIX A). *Knowledge*, is defined as the underlying memory structures used to recognize and utilize environmental information, *skill,* as what RPA operators need to have in order to complete necessary tasks, and *attitudes*, as the affective states and differences of team members. Problem solving is listed as a skill for RPA operators, needed to complete often erratic and fluctuating missions.

Howse (2011), reviewed more than 200 publications with the goal of identifying KSAOs for the purpose of RPA system design specification or for RPA personnel selection. As a baseline reference and operational definition tool for comparing KSAOs, Howse utilized Fleishman's Taxonomy of Human Abilities. In total, eight publications contained relevant lists of KSAOs for RPA pilot and sensor operator positions. Problem solving was identified in three of the eight separate KSAOs lists. Additionally, in an extrapolation of present KSAOs identified, problem solving was cited as a cognitive ability expected to increase in need in future RPA operations; given the proposed growing complexity of RPA missions. Howse concluded by suggesting that services consider conducting studies of training failure rates to determine if the costs and development and fielding selection instruments are justifiable.

Mapping of RPA operator KSAOs to existing USAF aptitude test batteries. Paullin, Ingerick, Trippe, and Wasko (2011) sought to first select "best bet" predictor measures to assist the USAF in identifying early career airmen likely to succeed as RPA pilots and sensor operators. They first compiled a comprehensive list of KSAOs, utilizing the Department of Labor's Occupational Information Network (O*NET) taxonomy as an operational definition tool; this will be discussed in further detail in subsequent paragraphs (Peterson, Mumford, Borman, Jeanneret, & Fleishman, 1999; see APPENDIX B). After compiling KSAOs, Paullin and colleagues (2011) selected predictor measures associated with the KSAOs and identified existing test batteries (inclusive of those owned by the USAF and those accessible to the USAF) which measured the critical KSAOs identified. Overall, Paullin, et al., (2011), recommended two possible batteries of predictor measures (one for pilots and one for sensor operators), as well as a combined test battery to screen for either position. They also addressed measurement gaps by first developing a new measure of time-sharing ability that involves performing multiple tasks tapping working memory, task prioritization, and selective attention (Paullin et al., 2011). The second measure is an RPA-specific person-environment fit measure, to help potential recruits determine if the RPA work context would be a good fit for their work preferences. The intent of this measure was to serve as a self-assessment tool prior to enlistment or accessioning (Paullin, et al., 2011). Critical thinking skills were identified as a critical KSAO, with no Air Force-owned predictor presently in place to measure it. They did not suggest including a measure of critical thinking skills in current test batteries out of the concern that it might not provide enough incremental validity beyond measures of fundamental cognitive abilities to be worth the extra screening time.

In addition to the work accomplished by Paullin, et al., (2011), Williams, et al., (2014) conducted a joint Air Force, Navy and Army review of skills, abilities and other characteristics (SAOC) needed for successful RPA pilot performance. The authors recommended an updated test battery to assess the RPA SAOCs and further concurred with the gap in critical thinking and problem solving measures in existing Department of Defense (DOD) RPA operator test batteries. Critical thinking was operationally defined in the study as the "ability to analyze the strengths and weaknesses of specific actions or decisions" (Williams, et al., 2014). Further, critical thinking was ranked the 6^{th} most important KSAO (by subject matter experts) not currently measured for in existing DOD proprietary tests. However, much like Paullin, et al., 2011, the authors foresaw little benefit from the addition of an explicit critical thinking measure.

As evidenced in Paullin, et al., (2011) and Williams, et al., (2014), critical thinking skills are a skill necessary for RPA operators. However, due to time constraints in testing, as well as a belief that explicitly screening for critical thinking skills would not provide enough incremental validity beyond current measures, there has not been a push to screen for these skills upfront in RPA operators **(Research Gap 1).** Additionally, Paullin, et al., (2011) focused on selecting "best-bet" predictor measures that could be used to identify entry-level or early career officers and airmen for RPA pilot or sensor operator positions. Whereas screening for entry-level positions is important, there is a gap in screening for subsequent training outcomes, such as advanced courses (e.g., Weapons School) **(Research Gap 2).** Williams, et al., (2014) also focused on identifying entry-level SAOCs required for RPA operators, rather than considering long-term outcomes **(Research Gap 2).**

Psychological and Personality Measures as Predictors for RPA Pilots

The RPA work environment operates in stark contrast to traditional, manned cockpits. Void of any haptic feedback or environmental cues, RPA pilots and sensor operators fly their air vehicles from ground control stations for shifts averaging anywhere between eight to twelve hours at a time (Chappelle, McDonald, & McMillan, 2011). This unique cockpit environment requires a special personality type in order to thrive; a person could have the psychomotor and cognitive abilities necessary to complete the unique RPA mission, but not the right personality or psychological attributes necessary to excel in the work environment. Cognitive aptitudes do not account for all of the factors associated with training and operational success, and Chappelle, et al., (2011) believe that this gap opens the possibility that other factors such as personality traits and motivation can provide additional insight into RPA pilot success. Chappelle, et al., (2011) set out to first identify important psychological attributes of USAF RPA pilots according to inputs from 82 RPA subject matter experts. Attributes are distinguished from knowledge and skills as, "the inherent aptitudes, traits, and motivation that must be present to acquire the level of knowledge and skills needed to successfully operate as a pilot and adapt to the unique demands of the RPA platform" (Chappelle, et al., 2011). Their analysis identified these important psychological attributes in four different domains: cognitive, intrapersonal, interpersonal, and motivation. Within the reasoning facet of the cognitive ability domain, problem solving was identified as a critical psychological attribute required of RPA pilots for success. The profile of critical psychological attributes for USAF RPA pilot is summarized in Table 3 below:

Table 3

Critical Psychological Attributes for RPA pilots

Cognitive Ability	**Intrapersonal personality traits**	**Interpersonal personality traits**	**Motivation**
Speed of information processing and accuracy; visual-perceptual recognition, tracking, and analysis; sustained and divided attention to visual/auditory information, ***problem solving***	Emotional composure, resilience, self-certainty, conscientiousness, perseverance, success-orientation, decisiveness, and adaptability	Humility, comfort and confidence in working in groups, social cautiousness and prudence, and team orientation	Moral and occupational interest in saving lives and sense of duty as a military officer

Table reproduced from Chappelle, et al., (2011)

Rose, et al. (2014), investigated the extent personality measures (self-description inventory, SDI+) could improve prediction of remotely piloted aircraft (RPA) training outcomes. Participants were 280 USAF officers, required to take the AFOQT and TBAS as part of their commissioning requirements. The participant pool was further broken up into 170 participants with RPA Initial Flight Screening (IFS) outcomes and 110 participants with RPA Instrument Qualification (RIQ) outcomes. Both courses focus on skills historically required for manned pilots, and which have since been identified as essential for RPA pilots. The SDI+ is currently administered as an experimental measure in the AFOQT, and is being evaluated for operational use. The Big Five Personality traits, per the SDI+ are agreeableness, neuroticism, extraversion, conscientiousness, and openness, plus a sixth factor resembling machiavellianism (Manley, 2011). Results demonstrated significant negative relationships between the Big Five personality trait of Openness and several RPA training outcomes. Traits commonly associated with openness include being reflective, introspective and curious. In general, these are not undesirable traits to have, however, in RPA missions, marked by periods of high workload levels followed by lulls in activity, pilots who tend to become introspective during downtime could have difficulties re-engaging in missions when required. Overall, the study supported incremental validity for personality in predicting RPA pilot training outcomes (Rose, et al., 2014). Findings from this study are useful in further refining selection tests. Although the Air Force has historically treated unmanned pilots like manned pilots, there are undeniable differences in the work environments, these differences need to be accounted for when assigning pilots to RPAs.

Chappelle, Swearingen, Goodman, and Thompson (2014) further investigated personality test scores in Remotely Piloted Aircraft pilot training candidates. Participants entered manned or unmanned pilot training between 2009 and 2013. Primarily, the study evaluated the differences in personality test scores for three distinct groups of pilot training candidates, summarized in the table below:

Table 4

Pilot Training Candidate Categories

Group 1	Pilot candidates who volunteered to fly RPA upon commissioning in the USAF
Group 2	Pilot candidates who had completed manned undergraduate pilot training, but were forced to fly RPA due to personnel gaps
Group 3	Pilot candidates who had completed manned undergraduate pilot training and were assigned manned airframes

Table reproduced from Chappelle, Swearengen, Goodman, and Thompson (2014)

Findings suggested that as a group, the RPA training candidates who had volunteered for the RPA career field (Group 1), tend to be more methodical, cautious, harbor a “team-player” mentality, as well as have a more defined value and belief system. Moreover, Group 3 individuals had a higher level of frustration tolerance and less of a need for excitement and stimulation when compared to Group 2. (Chappelle, et al., 2014). The findings from this personality assessment suggest significant, and potentially problematic person-job fit issues for those individuals assigned RPAs after being trained to fly manned aircraft (group 2). In light of these findings, it stands to reason that the person-job fit issue is one the USAF needs to keep at the forefront of personnel assignment (manned vs. unmanned) in order to mitigate retention issues for RPA pilots.

The personality literature can be directly tied back to the gaps identified within the RPA selection literature. Chappelle, et al., (2011), identified psychological attributes for RPA operators, with problem solving listed as a cognitive ability **(Research Gap 1).** Further, Rose, et al., (2014), investigated the extent to which personality measures could improve prediction of RPA training outcomes, with the focus on initial RPA training outcomes, and not outcomes in subsequent courses (**Research Gap 2).** The findings of

Chappelle, et al., (2014) served as a step towards screening for broader job criteria, which is not currently done **(Research Gap 2).** Group 2 in the study – pilot candidates who completed manned undergraduate pilot training, but were forced to fly RPA due to personnel gaps, had a lower level of frustration tolerance and a higher need for excitement and stimulation than individuals in Group 1 – pilot candidates who volunteered to pursue the RPA pilot career field from the beginning. This study could be potentially viewed as the foundations of a long-term outcome screening tool; individuals who came into the Air Force with the desire to fly manned aircraft and who have completed manned training could be perceived as a poor fit for the RPA flight environment, and less likely to remain in the Air Force after the completion of their active duty service commitments.

Validation of Existing Test Measures

The USAF has experience in the development and validation of selection methods for other aircrew occupations such as pilots, combat system operators and air battle managers. Additionally, the USAF has validated manned pilot selection methods for the unmanned pilot career field. Carretta (2013) validated manned pilot selection instruments, which are also used to screen for unmanned pilot candidates for RPA pilot training outcomes. Presently, the Undergraduate RPA Training (URT) course is structured similarly to the Specialized Undergraduate Pilot Training (SUPT) pipeline for manned aircraft pilots. The Air Force Officer Qualification Test (AFOQT) and the Pilot Candidate Selection Method (PCSM) are both screening tests used for both manned and unmanned pilot training programs. In his 2013 study, Carretta determined that the

AFOQT and PCSM demonstrated *moderate* predictive validity for URT completion, with r-values of 0.378 and 0.480, respectively. Although the AFOQT and PCSM demonstrated moderate predictive validity, he highlighted the fact that the Air Force is still continuing to examine the utility of other possible measures as a supplement to current methods.

Critical Thinking

Critical thinking is a skill in demand for nearly all professions, be it academia, industry or the military. As Sternberg, Roediger and Halpern (2007) assert – we all want a workforce and a citizenry that can do more of it. Critical thinking skills are even more desirable in complex RPA mission environments, and is cited as one of the leading causes of student attrition from USAF MQ-9 Weapons School (B. Callahan, personal communication, January 28, 2016). The term critical thinking refers to the use of cognitive skills or strategies that increase the probability of a desirable outcome (Halpern 1998). Furthermore, it is defined as thinking that is "purposeful, reasoned, and goal-directed," and it is the type of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions (Halpern, 1998). This definition aligns with the Department of Labor's O*NET framework definition of critical thinking skills, describing individuals as "skilled at using logic and reasoning to identify the strength and weaknesses of alternative solutions, conclusions, or approaches to problems" (Peterson, et al., 1999).

Critically thinking individuals tend to evaluate their thinking process and examining the reasoning that led to the final conclusion. In other words, critical thinking involves high levels of metacognitive monitoring. Metacognition is informally defined as “what we know about what we know” (Halpern, 1998). Moreover, metacognition is the ability to take this knowledge to guide and improve the thinking and learning process. Students who correctly engage in critical thinking will consciously monitor their thinking process, check whether progress has been made towards their goal, ensure accuracy, and make informed decisions about the use of their time and mental effort (Halpern, 1998).

These characteristics of critically thinking students directly align with MQ-9 Weapons School instructor expectations. During the debrief portion after missions, instructors look for the ability of students to self-assess their performance. Specifically, instructors are assessing whether students can accurately reconstruct their missions, and can identify points of failure in their plan and subsequent execution of the plan. They are assessing students on their ability to identify all of the crucial points in their missions before the instructor highlights their mistakes and flaws in the planning and execution stages (B. Callahan, personal communication, January 28, 2016).

Halpern Critical Thinking Assessment. The USAF presently does not directly screen for critical thinking abilities in officer candidates **(Research Gap 1).** The USAF RPA research community has identified critical thinking skills as necessary for RPA pilots to be successful, but share a general concern that measures of critical thinking skills might not provide enough incremental validity beyond measures of fundamental cognitive abilities to be worth the extra testing time (Paullin, et al., 2011; Williams, et al., 2014). The USAF Weapons School does not screen for critical thinking ability in its candidates. Rather, students are selected based on total flight hours, total hours as an instructor pilot, flight history, officer performance reports from the last five years, and public speaking abilities (B. Callahan, personal communication, January 28, 2016). With lack of critical thinking skills cited as the leading cause of attrition in MQ-9 Weapons School, there is a possibility of a relationship between critical thinking skills and training performance. Further, there is a possibility that this relationship is further broken down between the Distinguished Graduate and the rest of the class. It is important to note, however, that lack of critical thinking skills has yet to be scientifically determined to be the leading cause of attrition in MQ-9 Weapons School.

The Halpern Critical Thinking Assessment (HCTA) is a reliable measure of critical thinking skills and has been validated with multiple populations and measures of academic success. Furthermore, the HCTA has been shown to be the first test of critical thinking that actually predicts what people (say they) do in real life (Butler, 2012). Moreover, a subset of the HCTA directly measures problem solving skills, identified as one of the leading causes of MQ-9 Weapons School attrition. Critical thinking is a multidimensional construct and the assessment of critical thinking follows this idea. The HCTA assesses five different dimensions of critical thinking: verbal reasoning, argument analysis, thinking as hypothesis testing, likelihood and uncertainty, and decision making and problem solving (Halpern, 1998). The five dimensions of critical thinking are further described in the table below:

Table 5

Five dimensions of critical thinking

Verbal reasoning skills	Skills needed to comprehend and defend against the persuasive techniques embedded in everyday language
Argument analysis skills	An argument is a set of statements with at least one conclusion and one reason that supports the conclusion. In real-life settings, arguments are complex, with reasons that run counter to the conclusion, stated and unstated assumptions, irrelevant information, and intermediate steps
Skills in thinking as hypothesis testing	People function like scientists to explain, predict, and control events. Skills include generalizability, recognition of the need for an adequately large sample size, accurate assessment, and validity
Likelihood and uncertainty	Since very few events in life can be known with certainty, the correct use of cumulative, exclusive, and contingent probabilities should play a critical role in almost every decision
Decision-making and problem solving skills	Generating and selecting alternatives and judging among them

Note. Adapted from "Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring," by D. F Halpern, 1998, *American Psychologist*, *53(4)*, p. 449.

Transfer

Preliminary discussions with Weapons School instructors, prior to conducting the study, revealed potential student issues with transfer. Transfer can be broadly construed as, "the ability of individuals to 'treat' a new concept, problem or phenomenon as similar to one(s) they have experienced before" (Chi & VanLehn, 2012). Research on transfer asks how people strike the balance between reusing previous learning to treat situations like old ones, while also avoiding the tendency to overgeneralize prior learning and miss what is new (Schwartz, Chase, & Bransford, 2012). Many approaches to instruction focus on helping the student realize the "old in the new," with the end goal of developing familiar patterns for students; patterns which would facilitate the reuse of prior learning

(Schwartz, et al., 2012). This aligns with one of the desired outcomes of USAF Weapons School, specifically the expectation that students refrain from “brain dumping” information, and applying more of a building block approach (B. Gyovai, personal communication, May 11, 2016).

Summary

The reviewed literature builds a comprehensive profile of a successful RPA pilot, by first identifying desirable knowledge, skills, abilities, and other characteristics and matching these to validated predictors of success in the RPA profession. Specifically, the literature highlights existing measures within the USAF proprietary test battery repository which tap into necessary knowledge, skills, abilities and other characteristics of RPA pilots. The review reveals gaps in selection and screening procedures, as well as recommended measures. The personality and psychological factors literature is important to acknowledge because certain traits of individuals could serve as moderating variables to desired cognitive abilities previously identified. Gaps in measures of critical thinking skills have continually been identified throughout the literature, with no suggestion to explicitly screen for this important ability. Validation of existing measures is important, but given the attrition rate of RPA pilots from advanced courses such as the MQ-9 Weapons School, as well as difficulties with long-term retention, future research should focus on finding measures that: (1) can potentially provide incremental validity beyond measures presently in place, such as a critical thinking measure; (2) predict future on-the-job outcomes (such as performance in advanced courses), rather than just initial training outcomes.

Despite the discussion on critical thinking in the literature and the concern expressed by instructors that Weapons School attrition is because of student inability to critical think, it is important to remember that this has yet to be validated scientifically. Based on the literature review, there are obvious gaps in RPA measures that may provide incremental validity beyond measures currently in place, and that may predict future on-the-job outcomes - more specifically Weapons School completion. These gaps have led to problems in selecting students who will go on to successfully complete advanced RPA training courses. The first step in solving this problem is determining, in an academic sense, the competencies which distinguish successful and unsuccessful students in advanced training courses, such as the USAF Weapons School. Are critical thinking skills the distinguishing factor in performance? This will first require expanding the concept of critical thinking, from the Weapons School perspective.

Critical thinking is a term used in many contexts. In a recent *Forbes Magazine* article, a staggering 60% of managers claimed the new graduates they see taking jobs within their organizations do not have the critical thinking and problem solving skills they feel are necessary for the job (Strauss, 2016). Although this mirrors the general concern of the USAF Weapons School, critical thinking is not universally defined across these differing contexts. The goal of the research was to expand upon the idea of critical thinking within USAF Weapons School. In this context, what does critical thinking mean? What are the desired student competencies and student brief and debrief behaviors? What are the “soft” and “hard” skills required to be successful in USAF

Weapons School? Hard skills are more technical, whereas soft skills tend to be more intrapersonal and interpersonal oriented (Laker & Powell, 2011).

METHODS

The methods section is divided into the standard participants, task and materials, measures, procedure, analysis, and discussion sections. However, since the study drew on three primary sources of information: Halpern Critical Thinking Assessment scores, instructor interviews, and student grade sheets, each of these study portions will have separate task and materials, measures, procedure, and analysis sections.

Participants

For all three sections of the study, participants were six active duty United States Air Force RPA pilots currently enrolled in the MQ-9 Weapons School Course at Nellis Air Force Base, NV and seven instructor pilots (IPs) at the MQ-9 Weapons School. Students were competitively selected for Weapons School billets, evaluated on their total flight hours, total hours as an instructor pilot, flight history, officer performance reports (last five years), and public speaking abilities (B. Callahan, personal communication, January 28, 2016). Students graduating from the MQ-9 Weapons School Program serve as trusted advisors to military leaders at all levels and furthermore are the instructors of the Air Force's instructors and the service's "institutional reservoir" of tactical and operational knowledge (United States Air Force, 2015). Demographic data for both the instructors and students were collected and are summarized in the below tables:

Table 6

MQ-9 Weapons School Instructor Demographics

Instructor Number	Aircraft Flown prior to MQ-9	UPT Base	WIC Duration as IP
1	T-6, T-38, B-1	Traditional	9 months
2	T-6, T-38, MQ-1	ENJJPT	4 months
3	T-6, T-38, F-15, F-16	Traditional	24 months
4	T-6, T-38, MQ-1	ENJJPT	24 months
5	T-6, T-38, MQ-9	Traditional	4 months
6	T-37, T-38, MQ-1	ENJJPT	16 months
7	T-6, T-1, MQ-1, MQ-9	Traditional	4 months

Table 7

MQ-9 Weapons School Class 16A Demographics

Age (years)	30.8
Marital Status	83% or 5/6 students married*
Children	1.33**
Time spent on Weapons School material (hours/week)	70.33
Flight Background	2/6 Core 18X, 1/6 Navigator, 3/6 Undergraduate Pilot Training (UPT)***

*of the 5 students that were married, 3 of them were stationed in Nevada. The other two traveled to Nevada to complete Weapons School without their families present.

**of the students that had children, only 1 of them was stationed in Nevada. The other two traveled to complete Weapons School without their children present.

**Undergraduate Pilot Training (UPT) completion vs. Undergraduate Remotely Piloted Aircraft Training (URT): required for Core 18Xers

Halpern Critical Thinking Assessment

Task and Materials. To address the concern that poor Weapons School performance or even attrition from the program was attributed to poor critical thinking skills, students were asked to complete the Halpern Critical Thinking Assessment. Links to complete the Halpern Critical Thinking Assessments were sent to all six students; three of the six students in the class completed the assessment. No additional tasks were required of the students, other than providing brief demographic information.

Measures. Three students completed the Halpern Critical Thinking Assessment, a measure of critical thinking ability measured in five areas: verbal reasoning, argument analysis, thinking as hypothesis testing, likelihood and uncertainty, and decision making and problem solving (Halpern, 1998). Due to constrained-student schedules, an abbreviated, 25 multiple-choice question version of the assessment was distributed via email to students. The dichotomous pass/fail outcome of Weapons School at the end of the six-month program was also measured, as was the Distinguished Graduate and the #2 student for the class.

Procedure. To address the critical thinking portion of the analysis, students were sent links to complete the Halpern Critical Thinking Assessment. Three out of the six students completed the assessment. The scoring was performed remotely by Schuhfried Inc., and individual scores were emailed directly to the experimenter.

Results. The average score for the three students that completed the Halpern Critical Thinking Assessment was 97.7% (97%, 98%, and 98%). Further, at the end of the six-month Weapons School program, it was reported that all six students successfully completed the program. With only three scores, it is not possible to draw any statistically supported conclusion. However, it is notable that all three students received scores over 97%, and that all three successfully completed the program.

Instructor Interviews

Task and Materials. Students at the USAF MQ-9 Weapons School complete a rigorous six-month training program to become Weapons Officers. Because Weapons School is a challenging environment, the researcher observed the routine mission brief, mission, and debrief cycle for each of the six students enrolled in the course, and interviewed instructors post-student brief and post-student debrief. This helped in gaining an in-depth understanding of the research context, the tasks required of Weapons School students, and the typical challenges students face in the program. Further, the goal of the instructor interviews was to expand upon the instructors' definition of critical thinking. During the instructor interview portion, the debrief flow was outlined, providing further understanding of evaluation criteria for the debrief portion. The debrief flow is depicted in APPENDIX G.

Measures. Prior to the May 2016 Weapons School visit to conduct instructor interviews, the researcher contacted the 26 WPS Director of Operations and asked the following questions: (1) in your experience, what are the characteristics of both successful and unsuccessful Weapons School students? (2) what are debrief events/occurrences indicative of successful and unsuccessful students? These questions were intended to direct researcher attention towards both positive and negative behaviors of students during the brief-fly-debrief cycle, as well as contribute to the development of categories from which to code instructor interview responses. The table below summarizes the answers aligned with the related academic theme:

Table 8

Preliminary Instructor Interview

Question	Answers	Academic Themes
1.) In your experience, what are the characteristics of both successful and unsuccessful students?	Successful students: • "able to generalize into a process" • "able to systematically break down problems" and "apply strategies" • "stay calm" and "look past emotion	• **Transfer** • **Problem Solving** • **Emotional self-regulation, emotional competence, and meta-motivation**
2.) What are debrief events/occurrence indicative of successful and unsuccessful WUGs? What are characteristics of successful and unsuccessful debriefs?	Successful students: • "facts should be verified using truth data like mission video or other data sources" • "instructional fixes should be offered that address the when, what, and how to correctly accomplish the deficient task" • Giving feedback where "shortcomings [are] logically and systematically addressed to flesh out all contributing factors so no learning is missed" • "tell the student the criteria that needs to be	• **Evidence-based reasoning (argumentation)** • **Formative Feedback**

	satisfied for the fix to be valid" • This kind of feedback also tells you "the perceptions that need to occur, the decisions that need to be made, and the subsequent actions that need to be performed, to the utmost level of detail"	

During the 15-20-minute instructor interview portion post student brief and debrief, the following questions were asked:

Table 9

Full Instructor Interview Questions

1. How long have you been back at Weapons School as an instructor?
2. What aircraft did you fly prior to the MQ-9 Reaper?
3. In your opinion, what is the biggest reason for student attrition from Weapons School?
4. How well did the student brief? What did the student miss?
5. How well did the student execute what was pre-briefed?
6. What did you attend to in analysis of student performance?
7. What did you observe that was good/bad?
8. Was there any defining event during the brief/mission/debrief (good/bad?)
9. How much of your assessment is based on the student performance in the mission versus their ability to reconstruct the mission accurately (ex: 50/50, 60/40) and effectively debrief?
10. On a 1-5 scale, with 1 being poor and 5 being excellent, how did this student perform today? Why?
11. Do you predict this student will complete the program? Why or why not?

Procedure. All MQ-9 Weapons School students and instructor pilots verbally Consented to participation prior to experimenter brief, mission, debrief observations and prior to instructor pilot interview. Because this study was longitudinal in nature, all students were assigned a number, which served as their identification number for the duration of the study. All student references in observer notes were by student identification numbers; students were not identified by name. At the completion of the course, students' completion/failure was reported to the principal investigator based on their assigned identification number.

Results. Instructor interview question responses were thematically coded into the previously outlined categories of: *transfer, problem solving, emotional self-regulation, emotional competence, meta-motivation, evidence-based reasoning (argumentation), and formative feedback.* The frequencies are detailed in Table 10 below.

Table 10

Instructor Interview Answer Frequencies

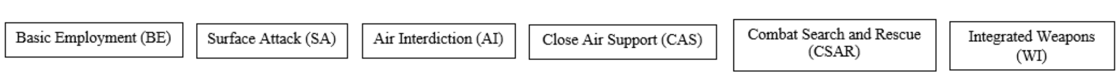
QUESTION	Critical Thinking	Transfer	Emotional Intelligence	Speech Pacing	Evidence-based reasoning
#3	2	4	1		
#4	1	1		2	
#5					
#6		2			2
#7		1			

Answers pertaining to transfer were mentioned most frequently by instructors. Question six, which asked instructors what they specifically attended to in analysis of student performance revealed additional categories: quality of instruction, self-identification of errors, demonstrating procedural knowledge (“how” to do something), use of truth data, artifact usage, adaptability, and adherence to the debrief focus point process. These categories also emerged in the analysis of student grade sheets. Comments pertaining to quality of instruction was addressed five different times, adherence to the DFP process were addressed four times, level of detail twice, truth data twice, self-identification of errors twice, and comments pertaining to the topics of transfer, artifact usage, preparation/planning, effective contributor and adaptability were each mentioned once. These were all categories identified during the grade sheet analysis.

Student Grade Sheets

Task and Materials. The instructor comment portion of student grade sheets for the six phases of the program were also used in the analysis. There were 2-3 missions per phase, equating to a total of 175 missions flown for all six students. Instructor comments were categorized into 21 different categories, with some comments falling into two or more categories. Comment categories are outlined in the measures section. The course phases are outlined below:

Figure 1
MQ-9 Weapons School Phases



Measures. Student grade sheets consisted of an objective analysis of a student's mission performance as well as an instructor comment section, which allowed instructors to provide additional subjective comments on student performance. Measures for this study were derived solely from the instructor comment section of student grade sheets.

Procedure. The 26 WPS provided the student grade sheet comment portion to the researcher in September 2016. Instructor comments pertained to the brief, fly, and debrief portion for each mission flown from January to June 2016. Students were responsible for briefing prior to executing the mission and for debriefing afterwards. After the debrief portion, student and instructor would "switch the pens" and the instructor would provide detailed feedback on the student's performance for the day, before completing the student grade sheet. During the instructor interviews, instructors were asked how much of their assessment was based on the student performance in the mission, versus their ability to reconstruct the mission accurately and effectively debrief. Instructor answers are provided in APPENDIX H, and provide additional information on what instructors focus on during the brief, fly, and debrief process. Based on instructor responses, it became clear that grading depends on how far the student is in the program, but in general, the emphasis in grading is on how well a student debriefs.

Instructor comments for 175 student missions were analyzed and thematically coded into 21 categories. The transcripts were not segmented into units, but rather tallies were made each time one of the 21 categories were mentioned. Unsatisfactory overall performance (brief/mission/debrief) resulted in students repeating failed missions, which is why each student flew different numbers of missions. Student #1 conducted 27 missions, Student #2 conducted 30 missions, Student #3 conducted 29 missions, Student #4 conducted 32 missions, Student #5 conducted 29 missions, and Student #6 conducted 28 missions.

Results. For the analysis of the student's grade sheets, all instructor comments pertaining to the brief, execution and debrief portions were reviewed and coded. A total of 21 different categories were established, with positive or negative occurrences of the behaviors within these categories tallied. The 21 categories are outlined in the below table. An example coded grade sheet is in APPENDIX I. The averaged frequency counts are presented in APPENDIX J. An undergraduate research assistant assisted with coding instructor comments into the various categories, and coded all six students grade sheet comment portions independent of the researcher. Prior to independently coding, the undergraduate research assistant underwent a two-hour coding training session with the researcher. The average scores for category frequency of both the researcher and undergraduate were calculated and used for the final analysis.

Table 11

Grade sheet instructor comment categories

Adherence to DFP Process
Flow
Weight of Effort & Pacing
Level of Detail/Specificity/Depth
Instructor Pilot (IP) Assistance
Efficiency
Organization
Preparation/Planning
Repeatable Process/Methodology
Decision Criteria
Triggers
Contracts
Quality of Instruction
Artifact/Tool Usage
Transfer
Identification of Errors
Collaboration
Weapons Officer Qualities
Valuable contributor during mass debrief
Adaptability
Flight and Area of Operations (AO) Leadership

Inter-rater reliability. To examine interrater reliability, a second coder counted the number of occurrences of each theme on each student's grade sheet. Across both raters, counts ranged from 0 to 31 with a mean count of 3.5. Differences between the counts of the two raters were taken for each theme. On average the two raters differed by 1.21 with a range of difference between 0 and 21. Because differences between raters was minimal, the mean of the two was used in cases of disagreements.

Table 12

Inter-rater reliability calculations

WUG-1 (27 msns)	**1.1**
WUG-2 (30 msns)	**1.5**
WUG-3 (29 msns)	**1.4**
WUG-4 (32 msns)	**1.1**
WUG-5 (29 msns)	**1.0**
WUG-6 (28 msns)	**1.1**

Coding Methodology. Due to the qualitative nature of the student grade sheets, a comprehensive coding methodology was needed to accurately categorize student behaviors. The complete coding methodology can be found in APPENDIX K. Instructor comments fell into 21 different categories, with some comments being dual coded if they fell into multiple categories. Examples of comments that were dual coded are provided in the below table:

Table 13

Overlapping comment examples

Overlapping instructor comments	**Applicable categories**
"did a good job at explaining how he was going to execute the briefed contracts based on specific triggers and criteria"	Procedural knowledge, contracts, triggers, and decision criteria
"debrief was efficient and covered all the learning from the sortie"	Efficiency, learning captured
"Reconstruction was not timely and was incomplete, requiring significant IP input to complete debrief"	Adherence to DFP Process, and IP inject
"Examples were well thought out, and instruction was specific and flowed in a logical manner"	Preparation, quality of instruction, and flow

MQ-9 Weapons School Student Profiles.

Profiles were generated for all six Class 16A students depicting the percentage of instructor comments pertaining to the 21 different categories identified in the data coding phase. The pie charts revealed student behaviors that were most salient to the instructors. For all six students, instructor comments fell broadly into four categories – quality of instruction (20.5%), adherence to the debrief focus point process (14.2%), level of detail (12.5%), and weight of effort (7.3%). The other 17 categories constituted 45.5% of the remainder of instructor comments. The averaged percentage breakdown for all of the remaining categories can be found in Table 21.

The USAF Weapons School primary mission is to "teach graduate-level instructor courses, which provide the world's most advanced training in weapons and tactics employment to officers of the combat air forces" (United States Air Force, 2016). With this in mind, it is logical that 20.5% of all instructor comments assessed the quality of student instruction. Additionally, delivering an efficient and effective debrief is a cornerstone of the instruction students receive in Weapons School; this aligns with the fact that 14.2% of all instructor comments pertained to adherence to the debrief focus point process that students are taught at the beginning of the course.

The goal of this research was to determine the competencies distinguishing the top student(s) from their classmates in Weapons School. Because Weapons School awards the honor of Distinguished Graduate at the end of the course, it was hypothesized that the competencies that the distinguished graduate exhibits differ from those of their classmates. For Class 16A, the top two students (student #5 and student #1) competed for the honor until the very end of the course, when student #5 received the title of Distinguished Graduate.

Profiles were generated for all six students in the form of pie charts displaying the various percentages of categories of instructor comments on each student. The pie chart profiles depict the nature of instructor comments, but not whether they are positive or negative comments within a category. These can be viewed in Figures 2, 3, 4, 5, 6, and 7. Percentages were calculated by taking the frequency of comments in any given category and dividing by the total number of comments for that particular student. For example, if the total number of comments (both positive and negative) in the Quality of Instruction category for a student was equal to 27, and the total number of comments for that student, across all of their grade sheets, was equal to 146, the percentage of comments pertaining to Quality of Instruction would be equal to 18.49%.

Because students #5 and #1 were essentially the top two students throughout the course, a combined Distinguished Graduate profile was created for them. A separate profile was created for students #2, #3, #4 and #6. Because the quality of instruction, adherence to DFP process, level of detail/specificity/depth, and weight of effort and pacing categories constituted 54.5% of the instructors' comments for all six students (20.5%, 14.2%, 12.5%, and 7.3%, respectively), indicating that these were the most salient categories to the instructors, these will be the categories focused on in the discussion to follow. Further, these categories are broken down into the percentage of positive and negative comments within each category. Percentages of positive and negative comments for any category were calculated by taking the number of positive or negative comments within a given category, and dividing by the total number of comments for that particular category.

Contingency tables were created for comparisons across student groups (top 2 students vs. their peers and student 5 vs. student 1) to directly compare theme count differences across the four categories that constituted the majority of instructor comments: quality of instruction, adherence to DFP process, level of detail, and weight of effort. For every action a student took in their brief or debrief, the corresponding instructor comment on the grade sheet had an equal chance of being a positive or negative assessment of this action. In the calculation of observed values for the contingency tables, because each outcome within a subcategory had an equal chance of happening, observed counts were equal to 50% of the total count for a group within a subcategory. For example, in the below contingency table for the Quality of Instruction subcategory, the observed value of 10.25 for students 1/5 is half of 20.5, the total count for quality of instruction for students #1 and #5.

Table 14

Contingency Table for Quality of Instruction Subcategory

Quality of Instruction	**Students 1/5**	**Students 2/3/4/6**
Quality of Instruction (+)	12.25 (10.25)	6.875 (12.25)
Quality of Instruction (-)	8.25 (10.25)	17.625 (12.25)

Quality of Instruction. Figure 8 is a comparison of the top two students (students #5 and #1) vs. students #2, #3, #4, and #6 in the Quality of Instruction Category. Students #5 and #1 outperformed students #2, #3, #4, and #6 in the positive Quality of Instruction subcategory (43% vs 23%) and in the Learning Captured (positive) subcategory (25% vs

11%). For the negative categories students #2, #3, #4, and #6 exhibited poor quality of instruction 57% of the time vs. 29% of the time for students #5 and #1. Students #2, #3, #4, and #6 missed opportunities for learning 9% of the time vs. only 3% of the time for students #5 and #1.

Adherence to the Debrief Focus Point Process. Figure 9 shows a comparison of students #5 and #1 vs. students #2, #3, #4, and #6 in the Adherence to DFP Process Category. Students #5 and #1 outperformed students #2, #3, #4, and #6 in the positive Adherence to DFP process subcategory (62% vs 32%), and in the correctly pulled truth data subcategory (15% vs 3%). In the negative categories, students #2, #3, #4, and #6 failed to adhere to the DFP process 60% of the time vs. only 20% of the time for student #5 and #1. Additionally, students #2, #3, #4, and #6 incorrectly or struggled to collect truth data (negative) 5% of the time vs. 3% of the time for students #5 and #1. A contingency table was created for the adherence to DFP process subcategories, showing the positive and negative for both group; this can be viewed in Table 16.

Level of Detail/Specificity/Depth. Figure 10 is a comparison of students #5 and #1 vs. Students #2, #3, #4, and #6 in the Level of Detail/Specificity/Depth Category. Students #5 and #2 outperformed their classmates in the good amount of detail subcategory (17% vs. 9%). The good amount of detail and demonstrated procedural knowledge subcategories were the positive categories in the overall level of detail/specificity/depth category, whereas the lack of detail and the missing procedural knowledge subcategories were the negative categories in the overall category. Further, students #5 and #1 outperformed students #2, #3, #4, and #6 in the lack of detail subcategory, meaning that they still omitted details in briefing/debriefing that required it; just to a lesser extent (48% vs 57%), and in the procedural knowledge subcategory (9% vs. 6%). Students #5 and #1 also outperformed students #2, #3, #4, and #6 in the lack of procedural knowledge subcategory (26% vs. 28%). Both groups demonstrated deficiencies in this category; students #5 and #1 just exhibited deficiencies to a lesser extent. A contingency table was created for the level of detail subcategories, showing the positive and negative counts for both groups; this can be viewed in Table 17.

Weight of Effort and Pacing. Figure 11 shows a comparison of students #5 and #1 vs. Students #2, #3, #4, and #6 in the Weight of Effort and Pacing Category. Students #2, #3, #4, and #6 outperformed the top two students in the positive proper weight of effort subcategory; practicing proper weight of effort 4% of the time vs. 2% of the time for the top two students. The top two students outperformed their classmates in the positive good pacing subcategory (20% vs. 6%), and in the negative poor pacing subcategory, demonstrating poor pacing 24% of the time vs. 31% of the time for their peers. Lastly, the top two students outperformed their classmates in the negative misplaced weight of effort subcategory (54% vs. 59%). Both groups demonstrated a misplaced weight of effort; the top two students just did so to a lesser extent. A contingency table was created for the pacing subcategory, showing the positive and negative counts for both groups. This can be found in Table 18.

Distinguished graduate profile. A deeper analysis of the differences between the the #1 and #2 student further reinforced the importance of instructional quality in instructors' evaluation of Weapons School students. The distinguished graduate (student

#5) outperformed the #2 student (student #1) in positive behaviors in both the quality of instruction and adherence to DFP process categories. These were also the two categories that constituted much of instructor comments for all six students at 20.5% and 14.2%, respectively. The Distinguished Graduate demonstrated positive quality of instruction in 85% of the instructor comments for this category (vs. 50% for the #2 student), and demonstrated adherence to the DFP process in 77% of the instructor comments for this category (vs. 75% for the #2 student). Contingency tables were created for the quality of instruction and adherence to DFP process subcategories to compare positive and negative counts for both groups; these can be viewed in Tables 19 and 20.

The #2 student outperformed the #1 student in the Level of Detail category, providing an appropriate level of detail 27% of the time, vs. 25% of the time for the #1 student. Additionally, the #2 student outperformed the #1 student in the weight of effort category; demonstrating appropriate weight of effort and pacing 28% of the time vs. 19% of the time for the #1 student. A more detailed student profile comparison can be found in Figures 12, 13, 14, and 15.

Discussion

The purpose of this research was to first, determine the competencies which distinguished successful and unsuccessful students in USAF Weapons School. After determining these competencies and behaviors, the next goal was to determine whether critical thinking skills were truly the distinguishing factor in performance, which required expanding upon the concept of critical thinking from the Weapons School perspective. Lastly, the final purpose of the research study was to create a more practical guide of “soft” and “hard” skills required to be successful in USAF Weapons School. It was hypothesized that because Weapons School has a distinguished graduate, who is recognized as the top all-around student in the class, that the behaviors of this selected student differed from their peers. This hypothesis was supported in this study – the behaviors that the distinguished graduate (student #5) exhibited, differed from those of his peers. Since the top two students competed until the end of the course, both students exhibited differing behaviors from their peers.

Recommendations. Based on the analysis, it appears that Weapons School instructors are more focused on the outputs of critical thinking versus explicit critical thinking. There was no explicit mention of “critical thinking” in any of the student grade sheets; however, there was a focus on the products of critical thinking – specifically how it is manifested in students’ instruction to their audience, in both their briefs prior to flying and debriefs after flying. Based on this finding, it is recommended that the USAF Weapons School conduct further research before determining if a critical thinking assessment should be added to the current candidate selection battery.

The findings from the abbreviated, 25 question, multiple choice administration of the Halpern Critical Thinking Assessment are inconclusive. The respective scores for the three students that completed the Halpern Critical Thinking Assessment were 97%, 98%, and 98%. Further, all six students in the class successfully graduated from the Weapons School course. Since all six students graduated the program, with only three students taking the assessment and scoring an average of 97.7% on the assessment, it is recommended that the USAF Weapons School administer the full assessment.

The full version still assesses performance in the five dimensions of critical thinking: decision making and problem solving, thinking as hypothesis testing, argument analysis, likelihood and uncertainty, and verbal reasoning (Halpern, 1998). However, in this version, the dimensions are weighted differently, with the decision making and problem solving facet being weighted with more total points than the other categories (Halpern, 1998). The full HCTA consists of 20 everyday scenarios, briefly described to the test taker. Further, in each scenario, respondents are asked an open-ended question (short response), followed by a forced choice question (e.g. multiple choice, ranking, or rating of alternatives) (Halpern, 1998). The dual response format allows test takers to respond to a situation in their own words, then tests their ability to recognize a good response. The full version is designed in this manner to differentiate between free recall and recognition processes in memory, with the total score equally weighted between the constructed response and forced choice questions (Halpern, 1998).

Because of the instructor comment emphasis on the quality of instruction and adherence to debrief focus point process categories, it is recommended that an additional screening tool is implemented to focus on these areas. Rather than assessing candidates' abilities to perform in these areas during their Weapons School pre-evaluation, or "WIC look", which occurs during a separate visit to Nellis AFB, prior to candidate selection, it is proposed that they are evaluated with more rigor at their home units for these traits. The statistically significant difference between the top two students and their peers in the quality of instruction category, suggests underlying differences in these students' instruction performance and abilities prior to entering Weapons School. Students were required to have a minimum of 75 instructor pilot hours prior to the start of the course. Based on the results of this study, we can assume that while the 75-hour benchmark is sufficient in selecting candidates that will graduate the rigorous course, there is still a great deal of variability in the instructional quality among selected candidates. If the goal is to ensure that candidates entering Weapons School perform well in the quality of instruction and debrief focus point process categories, another recommendation is that these skills are taught more heavily at the home units, prior to Weapons School selection and attendance.

It was expected that because Weapons School has a distinguished graduate, who is recognized as the top all-around student in the class, that the competencies the distinguished graduate exhibited differed from those of their classmates. This expectation was validated through the data analysis. Because students #5 and #1 performed similarly throughout the Weapons School course, with student #5 ultimately earning the title of distinguished graduate, this suggests that both students possessed desirable Weapons School student traits. Overall, all six students were successful in the course, suggesting that they all possessed desirable traits; students #5 and #1 simply outperformed their peers, and demonstrated these traits to a greater extent. For the purposes of data analysis, the grade sheet data for both students were combined and analyzed against the data of their four peers.

Desired Weapons School Student Competencies and Behaviors. The top two students significantly differed from their peers in the quality of instruction and adherence to debrief focus point categories. Further, the distinguished graduate (student

#5) significantly differed from the #2 student (student #1) in these two categories; this was validated statistically. Thus, we can conclude that these two categories contained the behaviors and competencies required to distinguish oneself in performance at Weapons School. Further, these two categories separated the very best students from other successful students. The coding methodology section outlined select instructor grade sheet comments across all 21 categories. Example instructional quality comments from the grade sheets are listed in Table 34, and adherence to DFP process comments are outlined in Table 22. The table below transforms examples of positive instructor comments from the grade sheets in these categories into "hard" and "soft" skills, creating a more practical guide for outlining student success in Weapons School.

Table 14

Desired Weapons School competencies

Example instructor comment from grade sheet	**Underlying competency**
"Examples were well thought out, and instruction was specific and flowed in a logical manner" "Briefing was outstanding, hit highlights and good explanation to sensors and #2 (Wingman) on criteria, roles, and tasks"	Specific instruction which provides well thought out examples suggest **advanced presentation and communication skills**

"Highlight was WUG's presence and ability to scope debrief appropriately" "WUG did a nice job tailoring examples to the scenario"	Scoping and tailoring suggest **advanced instructional skills.** Novice teachers tend to go into too much detail, or not enough detail when instructing (M. Niemczyk, personal communication, October 24, 2016). The expectation in Weapons School is that students are able to tailor brief and debriefs to their audience.
"Debrief focus point (DFP) addressed the correct issue and student had correct contributing factors (CF)" "able to identify the DFP question and reconstruct the sortie in accordance with phase standards"	**Argument analysis skills –** Halpern (1998) defines an argument as a set of statements with at least one conclusion and one reason that supports the conclusion. Within the debrief focus point process, the various contributing factors for any given debrief focus point can be viewed as the set of statements in the argument. It is up to the student to conclude which contributing factor served as the root cause for the debrief focus point (derived from a mission error).

The competencies in the table were all categorized as "soft" skills, as they were not technical in nature. This stands to reason, considering that upon Weapons School entry, students are expected to have a baseline level of technical competency in their airframe. However, Weapons School is a unique environment, requiring combination of both "hard" and "soft" skills. As novice MQ-9 pilots, Weapons School candidates focused on first mastering the technical or "hard" skills of their aircraft – focusing on flying the plane and deploying munitions. Upgrading to instructor pilots at their home units required use of more "soft" skills, as they were learning how to instruct the technical skills to new MQ-9 pilots. After a minimum 75-hours as an instructor pilot, the most qualified instructor pilots are selected for Weapons School attendance – a program which expects students to have most of the technical skills upon entry. After close analysis of student grade sheets, however, it was the “soft” skills that proved to be the most important aspects of student performance.

The benefit of the distinguishing competencies all being “soft” skills is that these can be taught to students demonstrating Weapons School potential. Advanced presentation, communication, instructional, and argument analysis skills could be pre-taught at home units and evaluated by Weapons Officers. Presently, there is more of a focus on developing instructional skills at home units. The instructor upgrade program, which trains MQ-9 pilots to become instructors at their units, focus on instructing the types of learning, types of communication, how to present information, and how to brief, debrief, and instruct in the seat (M. Dunn, personal communication, January 27, 2017). The expectation should not be that students arrive at Weapons School with perfect presentation, communication, instructional, and argument analysis skills – only with average to above average skills, with the capacity to improve in these areas with further instruction and practice.

Study Limitations. The biggest concern in the study is the limited sample size, which makes it difficult to draw conclusions from the Halpern Critical Thinking Assessment that are statistically reliable. The current MQ-9 Weapons School class has only six members in it. Class sizes are intentionally kept small because of the rigorous course structure and limited resources, such as instructors and ground control stations. An additional limitation is the restricted range of the study sample – all six students completed the course, which is the goal. Because all six students graduated the course, this study more specifically distinguishes between excellent and very good candidates.

The findings from this study pertain to selecting candidates who will not likely only succeed in the Weapons School program, but who will be competitive for the Distinguished Graduate award.

CONCLUSION

The most valuable outcome from this study is that MQ-9 Weapons School instructors are assessing students for behaviors that directly align with the Weapons School primary mission, which is, "to teach graduate-level instructor courses, which provide the world's most advanced training in weapons and tactics employment to officers of the combat air forces" (United States Air Force, 2016). Further, students are primarily chosen based on, "their ability to instruct in their weapon system, which implies a high standard of credibility, integrity, and affability" (Rosales, 2006).

However, to further improve the process of candidate selection, it is recommended that the Weapons School conduct a more thorough and standardized pre-evaluation of candidates' quality of instruction. This should be done at candidates' home duty stations, with the hopes of conducting a more objective evaluation of potential Weapons School candidates. It is recommended that MQ-9 Weapons Officers develop the objective evaluation criteria; as they know the level of instructorship required to be successful at Weapons School, and subsequently after graduation. Additionally, in grooming potential Weapons School students, training at the home units should incorporate techniques for improving the quality of instruction as well as a more rigorous introduction into the debrief focus point framework. The DFP framework is taught to students at their home units, but with more of a "training wheels" approach in comparison to what they would face at Weapons School (M. Dunn, personal communication January 27, 2017). This could be implemented with specific Weapons School debrief focus point process instruction, or a more top-level root cause analysis type training, such as identifying the root cause of mishaps in books such as “Set Phasers on Stun: And Other True Tales of Design, Technology, and Human Error” (Casey, 1998).

The use of remotely piloted aircraft in military applications will continue to grow in upcoming years. From the United States Air Force’s perspective, RPAs will continue to be at the forefront of the fights of today and tomorrow; thus, recruiting and retaining RPA pilots will remain of paramount importance. With the Air Force researching methods to broaden the pool of potential RPA candidates (L. Barron, personal communication, February 1, 2016), equal effort needs to be placed in broadening the pool of knowledge, skills, and other characteristics to screen for. This will ensure that airmen continue to possess abilities to adapt and respond faster than our potential adversaries.

Tables

Adherence to Debrief Focus Point Process	**Students 1/5**	**Students 2/3/4/6**
Adhered to process	10 (6.625)	7 (10.0625)
Failed to adhere	3.25 (6.625)	13.125 (10.0625)

Table 15: Contingency Table for Students 1/5 and 2/3/4/6 in the Adherence to Debrief Focus Point Process

Level of Detail	**Students 1/5**	**Students 2/3/4/6**
Provided sufficient level of detail	3.25 (6.25)	9.875 (5.75)
Provided insufficient level of detail	9.25 (6.25)	1.625 (5.75)

Table 16: Contingency Table for Students 1/5 and 2/3/4/6 in the Level of Detail Category

Weight of Effort/Pacing	**Students 1/5**	**Students 2/3/4/6**
Proper Pacing	2 (2.25)	0.625 (2.125)
Improper Pacing	2.5 (2.25)	3.625 (2.125)

Table 17: Contingency Table for Students 1/5 and 2/3/4/6 in the Pacing Category

Quality of Instruction	**Student 5**	**Student 1**
Quality of Instruction (+)	14.5 (9.25)	10 (11.25)
Quality of Instruction (-)	4 (9.25)	12.5 (11.25)

Table 18: Student 5 vs. Student 1 in the Quality of Instruction Subcategory

Adherence to Debrief Focus Point Process	**Student 5**	**Student 1**
Adhered to process	10.5 (6.75)	9.5 (6.5)
Failed to adhere	3 (6.75)	3.5 (6.5)

Table 19: Student 5 vs. Student 1 in the Adherence to DFP Process Subcategory

Category	Percentage of Comments (Average for all students)
Adherence to DFP Process	14.07%
Flow	5.20%
Weight of Effort/Pacing	7.28%
Level of Detail	12.52%
Instructor Pilot Assistance	5.43%
Efficiency	1.75%
Organization	0.11%
Preparation/Planning	3.26%
Repeatable Process/Methodology	4.08%
Decision Criteria	2.73%
Triggers	1.99%
Contracts	2.15%
Quality of Instruction	20.69%
Artifact/Tool Usage	0.95%
Transfer	4.62%
Identification of Errors	2.98%
Collaboration	1.50%
Weapons Officer Quality	2.42%
Valuable contributor	2.47%
Adaptability	1.94%
Flight Leadership	1.85%

Table 20: Averaged percentage breakdown of instructor comments for all six students

Figures

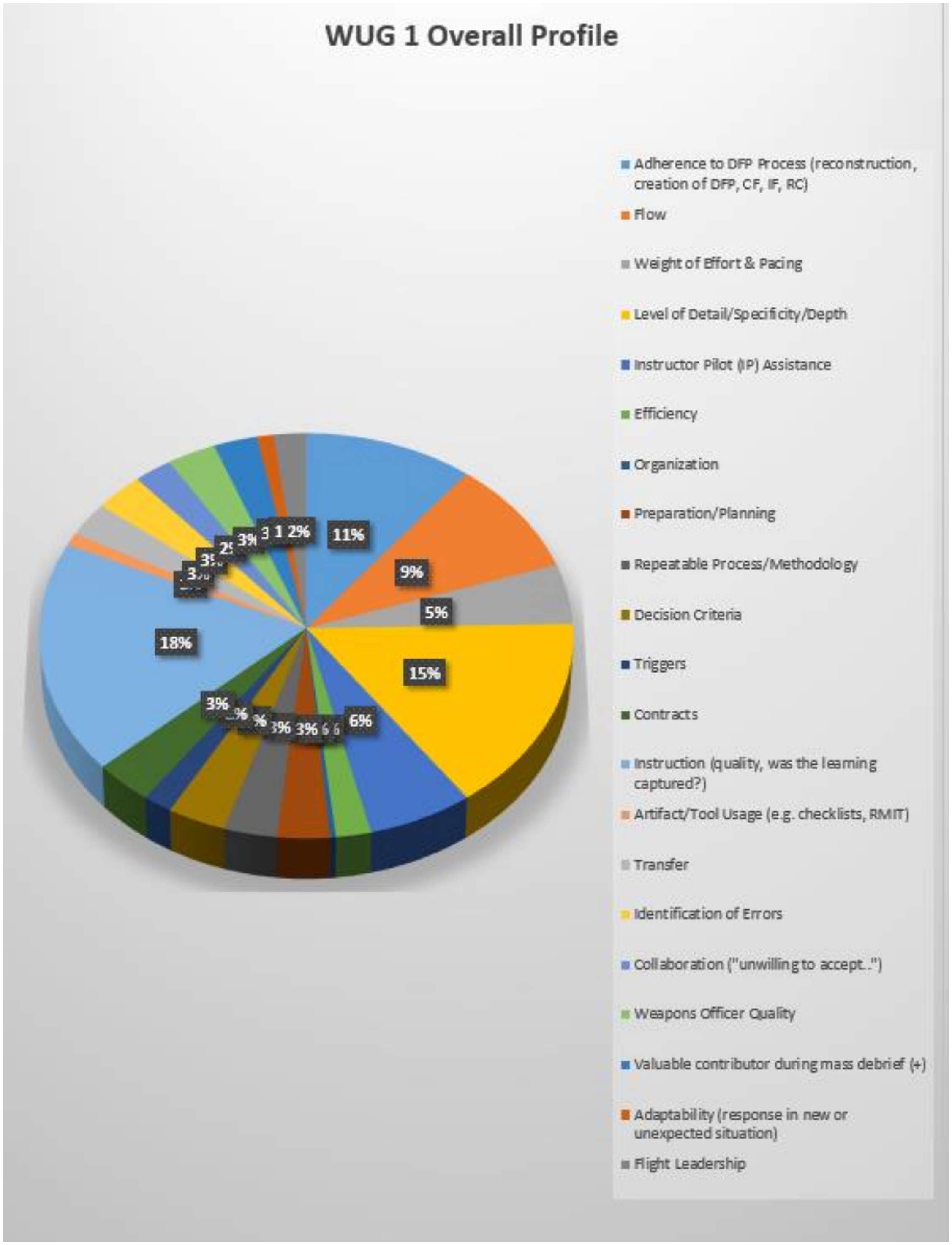


Figure 2: Student #1 Profile

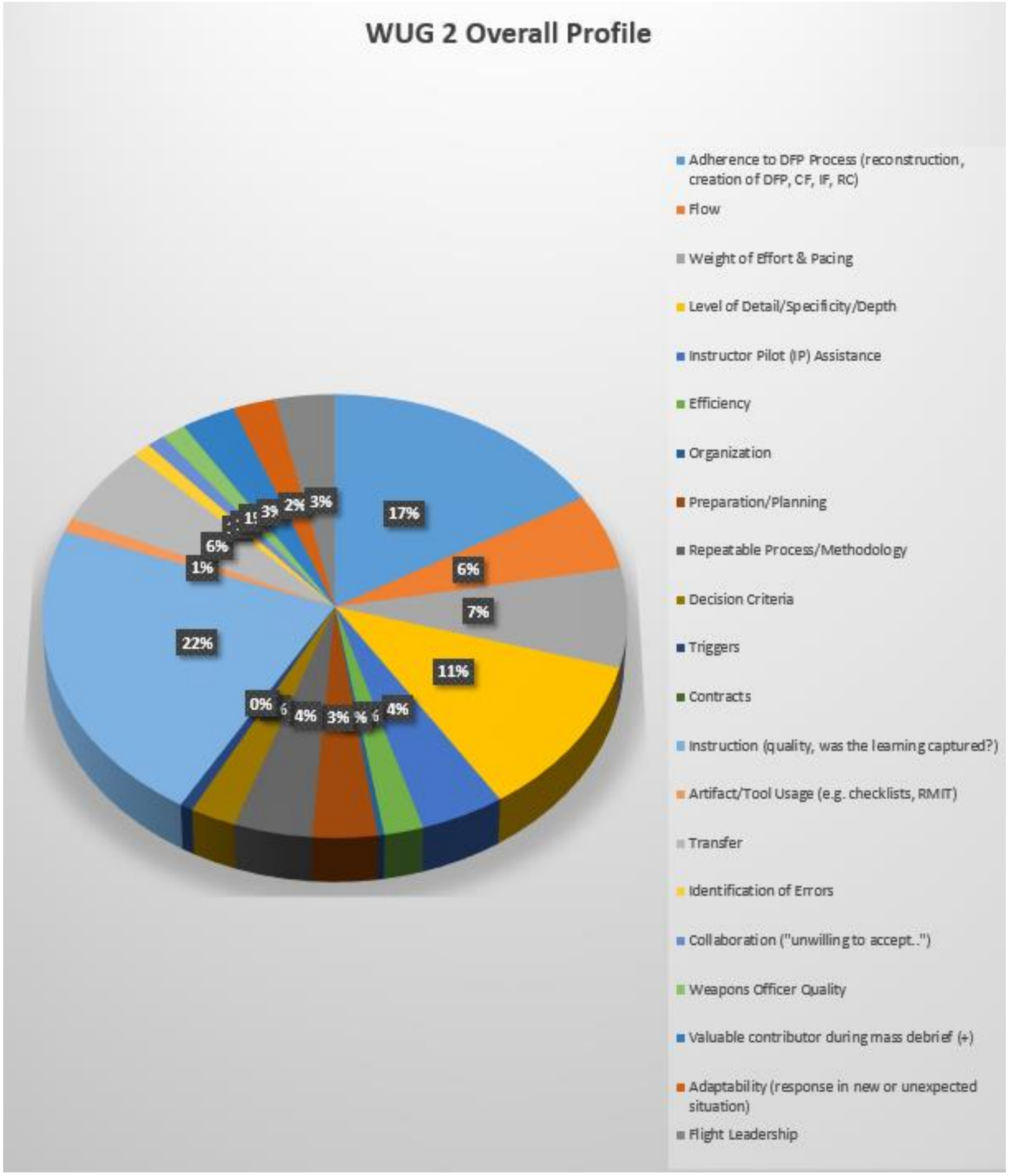


Figure 3: Student #2 Profile

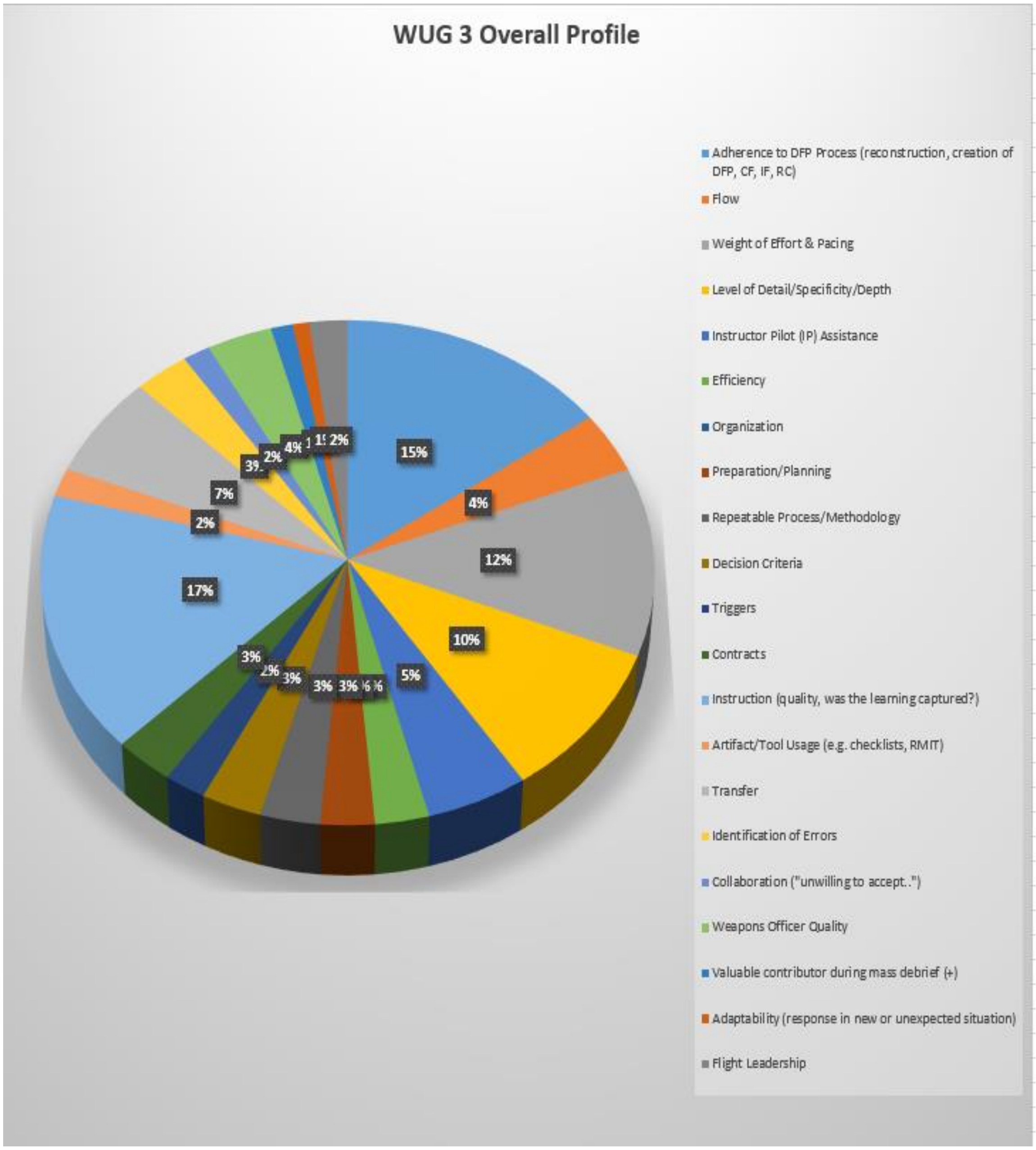


Figure 4: Student #3 Profile

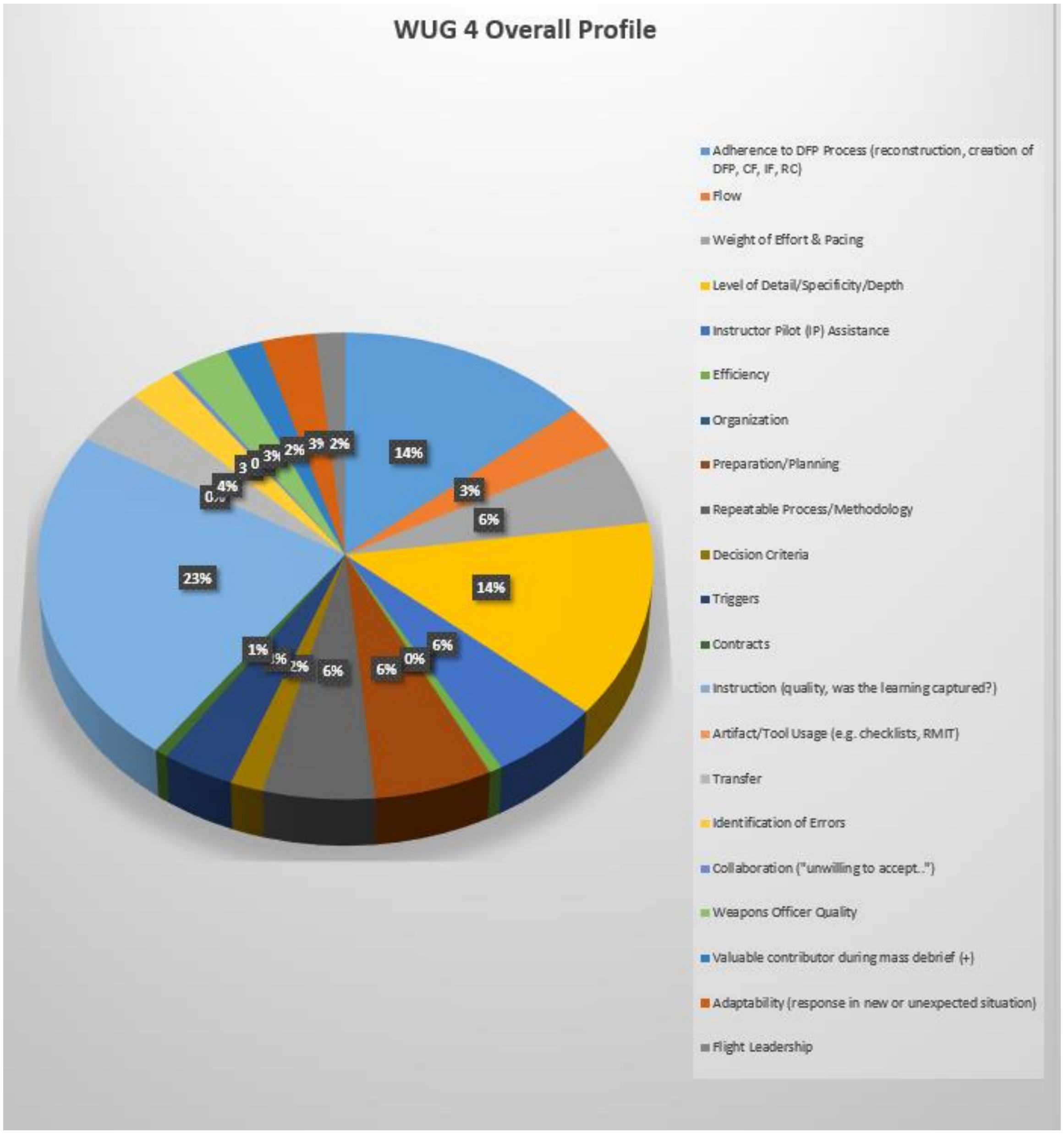


Figure 5: Student #4 Profile

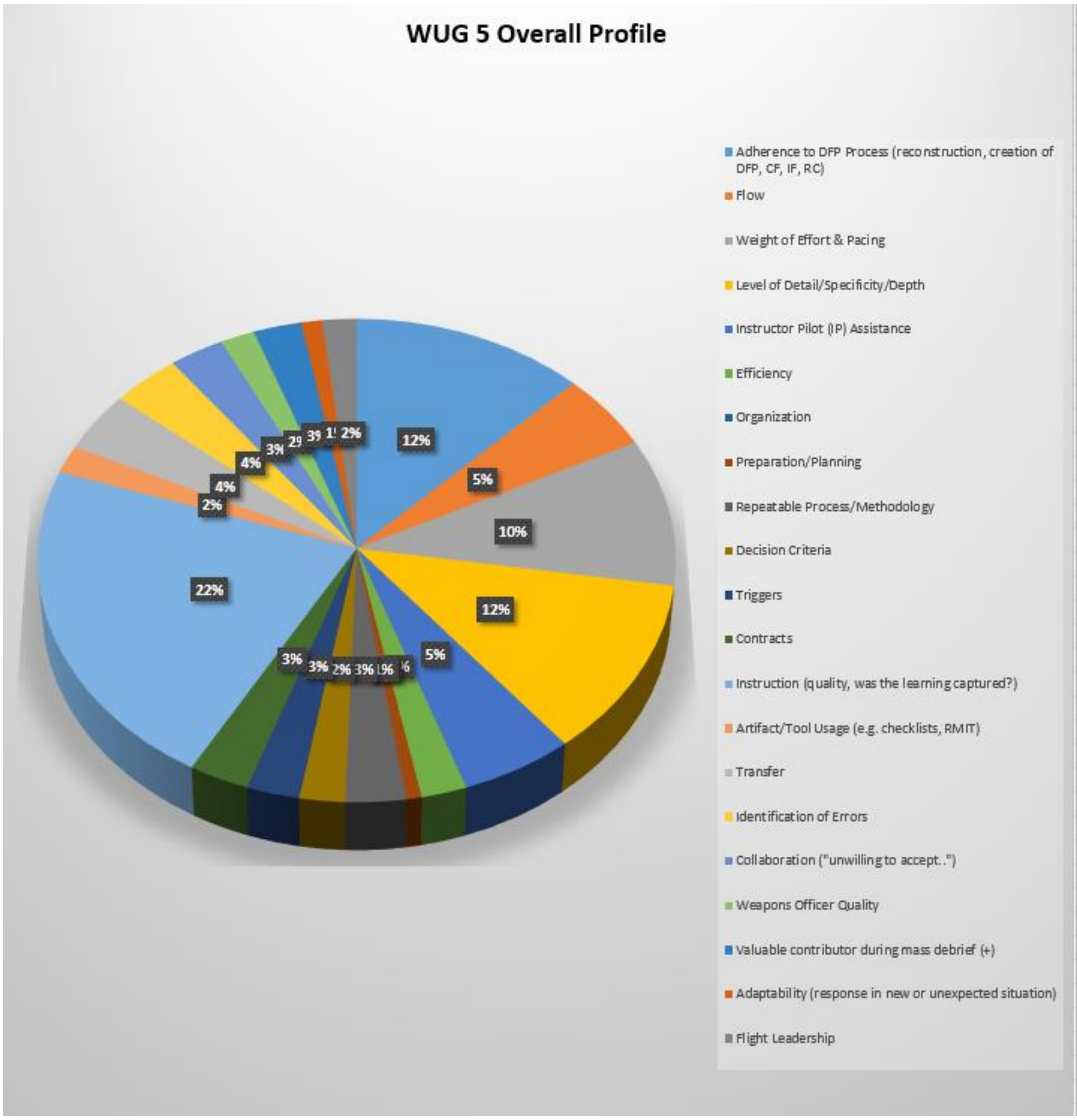


Figure 6: Student #5 Profile

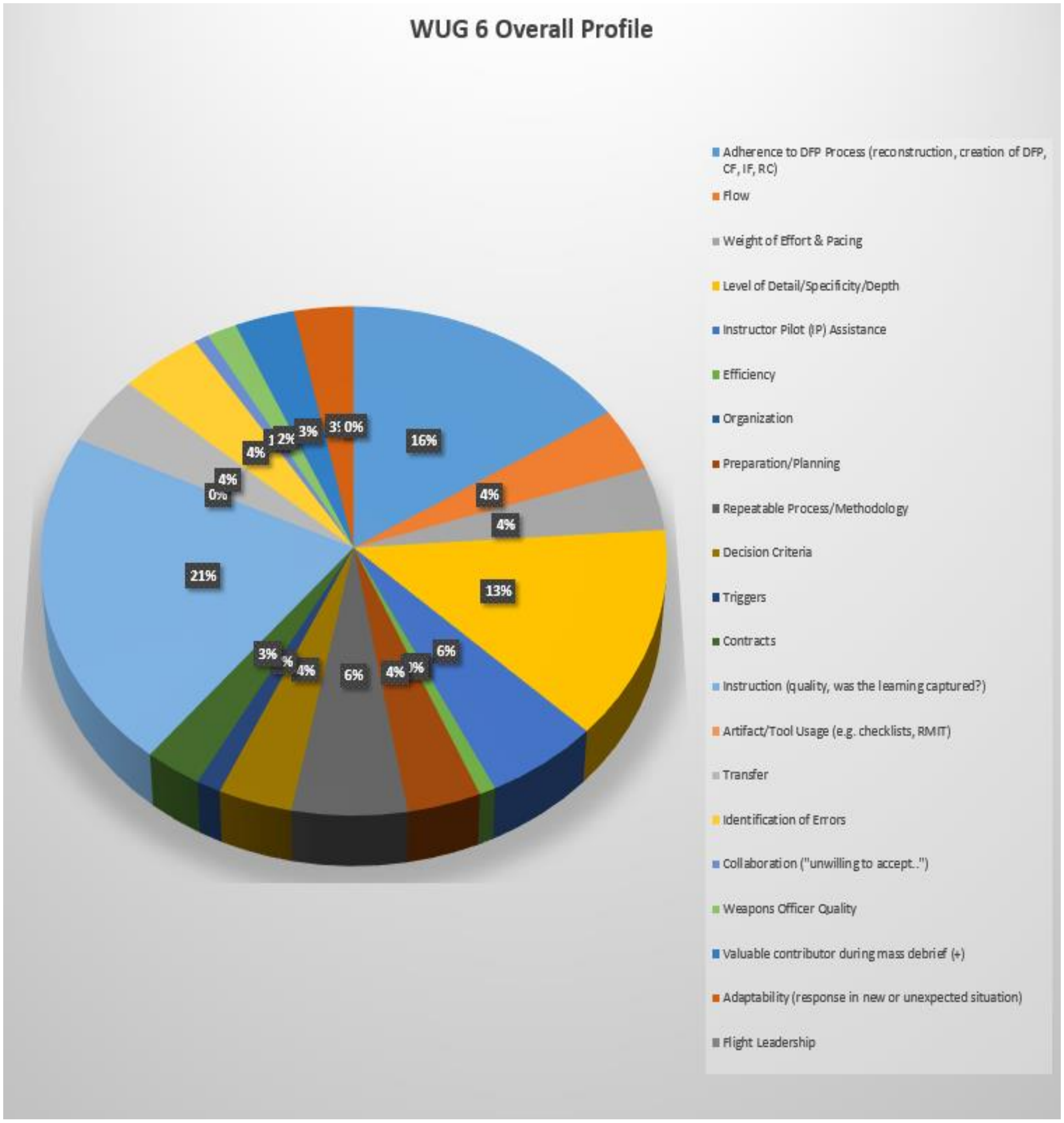


Figure 7: Student #6 Profile

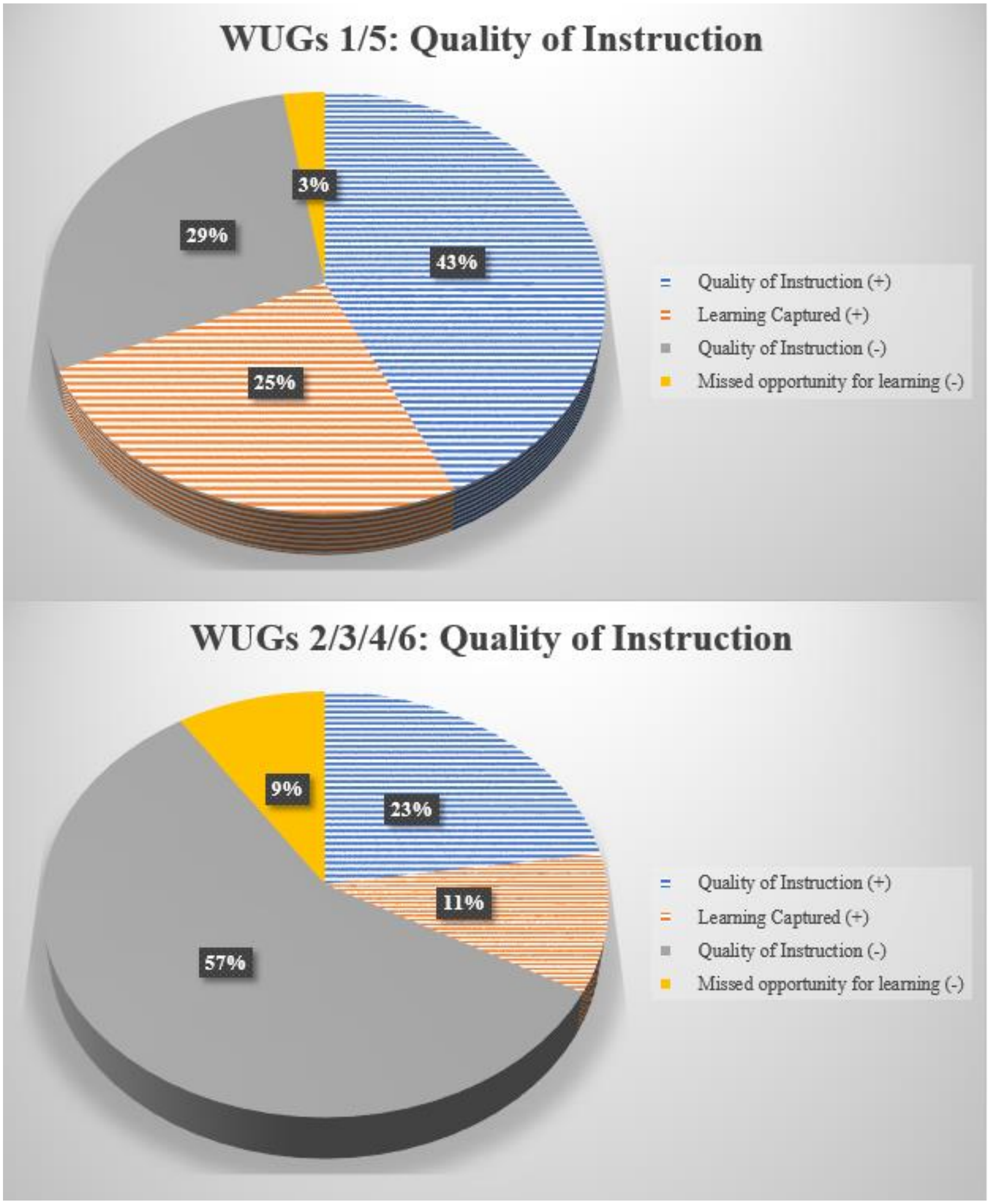


Figure 8: Students 1 and 5 vs. Students 2, 3, 4, and 6 in Quality of Instruction Category

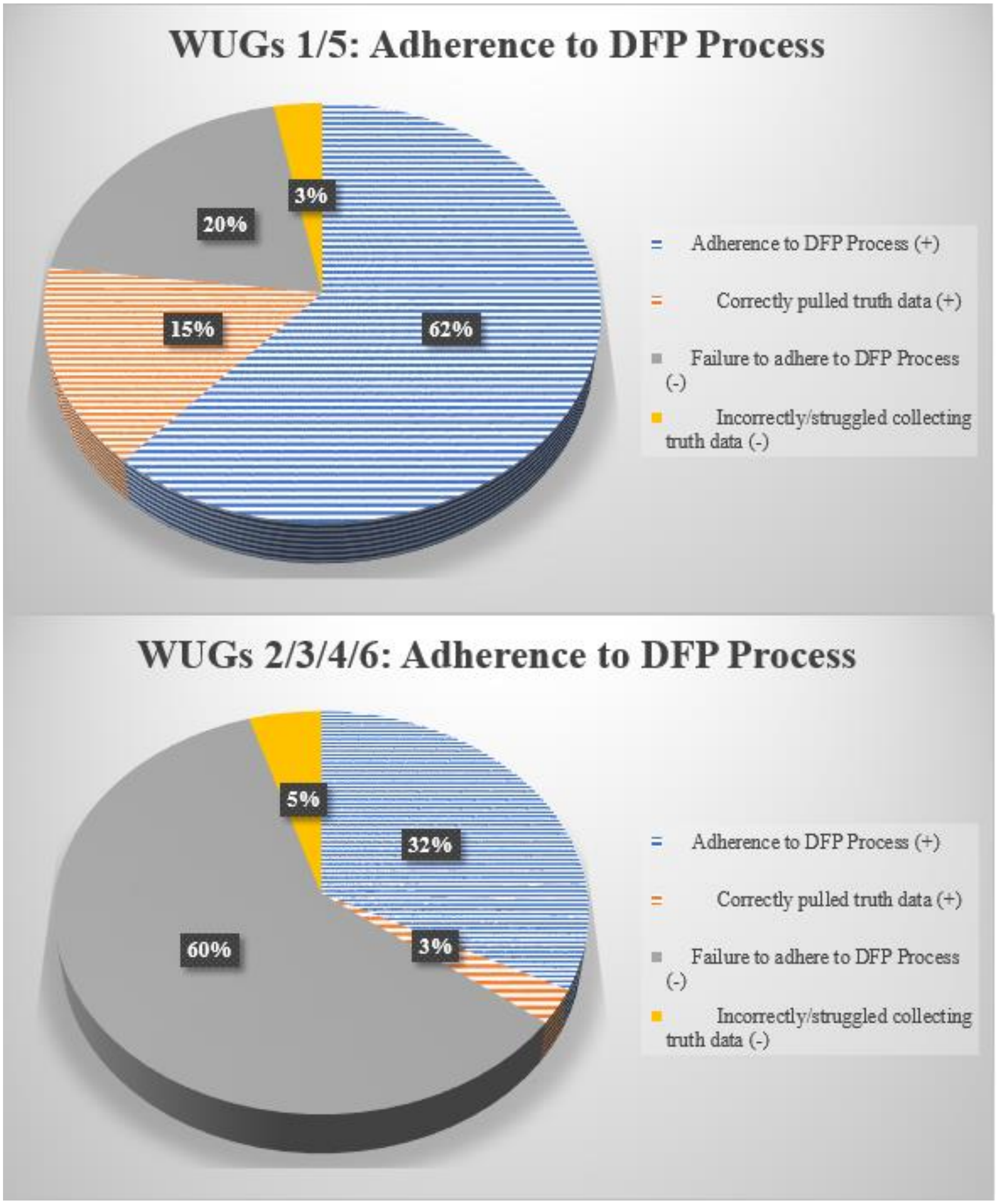


Figure 9: Students 1 and 5 vs. Students 2, 3, 4, and 6 in Adherence to DFP Process Category

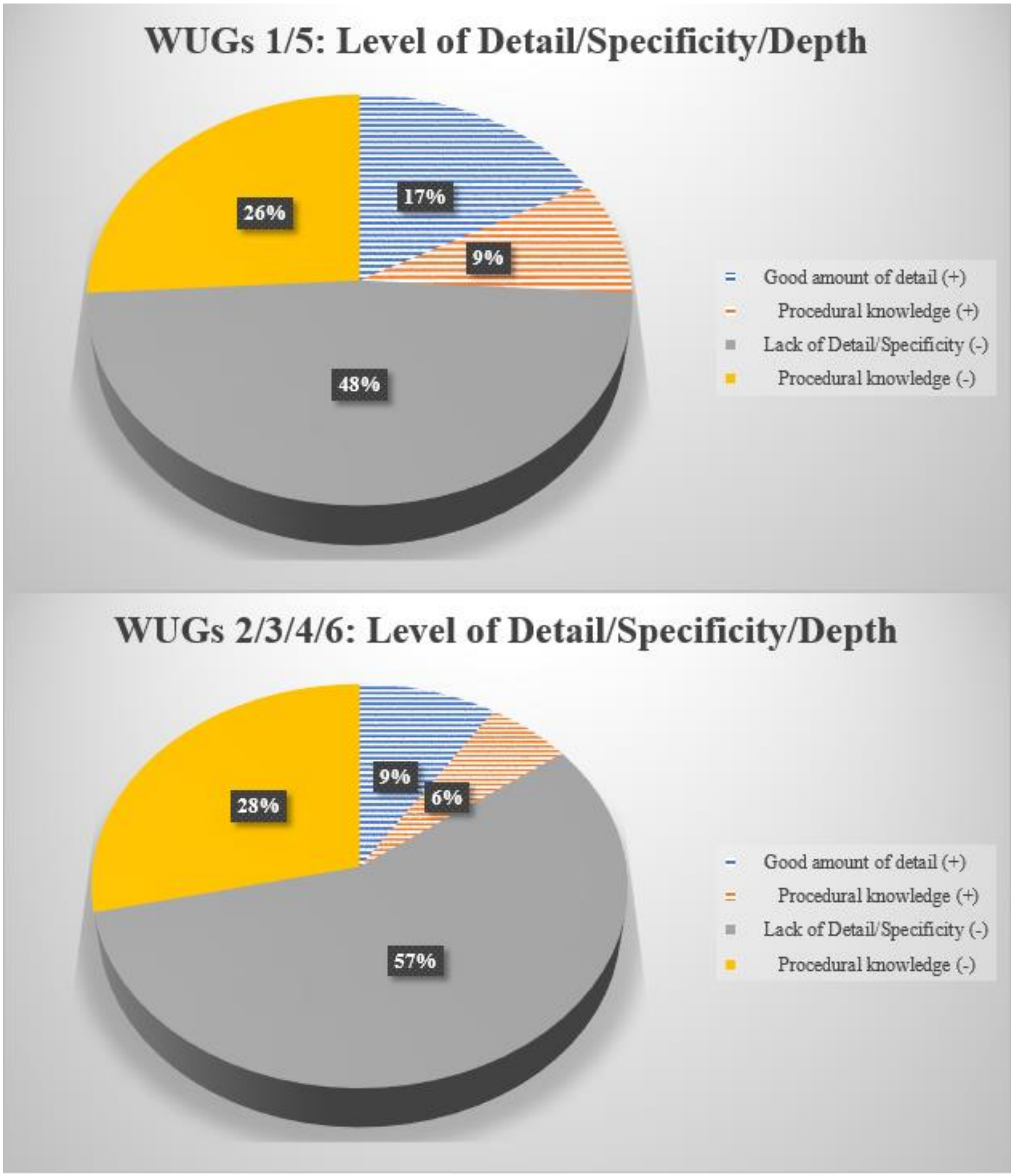


Figure 10: Students 1 and 5 vs. Students 2, 3, 4, and 6 in Level of Detail Category

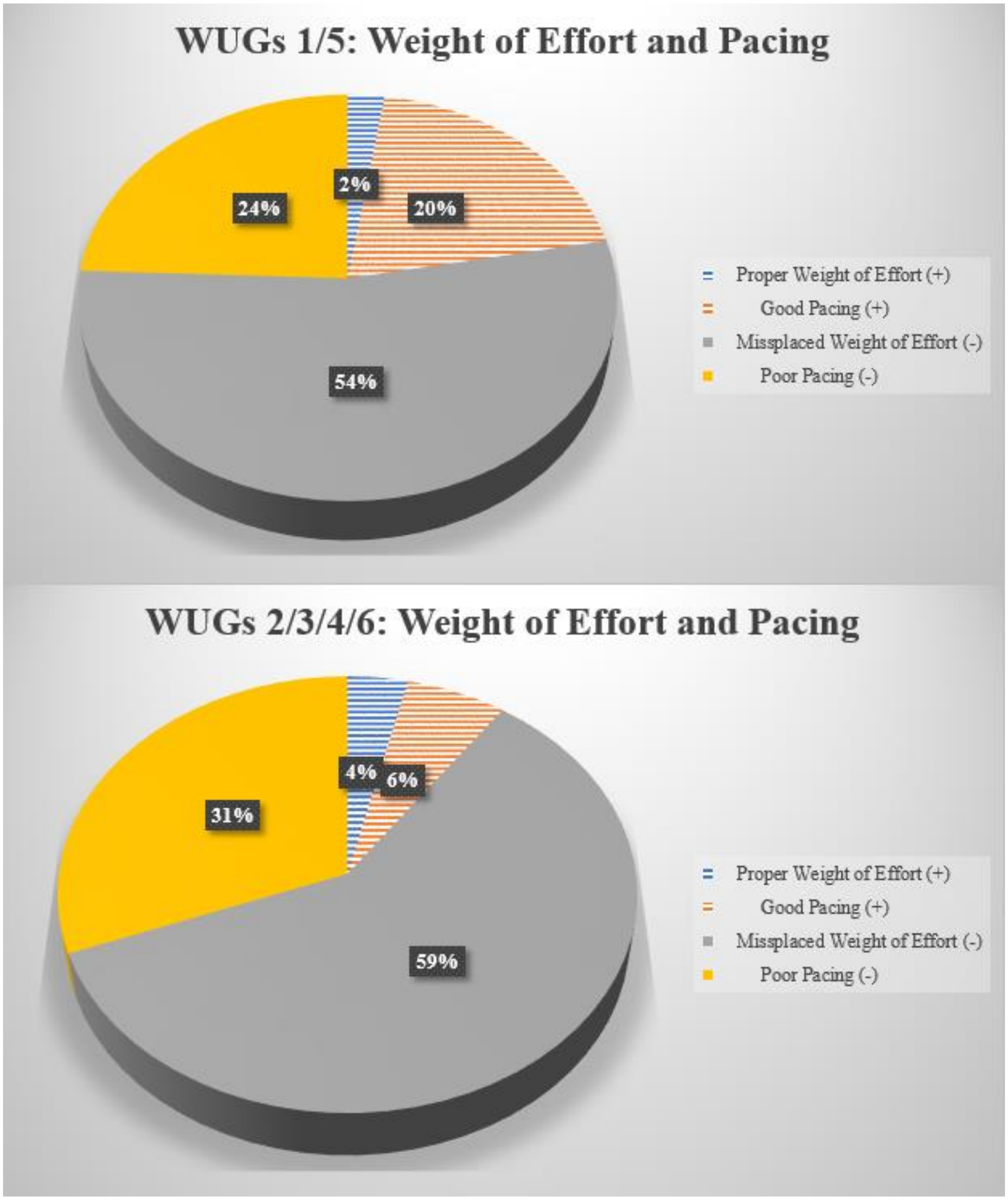


Figure 11: Students 1 and 5 vs. Students 2, 3, 4, and 6 in Weight of Effort/Pacing Category

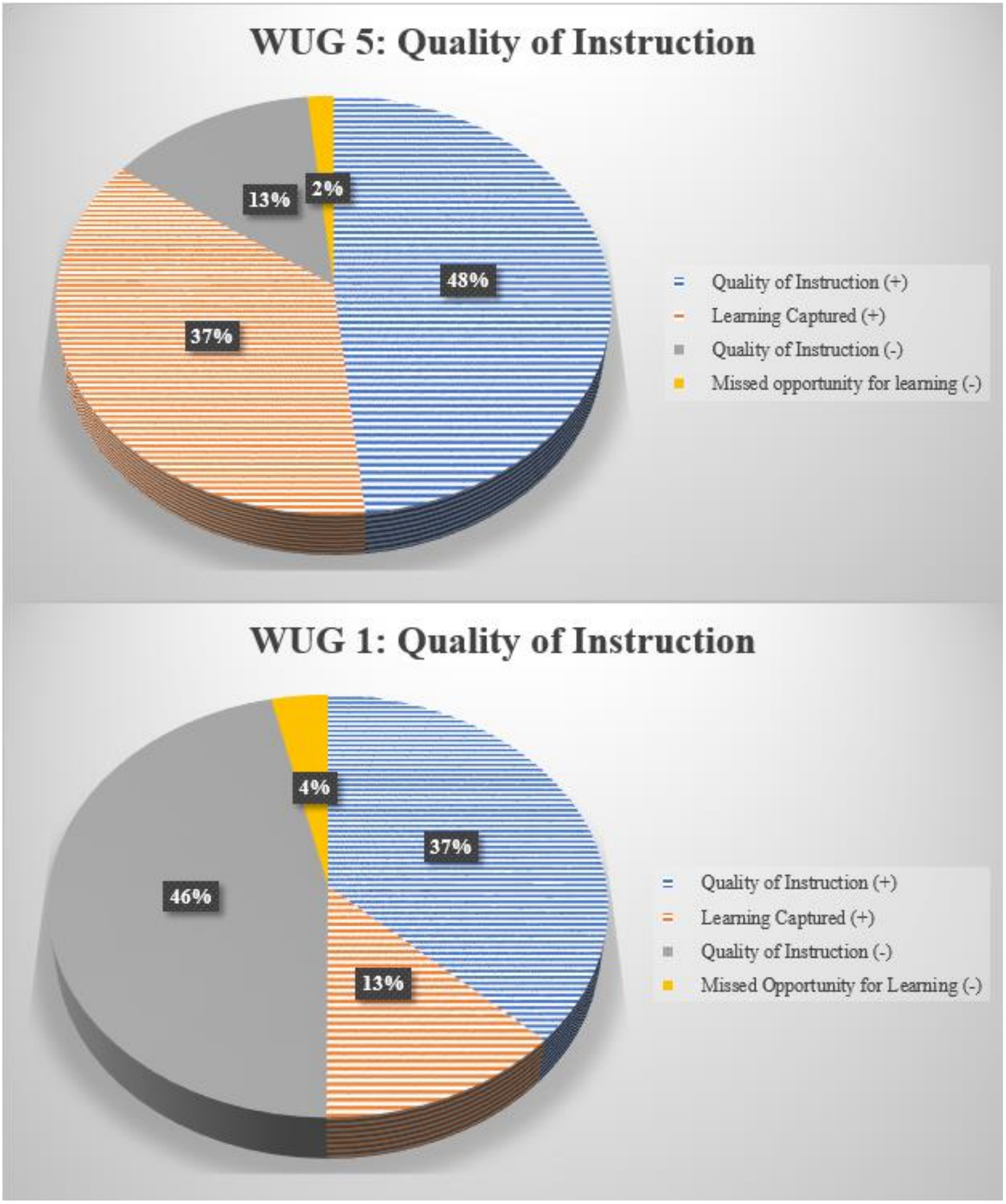


Figure 12: Student 5 vs. Student 1 in Quality of Instruction Category

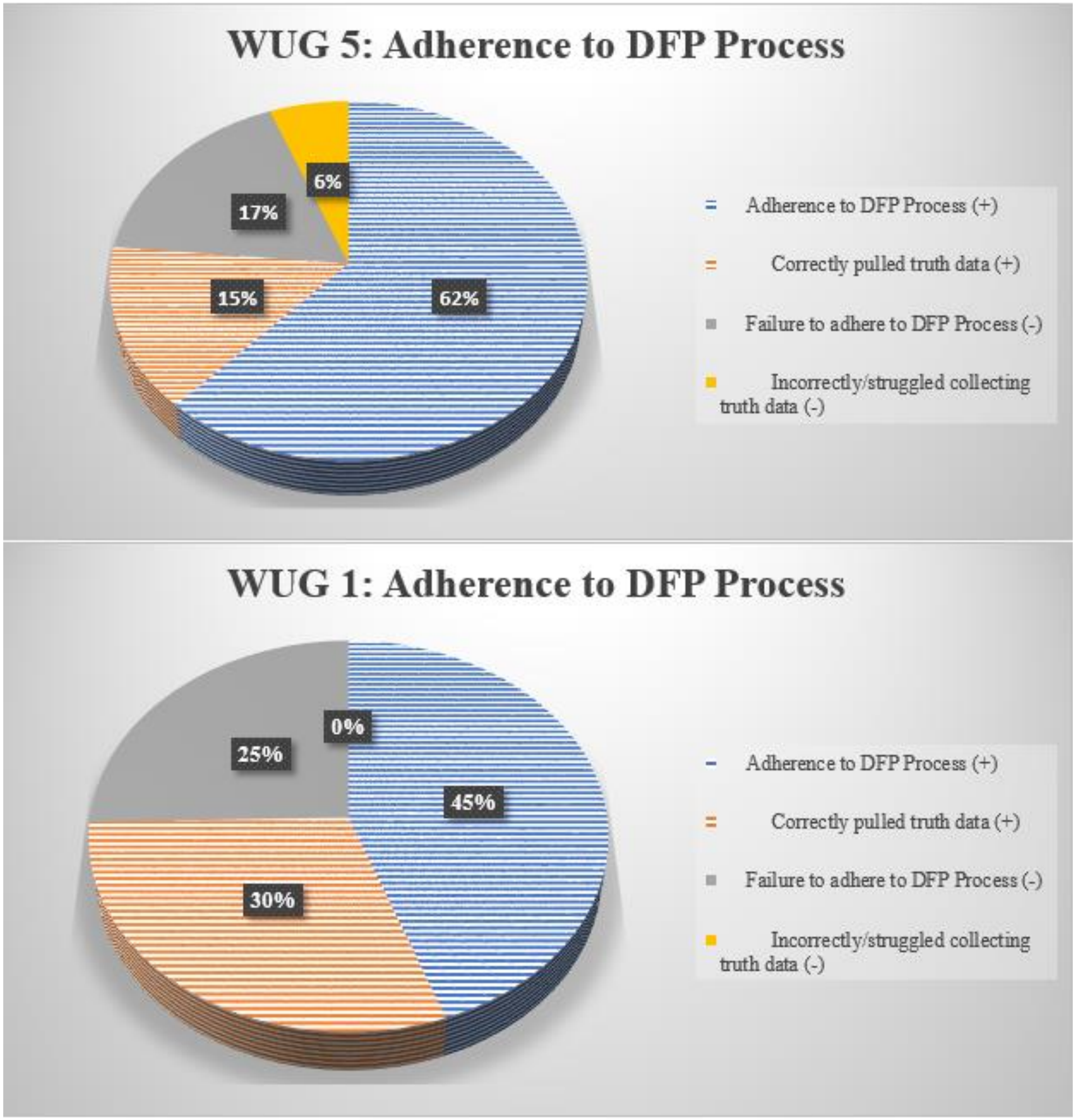


Figure 13: Student 5 vs. Student 1 in Adherence to DFP Process Category

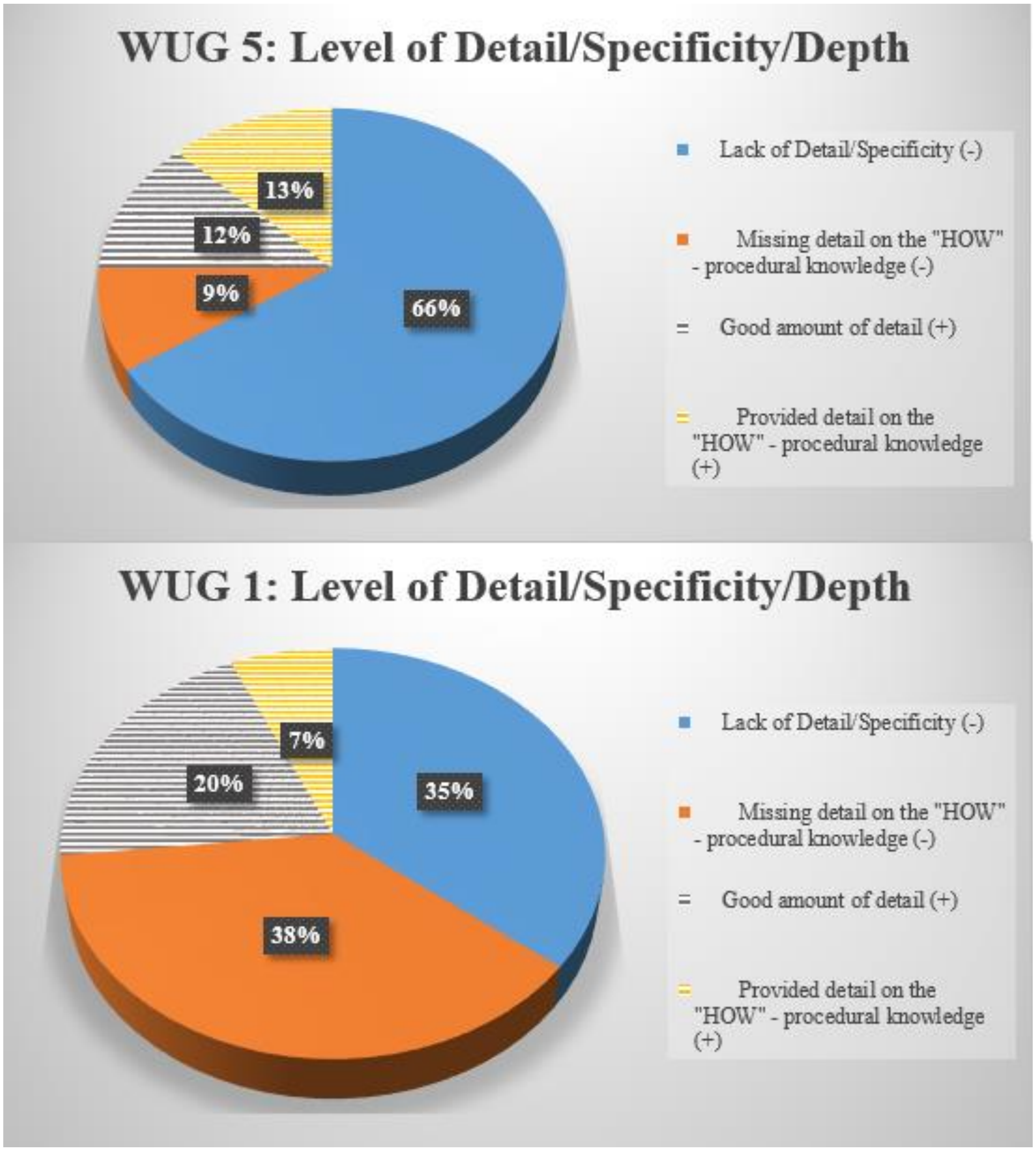


Figure 14: Student 5 vs. Student 1 in Level of Detail Category

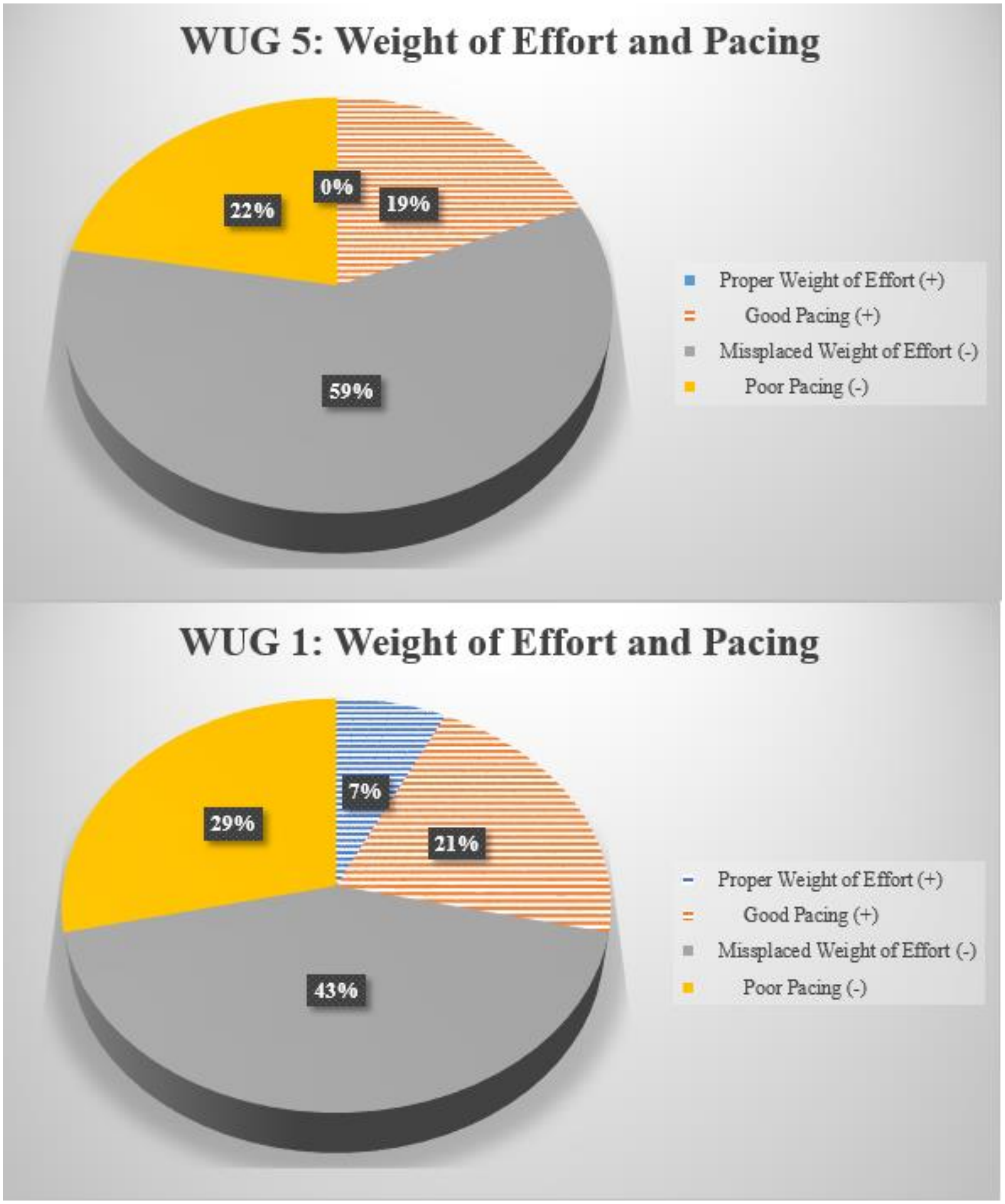


Figure 15: Student 5 vs. Student 1 in Weight of Effort and Pacing Category

References

Airman Magazine (2014). America's Air Force: A call to the future. Retrieved from https://airman.dodlive.mil/files/2014/07

Barron, L. G., Carretta, T.R., & Rose, M.R. (2016). Aptitude and Trait Predictors of Manned and Unmanned Pilot Job Performance. *Military Psychology*, 28(2), 65.

Butler, H. A. (2012). Halpern Critical Thinking Assessment Predicts Real-World Outcomes of Critical Thinking. *Applied Cognitive Psychology*, 26(5), 721-729.

Carretta, T. R. (2013). Predictive validity of pilot selection instruments for remotely piloted aircraft training outcome. Aviation, space, and environmental medicine, 84(1), 47-53.

Casey, S. M. (1998). Set phasers on stun: And other true tales of design, technology, and human error. Santa Barbara: Aegean.

Chappelle, W., McDonald, K., & McMillan, K. (2011). Important and critical psychological attributes of USAF MQ-1 Predator and MQ-9 Reaper pilots according to subject matter experts (No. AFRL-SA-WP-TR-2011-0002). SCHOOL OF AEROSPACE MEDICINE WRIGHT PATTERSON AFB OH AEROSPACE MEDICINE DEPT/AEROSPACE MEDICINE CONSULTATION DIV.

Chappelle, W., Swearengen, J., Goodman, T., & Thompson, W. (2014). Personality Test Scores that Distinguish US Air Force Remotely Piloted Aircraft Drone Pilot Training Candidates (No. AFRL-SA-WP-TR-2014-0001). SCHOOL OF AEROSPACE MEDICINE WRIGHT PATTERSON AFB OH AEROSPACE MEDICINE DEPT.

Chi, M. T., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. Educational Psychologist, 47(3), 177-188.

Clark, R. E., Feldon, D., van Merriënboer, J. J., Yates, K., & Early, S. (2008). Cognitive task analysis. Handbook of research on educational communications and technology, 3, 577-593.

Cooke, N. J. (1992). Modeling human expertise in expert systems (pp. 29-60). Springer New York.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. International Journal of Human-Computer Studies, 41(6), 801-849.

Damos, D.L. (2011). A summary of the technical pilot selection literature, (Report No. AFCAPS-FR-2011-0009). Randolph AFB, TX: Air Force Personnel Center, Strategic Research and Assessment Branch.

Davis, R. A., Flett, G. L., & Besser, A. (2002). Validation of a new scale for measuring problematic Internet use: Implications for pre-employment screening. CyberPsychology & Behavior, 5(4), 331-345.

Everstine, B. (2015, August). DOD Plans 50 percent increase in RPA CAPS by 2019. *Air Force Magazine.* Retrieved from http://www.airforcemag.com/DRArchive/Pages/2015/August%202015/August%2018%202015/DOD-Plans-50-Percent-Increase-in-RPA-CAPs-by-2019-.aspx

Gagne, R. M. (1984). Learning outcomes and their effects: Useful categories of human performance. American Psychologist, 39(4), 377.

Halpern, D. F. (2010). Halpern Critical Thinking Assessment. Schuhfried (Vienna Test System). http://www.schuhfried.at

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449.

Howse, W.R. (2011). Knowledge, skills, abilities, and other characteristics for remotely piloted aircraft pilots and operators. (AFCAPS-FR-2011-0006). Randolph AFB, TX: HQ AFPC/DSYX Strategic Research and Assessment Branch.

Laker, D. R., & Powell, J. L. (2011). The differences between hard and soft skills and their relative impact on training transfer. Human Resource Development Quarterly, 22(1), 111-122.

Losey, S. (2015, December). To meet anti-terror demand, Air Force offers $125K bonuses to keep drone pilots. Air Force Times. Retrieved from http://www.airforcetimes.com/story/military/benefits/pay/2015/12/15/rpa-pilots-now-eligible-for-125k-retention-bonus/77370056/

Manley, G. G. (2011). Development of domain and facet level scales for the Self-Description Inventory (Rep. No. AFCAPS-TR-2011-0007). Randolph AFB, TX: HQ AFPC/DSYX Strategic Research and Assessment Branch.

Martin, A. J., Nejad, H. G., Colmar, S., & Liem, G. A. D. (2013). Adaptability: How students' responses to uncertainty and novelty predict their academic and non-academic outcomes. Journal of Educational Psychology, 105(3), 728.

Paullin, C., Ingerick, M., Trippe, D. M., & Wasko, L. (2011). Identifying Best Bet Entry-Level Selection Measures for US Air Force Remotely Piloted Aircraft (RPA) Pilot and Sensor Operator (SO) Occupations (No. HRRO-FR-11-64). HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA.

Pavlas, D., Burke, C. S., Fiore, S. M., Salas, E., Jensen, R., & Fu, D. (2009). Enhancing unmanned aerial system training: A taxonomy of knowledge, skills, attitudes, and methods. Human Factors and Ergonomics Society Annual Meeting Proceedings, 53(26), 1903-1903. doi:10.1518/107118109X12524444083159

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P., & Fleishman, E. A. (1999). An occupational information system for the 21st century: The development of O* NET. American Psychological Association.

Rosales, J.R. (2006). A Legacy of Excellence: *The USAF Weapons School's Challenge to Maintain Standards.* (Doctoral dissertation, The Naval Postgraduate School).

Rose, M. R., Arnold, R. D., & Howse, W. R. (2013). Unmanned aircraft systems selection practices: Current research and future directions. Military Psychology, 25(5), 413.

Rose, M. R., Barron, L. G., Carretta, T. R., Arnold, R. D., & Howse, W. R. (2014). Early identification of unmanned aircraft pilots using measures of personality and aptitude. The International Journal of Aviation Psychology, 24(1), 36-52. doi:10.1080/10508414.2014.860849

Schogol, J. (2015, February). Air Force increases combat air patrols for Reaper pilots. Air Force Times. Retrieved from http://www.airforcetimes.com/story/military/2015/02/02/more-missions-for-reaper-pilots/22752129/

Schwartz, D. L., Chase, C. C., & Bransford, J. D. (2012). Resisting overzealous transfer: Coordinating previously successful routines with needs for new learning. *Educational Psychologist*, *47*(3), 204-214.

Sternberg, R. J., Roediger III, H. L., & Halpern, D. F. (2007). Critical thinking in psychology. Cambridge University Press.

Strauss, K. (2016). 9 Skills Bosses See Lacking in New Graduates in 2016. *Forbes Magazine.* Retrieved from http://www.forbes.com/sites/karstenstrauss/2016/05/17/these-are-the-skills-bosses-say-new-college-grads-do-not-have/#1cf74333596e

Tirpak, J., & Deptula, G. D. A. (2010). The RPA Boom. *Air Force Magazine*, 93(8), 36-42.

Tofel-Grehl, C., & Feldon, D. F. (2013). Cognitive task analysis–based training a meta-analysis of studies. Journal of Cognitive Engineering and Decision Making, 7(3), 293-304.

Tvaryanas, A. P. (2006). Human systems integration in remotely piloted aircraft operations. Aviation, space, and environmental medicine, 77(12), 1278-1282.

United States Air Force (2015, May). Air Force moves to bring about RPA mission relief. Retrieved from http://www.af.mil/News/ArticleDisplay/tabid/223/Article/589196/air-force-moves-to-bring-about-rpa-mission-relief.aspx

United States Air Force (2015, August). United States Air Force Weapons School. Retrieved from http://www.nellis.af.mil/library/factsheets/factsheet.asp?id=19837

United States Government Accounting Office (2017). Air Force and Army Should Improve Strategic Human Capital Planning for Pilot Workforces. Retrieved from http://www.gao.gov/assets/690/682413.pdf

Walker L. (2015). 90 Missions a Day: Pentagon Wants to Up Its Use of Drones. Newsweek. Retrieved from http://www.newsweek.com/90-missions-day-pentagon-wants-its-use-drones-363428

Williams, H. P., Carretta, T. R., Kirkendall, C. D., Barron, L. G., Stewart, J. E., & Rose, M. R. (2014). Selection of UAS Personnel (SUPer) Phase I Report: Identification of Critical Skills, Abilities, and Other Characteristics and Recommendations for Test Battery Development (No. NAMRU-D-15-16). NAVAL MEDICAL RESEARCH UNIT (DAYTON) WRIGHT-PATTERSON AFB OH.

Wong, C. A., & Cummings, G. G. (2007). The relationship between nursing leadership and patient outcomes: a systematic review. Journal of nursing management, 15(5), 508-521.

APPENDIX A: KNOWLEDGE, SKILLS, AND ABILITIES

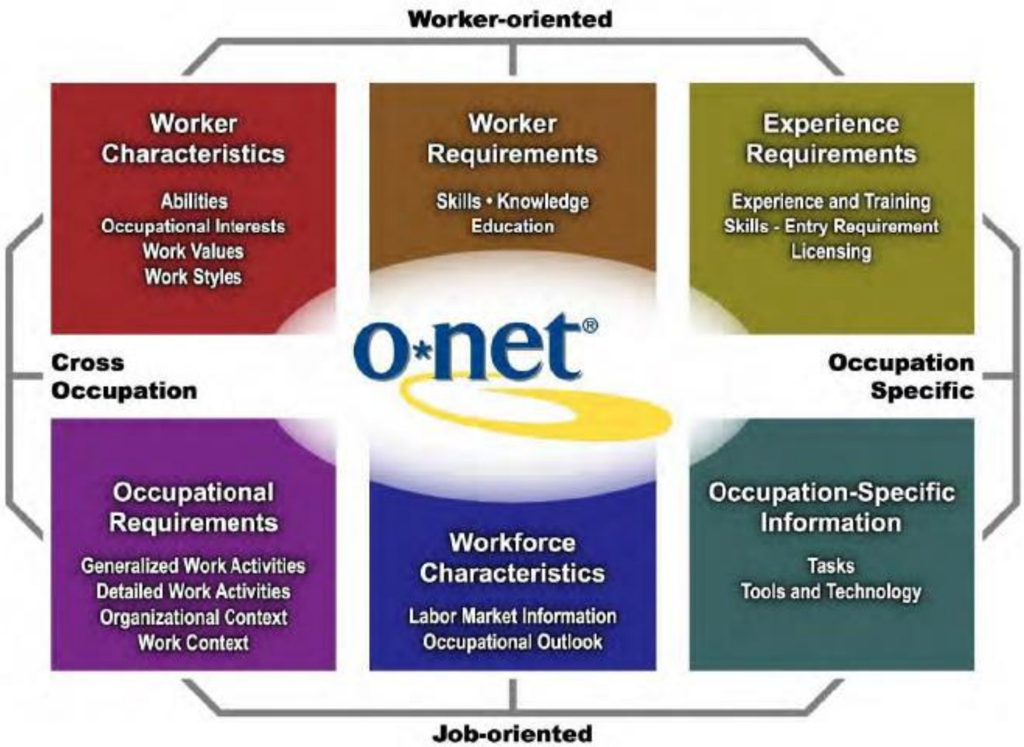
Knowledge	**Skill**
Human-Focused	Flight skill
Culture	Long-term monitoring
Contingency behavior	Target search
Workload	Instrument monitoring
Fatigue	Mission monitoring
Information overload	Navigation
Distraction	Team leadership
Mission awareness	Delegation
Human locations	Team composition
Human activities	Mission planning
Human identities	Plan understanding
Human characteristics	Replanning
Human intentions	Stress management
Human activity dependencies	Decision making
Larger mission	Problem solving
Commander's/mission intent	Risk assessment
Shared situation awareness	Visual scanning strategy
Equipment-Focused	Handoff
System feedback	Communication
Operational threats	Inter-team communication
Command set	Information flow
Constraints	Collaboration
Automation reliability	Coordination
4D Spatial Relationships	Teamwork skills
Performance envelope	**Attitude**
Level of automation	Risk perception
Latency	Risk taking behavior
Equipment knowledge	"Kill Chain" stress
Task knowledge	Complacency
Shared situation awareness	Overtrust
	Undertrust
	Calibrated trust
	High value opportunity

Table 1: UAS-specific KSAs.

General Methods	**Knowledge Methods**
Event-Based Training	Cross-Training
Scenario-Based Training	Multicultural Training
Self-Correction Training	**Skill Methods**
Guided Self-Correction Training	Team Coordination Training
On-the-job Training	Team Leader Training
Role Play	Virtual Team Communication Training
Simulations and Games	Behavior Modeling
Automation-enabled full flight simulation	Stress Exposure Training
UAS Simulation	Stress Inoculation
Embedded Instructional Agent	**Attitude Methods**
Simulation-based handoff scenario	Trust Tuning

Table 2: Identified training methods.

APPENDIX B: DEPARTMENT OF LABOR'S O*NET FRAMEWORK


Worker-oriented
Worker Characteristics
Abilities
Occupational Interests
Work Values
Work Styles
Worker Requirements
Skills • Knowledge
Education
Experience Requirements
Experience and Training
Skills - Entry Requirement
Licensing
Cross Occupation
o*net®
Occupation Specific
Occupational Requirements
Generalized Work Activities
Detailed Work Activities
Organizational Context
Work Context
Workforce Characteristics
Labor Market Information
Occupational Outlook
Occupation-Specific Information
Tasks
Tools and Technology
Job-oriented

APPENDIX C: CONSENT FORM

CONSENT FORM

TOWARDS DEVELOPING A PREDICTIVE MODEL FOR REMOTELY PILOTED AIRCRAFT TRAINING COMPLETION

INTRODUCTION

The purposes of this form are to provide you (as a prospective research study participant) information that may affect your decision as to whether or not to participate in this research and to record the consent of those who agree to be involved in the study.

RESEARCHERS

Nancy J. Cooke, Professor, Arizona State University, Nathan J. McNeese, Post-Doctoral Researcher, Arizona State University, Jade E. Best, MS Student, Arizona State University

STUDY PURPOSE

The purpose of the research is to determine if there is a relationship between United States Air Force (USAF) Remotely Piloted Aircraft (RPA) Weapons School Completion and critical thinking skills.

DESCRIPTION OF RESEARCH STUDY

If you decide to participate, you will join a study which will provide valuable insight on the potential correlation between critical thinking skills, measured by the Halpern Critical Thinking Assessment and performance in USAF MQ-9 Weapons School. The Halpern Critical Thinking Assessment is a 25 question multiple choice cognitive assessment. It measures critical thinking skills in five different domains: verbal reasoning, argument analysis, thinking as hypothesis testing, likelihood and uncertainty, and decision making and problem solving. You will take the assessment at the beginning of your training program (no later than 1 March 2016), and again at the end of your training program (estimated 1 June 2016). You will have 20 minutes to take the assessment each time. Your participation is completely voluntary and you may cease participation at any time. Should you choose to participate, training completion/failure will be tracked. We will be using a master list to link pre/post training test measures and training completion/failure rates.

RISKS
There are no known risks from taking part in this study, but as in any research, there is some
possibility that you may be subject to risks that have not yet been identified. Your participation in
this research study will be confidential. Your participation in this study will **not** impact your
standing with United States Air Force (USAF).

BENEFITS
This research will have implications for developing a comprehensive, predictive test battery for
screening potential United States Air Force RPA Pilots. You will learn if there is a relationship
between critical thinking skills and performance in MQ-9 Weapons School.

CONFIDENTIALITY
All information obtained in this study is strictly confidential. Weapons School instructors will
know who participated in the study, but will not know the identification numbers assigned to
students, or have access to individual students' Halpern Critical Thinking Assessment scores.
The results of this research study may be used in reports, presentations, and publications, but
the researchers will not identify you. In order to maintain confidentiality of your records, Dr.
Nancy J. Cooke will follow these procedures: (1) Each participant will be assigned a number; (2)
The researchers will record any data collected during the study by number, not by name; (3) Any original data files (to include the master list), will be stored on a hard drive secured in the Cognitive Engineering Research on Team Tasks (CERTT) lab until completion of thesis defense (1 Feb 2017) accessed only by authorized researchers; (4) the master list will be kept in a secure, separate location from the rest of the study data and will be destroyed upon data analysis; (5) consent forms will not link names to ID numbers. Consent forms will also be secured in a separate file, maintained in the CERTT lab, until completion of thesis defense (1 Feb 2017).

WITHDRAWAL PRIVILEGE
Participation in this study is completely voluntary. It is ok for you to say no. Even if you say yes
now, you are free to say no later, and withdraw from the study at any time. Your participation is
voluntary and that nonparticipation or withdrawal from the study will not affect your status in
class.

VOLUNTARY CONSENT

Any questions you have concerning the research study or your participation in the study, before
or after your consent, will be answered by Nancy J. Cooke at ASU Polytechnic, 480-727-2418.
If you have questions about your rights as a subject/participant in this research, or if you feel
you have been placed at risk; you can contact the Chair of the Human Subjects Institutional
Review Board, through the ASU Office of Research Integrity and Assurance, at 480-965 6788.
This form explains the nature, demands, benefits and any risk of the project. By signing this
form you agree knowingly to assume any risks involved. Remember, your participation is voluntary. You may choose not to participate or to withdraw your consent and discontinue
participation at any time without penalty or loss of benefit. In signing this consent form, you are
not waiving any legal claims, rights, or remedies. A copy of this consent form will be given
(offered) to you.

Your signature below indicates that you consent to participate in the above study.

______________________	______________________	__________
Subject's Signature	Printed Name	Date

INVESTIGATOR'S STATEMENT

"I certify that I have explained to the above individual the nature and purpose, the potential
benefits and possible risks associated with participation in this research study, have answered
any questions that have been raised, and have witnessed the above signature. These elements
of Informed Consent conform to the Assurance given by the Arizona State University
Institutional Review Board (IRB) to protect the rights of human subjects. I have provided
(offered) the subject/participant a copy of this signed consent document."

Signature of Investigator______________________________________
Date_____________

APPENDIX D: RECRUITMENT SCRIPT

RECRUITMENT SCRIPT

I am a graduate student under the direction of Dr. Nancy Cooke in the Human Systems Engineering Department at Arizona State University. I am conducting a research study to determine if there is a relationship between critical thinking skills, measured by the Halpern Critical Thinking Assessment, and United States Air Force MQ-9 Weapons School completion/failure.

I am recruiting individuals to complete a pre-training (no later than 1 March 2016) and post-training (estimated 1 June 2016) Halpern Critical Thinking Assessment, which will take approximately 40 minutes total. The assessment can be accessed online, and consists of 25 multiple choice questions (per test administration). I will be linking Halpern Critical Thinking Assessment pre-training and post-training scores, with student Weapons School completion/failure.

Your participation in this study is voluntary. If you have any questions concerning the research study, please call me at (916) 969-6713 or email me at jade.best.1@us.af.mil.

Your help is greatly appreciated,

Jade Best

APPENDIX E: HALPERN CRITICAL THINKING ASSESSMENT

Halpern Critical Thinking Assessment

1. 25 question, multiple choice exam administered online, by email invitation only.

2. Assessment will measure critical thinking in five domains: verbal reasoning, argument analysis, thinking as hypothesis testing, likelihood & uncertainty, and decision making & problem solving.

3. Sample test questions are **not** included for proprietary reasons. Test is only available through purchase from Schuhfried Publishing.

APPENDIX F: DEBRIEFING SCRIPT

<u>Debriefing</u>

Thank you for your participation in our study.

The study you have completed was to help us to develop a predictive model for successful RPA mission performance, by conducting a correlational analysis between Halpern Critical Thinking Assessment (HCTA) scores and USAF MQ-9 Weapons School completion. RPA are increasing in use in both commercial and military applications, so therefore it is important to ensure that the most qualified individuals are charged with their operation.

As this is ongoing research, please do not discuss this information with anyone that will be participating. You are free to discuss this study with anyone that will not be participating.

Thank you again. Please ask if you have any additional questions.

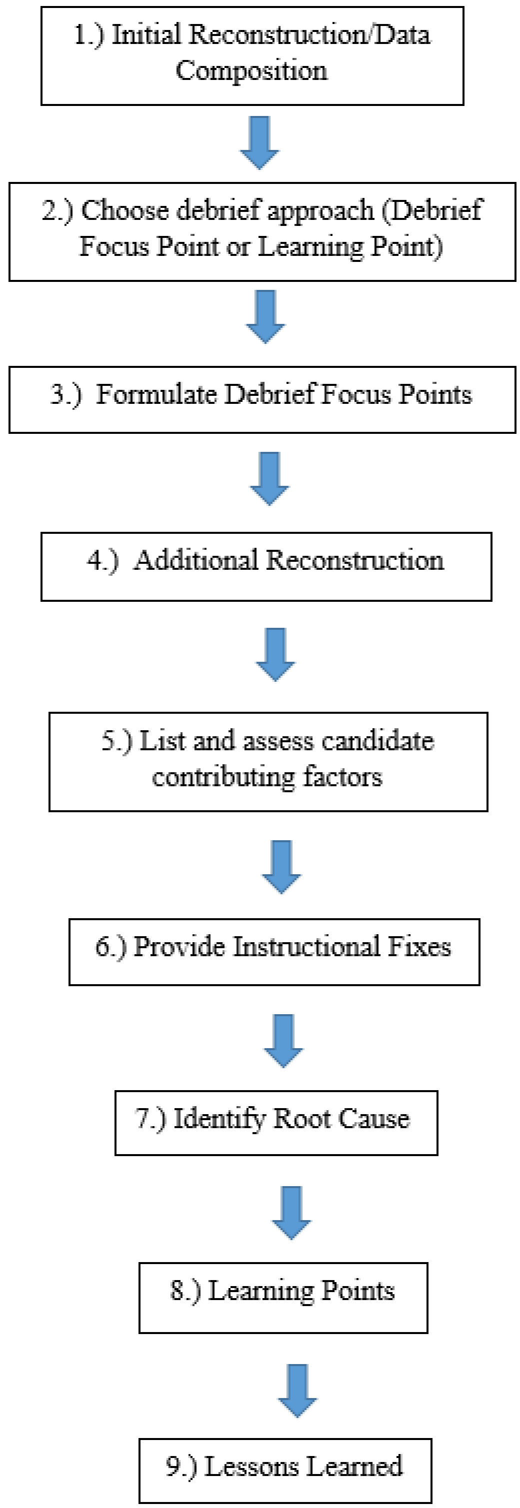
Sincerely,

Nancy J. Cooke

Professor, ASU

APPENDIX G: WEAPONS SCHOOL DEBRIEF FLOW

Weapons School Debrief Flow



APPENDIX H: INSTRUCTOR EXPLANATION ON STUDENT PERFORMANCE AND WEIGHT OF EFFORT IN GRADING

Question 8: How much of your assessment is based on the student performance in the mission versus their ability to reconstruct the mission accurately (ex: 50/50, 60/40) and effectively debrief?

- **IP 1:** Student will FAIL a ride based on a debrief before failing a ride based on actual flying performance
- **IPs 1/2:** Earlier phases about skills (70 mission exec/30 debrief), as students progress to later phases (50 mission exec/50 debrief). What should remain consistent throughout – PROFESSIONALISM, SOLID PRESENCE, ARTICULATE. Cognizant of other people in room and their potential debrief inputs.
- **IP 3:** It shifts depending on the phase a student is in the program. Early on, more even (50/50) debrief/mission focus for evaluation. In later phases, this shifts to more of a (90/10) debrief/mission focus for the evaluation. A typical for a student to not have a lot to talk about in debrief because they flew well. Level of detail; can't debrief out of, but if you can identify learning – you will likely pass the ride.
- **IPs 4/5** There is a minimum performance that students have to hit, assuming basics are met, a lot of stuff is savable in the debrief. Actually, a MAJORITY is savable. Looking for – what can you get good learning out of and here's what I SHOULD'VE done.
- **IP 4:** There is a minimum performance that students have to hit, assuming basics are met, a lot of stuff is saveable in the debrief. Actually, a MAJORITY is saveable. Looking for – what can you get good learning out of and here's what I SHOULD'VE done.
- **IPs 6/7:** "3 Round Fight" analogy. Brief/Execution/Debrief. Student can mess up a ton, but if they can nail the debrief, this is good. Since this is an advanced instructor course, the debrief portion is the most important and the heaviest weighted portion.

APPENDIX I: SAMPLE CODED DATA FROM WEAPONS SCHOOL STUDENT GRADE SHEET

AI-3(2): Sortie Accomplished IAW 26 WPS MQ-1/MQ-9 WIC Syllabus

AI-3(2): (2) indicates that this is the second time the student has attempted this mission

Overall: Average

Specifics: Briefing; Execution; Debriefing

Briefing: Slightly Below Average. Brief lacked coherent flow through motherhood/tac admin. WUG found himself talking about FENCE tasks in the first phase of SCAR. He did not fully describe how he would battle track with his decon sheet. WUG had no macro level view of the AO on his board so it was difficult to follow where the aircraft was going throughout the sortie. The coordinated attack example lacked coordination and was just 2 aircraft operating in close proximity to each other and lacked triggers for when to initiate a coordinated attack.

Execution: Slightly Above Average. Some battle tracking errors, double tapped some targets, and never elevated a TBM found with CDE concerns. Overall no major deconfliction errors, but there were some inefficiencies as the SCAR

Debriefing: Average. Debrief captured all the truth data but WUG ended up fishing for a DFP since there was no major mission failure. He found the appropriate learning but his debrief lacked efficiency. WUG was debriefed on how to frame DFPs in situations where there was not a mission failure. WUG was also debriefed on how to ask the next level why questions to ID all the mistakes and find the true root cause of the DFP in order to provide an appropriate level of instruction.

Recommendation: Press to AI-4

Press to AI-4 indicates that the student has passed the AI-3 mission and can proceed w/the syllabus

APPENDIX J: AVERAGED FREQUENCY COUNTS FOR ALL SIX STUDENTS

Themes from student gradesheets:	WUG-1 (27 msns)	WUG-2 (30 msns)	WUG-3 (29 msns)	WUG-4 (32 msns)	WUG-5 (29 msns)	WUG-6 (28 msns)
Adherence to DFP Process (reconstruction, creation of DFP, CF, IF, RC)						
Adherence to Weapons School Process (+)("good DFP/CF")	9.5	6	7	7.5	10.5	7.5
Correctly pulled truth data (+)	2.5	0	1.5	1	2.5	0
Failure to adhere to Weapons School Process (-) ("poor DFP/CF development")	3.5	15.5	15.5	10.5	3	11
Incorrectly/struggled collecting truth data ("relied on memory," "injected opinion") (-)	0	2.5	1	2	1	1
Flow						
Good Flow (+)	6.5	2	1	2.5	4.5	2
Poor Flow (-) ("disjointed, tough to follow, incoherent, not logical")	7	6	5.5	2.5	2.5	3
Weight of Effort & Pacing						
Proper Weight of Effort (+)	0.5	0.5	0	0	0	0.5
Good Pacing (+)	1.5	1	1	0	2.5	0.5
Missplaced Weight of Effort (-)	3	6.5	11	6	8	2.5
Poor Pacing (-) ("brief went long," "pacing was off")	2	2.5	8	2.5	3	1.5
Level of Detail/Specificity/Depth						
Lack of Detail/Specificity (-)	8	9	13	13	10.5	4.5
Missing detail on the "HOW" - procedural knowledge (-)	8.5	6	1.5	3.5	1.5	8.5
Good amount of detail (+)	4.5	1	0.5	2.5	2	2.5
Provided detail on the "HOW" - procedural knowledge (+)	1.5	0.5	1	1.5	2	1
Instructor Pilot (IP) Assistance						
IP inject/assistance/prodding required (-)	9	6.5	8.5	8.5	6	7
Minimal IP inject required (+)	0	0	0	0	1.5	0
Efficiency						
Efficient (+)	2	1	1.5	0	3	0
Inefficient (-) ("led to one DFP taking over 3 hours")	1	2	3	1	0	1
Organization						
Organized (+)	0	0	0	0	0	0
Lack of Organization (-)	0.5	0.5	0	0	0	0
Preparation/Planning						
Adequately Prepared (+) ("well thought out")	0	0	1.5	1.5	0	0
Adequate Planning (+) (briefed contingencies, "well thought out")	0	0	0.5	2	0	1
Lack of preparation, noticeably unprepared (-) ("examples were not rehearsed and smooth")	1.5	1	1	2.5	0	1.5
Inadequate Planning (-) (didn't brief contingencies, provided little direction for #2 (Wingman))	3	4	1.5	3	1	2
Repeatable Process/Methodology						
Use of Repeatable Process/Methodology (+)	2.5	1.5	2.5	2.5	2.5	2
Lack of Methodology/Repeatable Process (-)	2	4.5	2.5	6	1.5	5
Decision Criteria						
Presented Decision Criteria (+)	1	0.5	1	0.5	1	1.5
Omitted/Presented Incorrect Decision Criteria (-)	4	3	4	2	2	3
Triggers						
Presented Triggers (+)	1	1	1	2	0	1.5
Omitted/Presented Incorrect Triggers (-)	1.5	0	2.5	3.5	3.5	0
Contracts						
Correct Usage (+)	2.5	0	2	0	1	0.5
Incorrect Usage (or lack of use) (-)	2.5	0	3	1	3	3
Instruction (quality, was the learning captured?)						
Quality of Instruction (+) "provided good instructional fixes (IF)"	10	8.5	6.5	7	14.5	5.5
Quality of Instruction (-) "tended to be academic at times" "instructional fixes (IF) poor, rushed, "diffic	12.5	20.5	12.5	22	4	15.5
Learning Captured (+)	3.5	2	3.5	3.5	11	3.5
Missed opportunity for learning (-) ("left some learning on the table")	1	1	5	3	0.5	2
Artifact/Tool Usage (e.g. checklists, RMIT)						
Appropriate Artifact Usage (+)	0.5	0.5	0	0	0.5	0
Inappropriate or (lack of) Artifact Usage (-)("understanding 3-3 manual" "IAW 3-1 Shot Kill (SK)))	1	1	3	0	2	0
Transfer						
Mission to Mission Transfer ("improved from last mission, sortie" "no issues carried over") (+)	1.5	2.5	3	2.5	1.5	3.5
Mission to Mission Transfer ("made same mistakes in last mission, sortie") (-)	0.5	1.5	2	1	2	0.5
Intramission Transfer (showed improvement within same mission) (+)	0.5	1	2	0	1.5	1
Intramission Transfer (repeated same mistakes within the same mission) (-)	0	0	0	0	0	0
Transfer occuring over the course of the Program (across phases) (+)	0.5	1.5	0	0.5	0	0
Transfer occuring over the course of the Program (across phases) (-)	1	2	4	2	0.5	0.5
Identification of Errors						
Correct Identification of errors/flaws in their own plan or execution in mission (+)	3	1	3.5	3.5	5	4
Incorrect (or unable to ID) Identification of errors/flaws in their own plan or execution in mission (-)	1.5	0.5	1.5	0.5	0	1.5
Collaboration ("unwilling to accept..")						
Utilization or collaboration with teammates during debrief (reconstruction process)/mission (+)	1.5	0	0.5	0.5	3	1
Neglect (failure to collaborate) with teammates/underutilization of teammates (-)	2	1.5	2	0	1	0
Weapons Officer Quality						
Presence (+)	4.5	0.5	2.5	2.5	2.5	0.5
Presence (-) ("watch monotonous tone")	0	1.5	0.5	0	0	0.5
Conduct (+)	0	0	0	0	0	1
Conduct (-)	0	0	3	2	0	0
Valuable contributor during mass debrief (+)						
Valuable Contributor (+)	4	3.5	2	3	3.5	4
Not Valuable (-)	0	1	0	0	0	0
Adaptability (response in new or unexpected situation)						
Adapted well to novel situation (+) ("first time in mass debrief," "first attempt at SCAR")	1.5	3.5	1.5	2.5	1.5	3
Did not adapt well to novel situation, new problem set (-)("obviously uncormfortable with the materia	0	0	0	2	0	1
Flight Leadership						
Demonstrated Flight Leadership (+)	2	0.5	1	0.5	1.5	0
Failed to demonstrate Flight Leadership (-)	1	4.5	2.5	2	1	0

APPENDIX K: GRADE SHEET CODING METHODOLOGY

Adherence to Debrief Focus Point Process. The adherence to debrief focus point process category was based on whether students appropriately executed the process instructed at the beginning of the Weapons School course – reconstruction/data composition (specifically the utilization of truth data), selection of debrief approach (debrief focus point vs. learning point), formulation of debrief focus point, additional reconstruction (if required), listing of contributing factors, providing instructional fixes, and identifying the root cause for the mission error/failure. The collection of truth data was treated as a subcategory within the adherence to DFP process category. Students' ability to adhere to the Debrief Focus Point process was coded as a binary outcome – they either adhered to the process, or they failed to adhere to the process. Additionally, students' ability to pull truth data was treated as a binary outcome – students were able to pull truth data, or they were unable to/struggled pulling truth data. Example instructor comments indicating students' adhering/failing to adhere to the process and capturing/struggling to capture truth data are provided in the table below:

Table 22

Adherence to DFP Process comment examples

Adhering to the DFP Process	**Failure to adhere to the DFP Process**
"student did a decent job reconstructing with the SIM tools"	"reconstruction was not timely and was incomplete"
"correctly identified the DFPs and CFs"	"failed to continue to task the why question to get to the root of many of the contributing factors"
"captured all data"	"started with a DFP question and began putting up CFs without any reconstruction"
"able to identify the DFP question and reconstruct the sortie in accordance with (IAW) phase standards"	"tends to narrowly focus on CFs that were endgame/near endgame errors"
"DFP addressed correct issue and student had correct CFs"	"some of the instructional fixes were "wave top" level"

Flow. The flow category was based on whether the student briefed and debriefed with a logical flow to their ideas. Good flow was characterized by the presentation of ideas in a logical manner (e.g. start to finish, finish to end timeline of events for a mission), and bad flow resulted in the presentation of ideas in a disjointed, difficult to follow manner. Flow was assessed as a binary outcome – students either practiced good flow or poor flow in their briefs and debriefs. Examples of good and poor flow noted in student grade sheets are highlighted in the below table:

Table 23

Flow Comment examples

Good Flow	**Poor Flow**
"brief was broken into phases and flowed logically"	"flow of brief was a little disjointed as he talked about each item"
"WUG had reflowed his brief and improved from last sortie"	"FRA brief was disjointed and tough to follow"
"had good process/flow for running debrief"	"lacked any real flow as to how the processes were going to be repeated"
"WUG did a pretty good job with building a logical flow to his DFP"	"flow to debrief started out a little off"
"WUG did a nice job flowing CF's via reverse chronological method"	"brief did not have a coherent flow"

Weight of Effort and Pacing. The weight of effort category was based on whether students practiced appropriate weight of effort on varying topics in their briefs and debriefs. This category was evaluated objectively and assessed how students managed the 55-minute brief time limit (prior to executing their mission). Because of the time cap, students needed to ensure they placed the appropriate weight of effort in their briefs, ensuring that they did not spend an excessive amount of time on administrative information, and that they focused on briefing "how to conduct the actual mission at hand" (N. Meyers, personal communication, November 28, 2016). Inadequate weight of effort resulted in topics being over or under explained, and adequate weight of effort resulted in all the involved actors (e.g. sensor operator, other student pilot flying as wingman) receiving all the required information within the 55-minute time cap, and knowing how to execute the mission.

Pacing was treated as a subcategory within weight of effort. Pacing ties into weight of effort, but focuses more on student rate of speech, misspeaks, etc. (N. Meyers, personal communication, November 28, 2016). Pacing and weight of effort were treated as binary outcomes – students either demonstrated appropriate pacing and weight of effort in their briefs or inappropriate pacing and weight of effort, which often led to inadequately covering the required material prior to flying. Examples of good and poor weight of effort are provided in the below table:

Table 24

Weight of Effort and Pacing Comment Examples

Good Weight of Effort and Pacing	**Poor Weight of Effort and Pacing**
"good motherhood/tactical admin pacing"	"brief was heavily weighted towards SAR employment with only 5 minutes dedicated to instructing specific detail…"
"work on verbal pauses, pacing, and removing redundancy"	"poorly allocated with standard 13 mins for motherhood/Tac Admin… leaving little time to talk contracts and execution expectations"
"decent pacing and scope with only 40 min available post mass"	"lacked a solid cadence/pacing"
"good pacing and presence"	"pacing was off, did not cover appropriate content within allotted 55 min"

Level of Detail. The level of detail category determined whether students went into the appropriate level of detail, specificity, and depth in their explanations (brief and debrief). Level of detail was further broken down into whether the student provided detail on "*how* to do something," or *procedural knowledge*, which is shown when, "a person is able to apply a sequence of concepts representing condition and action to a general class of situations" (Gagne, 1984). Level of detail was treated as a binary outcome – students either went into the required level of detail, or were too "wave top" and did not get down to the required level of detail for the mission. Procedural knowledge, a subcategory of level of detail, was also treated a binary outcome – students either presented procedural knowledge (how to do something) when required, or they did not. Examples of positive and negative level of detail behaviors are highlighted in the below table:

Table 32

Level of Detail comment examples

Good Level of Detail	**Poor Level of Detail**
"covered criteria and formations roles in phases and how to effectively execute with well thought out contingencies"	"instructional fixes lacked detail information on how to fix the problem"
"examples were well thought out, and instruction was specific and flowed in a logical manner"	"lacked a methodology for how he was going to accomplish attacks which led to confusion during sortie"
"the brief contained good detail for execution on the FRA"	"lacked specifics in terms of desired airspeed, bank angle, and cross-check"
"IFs had good detail and appropriate depth of instruction"	"dynamic targeting examples lacked context and a standard methodology and game plan to execute"

Instructor Pilot (IP) Assistance. The IP assistance category captured whether students required IP interjection, redirects (revectoring), or general assistance during their brief or debrief. The briefs and debriefs are structured to where the Weapons School student is required to brief and debrief. Only when progress is stagnated, the student misspeaks, or the student explicitly requests guidance from the instructor, does the instructor intervene. For this category, instructor comments indicating assistance were the only ones noted, since the course expectation is that students do not require instructor assistance. Examples of students requiring instructor assistance are provided in the below table:

Table 26

Instructor Pilot Assistance comment examples

IP Assistance Required
“2nd DFP was initially poor, but with moderate IP questioning, WUG correctly scrapped the poorly scoped DFP question”
“after IP revectors, he had a better methodology”
“IP prompting was needed to help him focus the debrief to get the best instruction”
“required IP input to help him transition from a single ship focused debrief”
“was able to re-cage after quick vector check from the IP”

Efficiency. The efficiency category captured whether students were efficient in their briefs and debriefs, with respect to time and various strategies employed. Inefficient strategies resulted in students taking excessively long to perform routine tasks. Efficient strategies were timely and appropriate for the task at hand. Efficiency was treated as a binary outcome – students either practiced efficient strategies in their briefs or debriefs or inefficient strategies. Examples of students demonstrating both efficient and inefficient behaviors are highlighted in the below table:

Table 27

Efficiency comment examples

Efficient	Inefficient
"WUG had solid debriefing guide and was efficient at reconstruction"	"student didn't focus debrief on lost effects due to inefficient taskings, confusing comm, and lack of asset capes/lims understanding"
"WUG led a deliberate and efficient debrief"	"inefficient brief with redundancy in preattack checks/attack pacing"
"Was effective/efficient at debriefing with the ground party"	"led to issues in execution due to being task saturated and trying to implement an inefficient plan"
"was able to efficiently complete shot-kill matrix and chalk lines for CP timing attack"	"reconstruction was inefficient and took about 2x longer than necessary"
"debrief was efficient and covered all the learning from the sortie"	"this led to confusing logic flow and an inefficient debrief"

Organization. The organization category captured organized student behaviors. Organization was treated as a binary outcome – students either were organized in their briefs and debriefs or disorganized. Examples of students exemplifying organized and disorganized actions are provided in the below table:

Table 28

Organization comment examples

Organized	Disorganized
"DFP was correct and CFs were logical and organized well"	"CF structure in his DFP was unorganized and this led to confusing logic flow"
	"Motherhood and TAC Admin were not well organized and lasted 20 minutes"

Preparation/Planning. The preparation/planning category captured whether the students were noticeably prepared for the brief, flight, and debrief. Preparation/planning was treated as a binary outcome – students either were prepared for their missions, or were visibly unprepared. Examples of student behaviors exemplifying both prepared/planned and unprepared/unplanned behaviors are highlighted in the below table:

Table 29

Preparation/planning comment examples

Prepared	Unprepared
"Well prepared with multiple flight DFPs post mass"	"WUG was not prepared for the debrief to start as soon as the JTAC returned"
"Provided valid inputs for brief and ensured flight lead was prepared"	"WUG elected to begin debrief without being fully prepared"

Repeatable Process/Methodology. The Repeatable Process/Methodology category captured if students utilized repeatable processes or a methodology in their brief, mission, and debrief. Methodology was defined as "a system of methods used in a particular area of study or activity" (Merriam-Webster, 2016). Use of a repeatable process or a methodology was treated as a binary outcome – students either employed a repeatable process or a methodology, or they neglected to. Examples of using/failing to use repeatable processes or a methodology annotated in student grade sheets are summarized in the below table:

Table 30

Repeatable Process/Methodology comment examples

Use of Repeatable Process/Methodology	Failure to use Repeatable Process/Methodology
"brief did outline repeatable process"	"student did not have a consistent repeatable process to execute PT HF attacks"
"WUG obviously had a very good and repeatable process for executing the planning tool"	"planning tool instruction was difficult to follow as it was not briefed as a repeatable process"
"high points included a repeatable process to instruct timing attacks, max-performance descent, LOWAT and vertical target attacks"	"ad hoc 2-ship attacked lacked methodology or repeatable process"

Decision Criteria. The Decision Criteria category assessed the depth of thought a student gives a problem to see if they have thought through different types of events/contingencies. This is important in briefs as it has the potential to save time during mission execution (N. Meyers, personal communication, November 28, 2016). Decision criteria was treated as a binary outcome – students either presented decision criteria in situations that required it, or did not provide decision criteria. Examples of using/failing to use decision criteria are highlighted in the below table:

Table 31

Decision Criteria comment examples

Presented Decision Criteria	**Failed to provide Decision Criteria**
"WUG didn't settle for easy solution, rather briefing more complex decision criteria"	"brief did not have a coherent flow and lacked the triggers and decision criteria to get into and out of specific phases of SCAR"
	"leaving little time to cover important processes/decision criteria to deconflict and organize assets"
	"needs to spend more time on decision criteria and then give an example of the criteria"

Triggers. The triggers category captured whether students briefed appropriate triggers prior to flying each mission. Triggers typically refer to the various phases of a mission, such as: marshall, ingress, target attacks, egress/DT attacks. They are the events that need to happen in execution to transition from one phase in a mission to the next (N. Meyers, personal communication, November 28, 2016). Triggers were assessed as a binary outcome – student either presented triggers when a situation required them or failed to present them. Examples of students using/failing to identify triggers are presented in the below table:

Table 32

Triggers comment examples

Presented Triggers	**Failed to Present Triggers**
"did a good job of explaining how he was going to execute the briefed contracts based on specific triggers and criteria"	"lacked triggers for when to initiate a coordinated attack"
"overall the WUG effectively briefed a phase based plan with triggers"	"failed to highlight decision triggers"
"good overall understanding of the phase based approach and triggers in/out of those phases"	"lacked appropriate comm examples and triggers necessary to go out and properly execute the ride"
"brief covered what-when-how and included triggers into the next phase"	"low point was lack of developed triggers and contracts to integrate with Viper and JTAC players"

Contracts. The contracts category captured if students established appropriate contracts with their teammates (sensor operator, wingman, etc.) prior to flying their missions. Contracts are a formal agreement between 2 or more actors outlining when to act (or not act). Contracts were assessed as a binary outcome – students either established appropriate contracts between actors, or failed to establish them. Examples of students using contracts/failing to use contracts are summarized in the below table:

Table 33

Contracts comment examples

Briefed Contracts	**Omitted/Inappropriate Contracts**
"did a good job at explaining how he was going to execute the briefed contracts based on specific triggers and criteria"	"didn't lay out contracts for proper CAS procedures"
"effectively briefed a phase based plan with triggers and contracts"	"briefing was marred by some illogical contracts for #2"
"brief was average, outlined key contracts"	"poor contracts in mission planning led to WUG accepting that MoD was killed"
"WUG had some well thought out contracts"	"organization plan fell apart due to ineffective contracts with the sensor operator"

Quality of Instruction. The quality of instruction category captured how well the students performed as instructors. Weapons School is an advanced instructor training course, and the incoming students are required to have a minimum of 75 hours as a flight instructor as of the class start date (N. Meyers, personal communication, November 28, 2016). Quality of instruction was assessed based on the quality of various instruction points by the student (plan briefed prior to flying, instructional fixes provided, misspeaks, resultant confusion from instruction, etc.), in addition to how well students captured the available learning in debrief focus points. Students are expected to capture mistakes in execution during the mission portion, and expand upon these errors, ultimately reaching the root cause for the errors committed during the mission. The expectation is that students select errors from the mission (even if there were numerous errors), that provide the best opportunity for learning. Procedural errors during the mission, inadequate planning, and lack of a shared mental model in the cockpit were commonly cited errors in the student grade sheets. However, it is up to the student to determine which error (if choosing between several) has the greatest opportunity for learning, and further develop that into a debrief focus point. Any shortcomings should be logically and systematically addressed to flesh out all contributing factors so that no learning is missed (B. Gyovai, personal communication, May 11, 2016). Further, in instructing the fix to errors, it is expected that students inform the audience of the perceptions that need to occur, the decisions that need to be made, and the subsequent actions that need to be performed, to the utmost level of detail (B. Gyovai, personal communication, May 11, 2016).

Creation of instructional fixes are a part of the DFP process taught in Weapons School, but since instructional fixes are intended to provide a detailed explanation of a fix to avoid repeat of mission errors, any comments pertaining to the quality of students' instructional fixes were counted in the quality of instruction category. Quality of instruction was assessed as a binary outcome – students either exhibited good instructional quality, or poor instructional quality. Further, the learning captured subcategory was also assessed as a binary outcome – students either captured all available learning in their debriefs, or they missed opportunities for additional learning. Examples of both good and poor instructional quality are highlighted in the below table:

Table 34

Instructional Quality comment examples

Good Instructional Quality	**Poor Instructional Quality**
"Examples were well thought out, and instruction was specific and flowed in a logical manner"	"WUG was "stuck" in determining how to approach instruction and DFP for event failure"
"DFP was appropriate for the sortie…IFS had good detail and appropriate depth of instruction"	"WUG failed to brief a coherent 2-ship SCAR fighter plan"
"Highlight was WUG's presence and ability to scope debrief appropriately"	"Flyout maneuver instruction was vague and confusing"
"Some minor misspeaks, but overall solid instruction. Above average integration of visual aids"	"Brief was overly basic and delved into academics at a few points"
"WUG did a nice job tailoring examples to the scenario"	"The DFP consisted of a list of thirty-three CFs that were really a compilation of errors committed throughout the sortie, which led to ineffective root cause analysis and instruction"
"Briefing was outstanding, hit highlights and good explanation to sensors and #2 (Wingman) on criteria, roles, and tasks"	"Needs to present expert level formation instructional debrief; student focused on procedural adherence"

Artifact/Tool Usage. The artifact/tool usage category captured whether students properly utilized various artifacts and tools such as routine checklists, and the various software tools available in the MQ-9 cockpits and in the debriefing rooms. Artifact/tool usage was assessed as a binary outcome – students either used artifacts appropriately, or they used them inappropriately. Examples of both appropriate and inappropriate artifact usage noted in student grade sheets are presented in the below table:

Table 35

Artifact Usage comment examples

Appropriate Artifact Usage	**Inappropriate Artifact Usage**
“WUG used decent examples on the products that would be used in the cockpit to chairfly the briefed scenario”	“WUG was missing matrices from the debriefing guide that could have helped capture the flaws in execution”
	“WUG had a debriefing guide but didn’t use it in an attempt to expedite debrief”
	“WUG accidentally gained the aircraft due to inappropriate checklist procedures”

Transfer. The transfer category captured whether students demonstrated negative and positive transfer. Weapons School uses a building block approach, and once a skill is mastered, introduction of a new problem should not imply “throwing out” previous learning (B. Gyovai, personal communication, May 11, 2016). Transfer, in an academic sense, is “the ability of individuals to ‘treat’ a new concept, problem, or phenomenon as similar to one(s) they have seen before” (Chi & VanLehn, 2012).

Transfer was further broken down into intra-mission transfer (positive or negative transfer within the same mission), mission-to-mission transfer (positive or negative transfer from mission to mission within the same phase), and transfer over the course of the program (positive or negative transfer across phases). Transfer was assessed as a binary outcome across all three types – students either exhibited negative transfer, or carried over bad habits from prior missions/phases, or positive transfer, which equated to students incorporating instructor feedback, and not repeating the same errors across missions/phases. Examples of both positive and negative transfer observed in the student grade sheets are highlighted in the below table:

Table 36

Transfer comment examples

Positive Transfer	Negative Transfer
"First attempts at CFs were confusing and were difficult to follow. WUG corrected the mistakes on the 2nd DFP" (Intra-mission transfer)	"Good DFP flow, but similar issues to debrief on SA-1X" (phase-to-phase transfer)
"Overall, WUG effectively addressed most deficiencies from previous sorties and put together a solid performance (phase-to-phase transfer)	"WUG still struggles with using the DFP question to focus reconstruction and learning" (mission-to-mission transfer)
"Showed improvement from previous trends" (mission-to-mission transfer)	"WUG was debriefed to not forget debrief basics even in a mass environment" (phase-to-phase transfer)
"WUG assigned to reaccomplish 2nd DFP to WIC level CFs and Ifs. WUG made progress upon second attempt (Intra-mission transfer)	"WUG initially continued previous trends of rambling without adding learning to the board" (phase-to-phase transfer)

Self-identification of errors. The identification of errors category captured if students could self-identify errors committed in the brief (tied to the planning stages prior to the mission) or in the mission and appropriately brief to and instruct to the error solution in the debrief. The self-identification of errors was treated as a binary outcome – students either self-identified their errors, or they failed to, and the instructor was required to intervene. Examples of students self-identifying and failing to self-identify are summarized in the below table:

Table 37

Identification of errors comment examples

Self-identification of errors	**Failure to identify errors**
"Flight debrief sufficiently covered relevant errors"	"WUG struggled with the first DFP due to focusing on the types of tracks SO was attempting to grow and where they were placed vs. the fact that a track never should have been attempted in the first place"
"WUG correctly identified the errors in the flight and the CFs"	"WUG has a tendency to throw down LPs early that are actually a part of larger, AO level DFP he's unable to see"
"WUG did a good job informing how he would've done things differently to the masses to optimize his effects"	"Second DFP misprioritized MTK and was very focused on single-ship execution errors, little focus on 2-ship/flight level reconstruction and fixes"

Collaboration. The collaboration category assessed if students effectively collaborated with their teammates. The MQ-9 is operated by a two-person team consisting of a pilot and a sensor operator. Inherently, the MQ-9 is a team environment and it is up to the pilot to work with the sensor operator to accomplish the mission. In the later phases of the Weapons School course, the MQ-9 students integrate with various other Air Force platforms and players, and collaboration becomes of paramount importance. Collaboration was treated as a binary outcome – students either collaborated with their teammates or failed to collaborate. Collaborative and non-collaborative behaviors are highlighted in the below table:

Table 38

Collaboration comment examples

Collaborated	**Failed to Collaborate**
"WUG was critical to integrated debrief, providing valuable instruction and leadership. RMC asked WUG no less than 26 times for his input"	"WUG needs to… incorporate SO into plan/do CRM"
"Good job at 1st integrated debrief (JTACS)… inclusion of guest help in developing lessons, and identifying important DFP elements"	"looked like strangers, not WIC studs pooling capes to solve a problem"
"Good interaction with JTACs with high SA demonstrated"	"WUG needs to work on fully developing CFs and involving SO more"

Weapons Officer Quality. The Weapons Officer Quality category captured if the students exuded the expertise and presence expected of a USAF Weapons Officer. The Weapons School creed is "humble, approachable, and credible" (United States Air Force, 2015). With the goal of training tactical experts and leaders in their airframe, students are assessed on their Weapons Officer potential throughout the course of the program. Weapons Officer presence was treated as a binary outcome – students either exhibited Weapons Officer presence or they did not. In extreme cases where students did not exhibit behavior becoming of a Weapons Officer, conduct discrepancies were noted in the grade sheets. Examples of Weapons Officer Presence and lack of presence (extended to conduct issues) are presented in the below table:

Table 39

Weapons Officer Presence/Conduct comment examples

Demonstrated Weapons Officer Presence/Conduct	**Lacked Weapons Officer Presence/Conduct**
"Solid Weapons Officer presence for SERE specialists"	"This led to long moments of silence and little Weapons Officer presence in the debrief"
"WUG had solid Weapons Officer conduct throughout sortie and debrief"	"WUG's self-deprecating humor detracted from Weapons Officer Conduct"

Valuable contributor. The valuable contributor category assessed if students were valuable debrief contributors. As the Weapons School syllabus progresses, students integrate with other Weapons School students flying different Air Force platforms, and with students filling integrated roles on the ground. During the integrated missions, the MQ-9 Weapons School student is not necessarily the individual that is leading the integrated debriefs. In these situations, where the MQ-9 Weapons School student is not leading the debrief, they are expected to be valuable contributors to those who are leading the integrated debrief. Valuable contributor was assessed as a binary outcome; students either were valuable contributors, or they weren't. Examples of valuable contributor behaviors, and one non-valuable contributor behavior are summarized in the below table:

Table 40

Valuable contributor comment examples

Valuable contributor	**Not a valuable contributor**
"WUG was an effective contributor during the mass debrief"	"The WUG told the mass debrief team his Link-16 was inoperative for the duration of the vul, however this was not accurate"
"Effective contributor in mass debrief, and focused learning appropriately in the flight debrief"	
"Effective contributor during the mass debrief. Although did not lead the debrief, he was a major contributor to one of his flight lead's DFPs"	

Adaptability. The adaptability category captured if students readily adapted to novel or unexpected/unplanned situations. Adaptability is defined as, "appropriate cognitive, behavioral, and/or affective adjustment in the face of uncertainty and novelty" (Martin, Nehad, Colmar, & Liem, 2013). Adaptive students did well in "first time" situations, whereas non-adaptive students had difficulty with new situations and problem sets. Adaptability was treated as a binary outcome; students either exhibited adaptive behaviors, or they failed to. Examples of adaptive behaviors and one non-adaptive behavior are highlighted in the below table:

Table 41

Adaptability comment examples

Adapted well	**Did not adapt well**
"Student gave good brief for first attempt at SCAR"	"WUG has difficulty applying learning for new problem sets when he cannot rely on experience"
"Good job at 1st integrated debrief (JTACs)"	

Flight/Area of Operation (AO) Leadership. The Flight and AO leadership category assessed how students led the wingman in their flight, how they effectively managed an entire formation, and how students adjusted plans upon recognition that the mission at hand was not going as planned (N. Meyers, personal communication, December 10, 2016). When the mission at hand is not going as planned, the student is expected to provide information/recommendations to the entire package (other aircraft platforms) to get the war back on track (N. Meyers, personal communication, December 10, 2016). Flight and AO leadership was treated as a binary outcome; students either demonstrated Flight/AO leadership or they did not. Examples of Flight/AO leadership are summarized in the below table:

Table 42

Flight/AO Leadership comment examples

Demonstrated Flight/AO Leadership	**Lacked Flight/AO Leadership**
“Flight leadership and generic SCAR tasking flow was solid”	“WUG did not provide the flight leadership necessary to ensure the -2 was in position and deconflicted throughout the sortie”
“He was able to show good AO leadership by adjusting to the briefed game plan by taking on the role of OSC for the two survivors”	“WUG had a fundamental lack of knowledge of what it meant to be the SCAR and how to be an effective/proactive AO leader to share the battlespace”
“WUG demonstrated flight leadership, comm contracts and refined 2-ship understanding”	“Flight leadership was lacking. There were many times throughout the sortie when wingman was untasked and didn’t have a specific place to hold”