Issues Related to Experience & Automated Agent Technology in Synthetic Task Performance

Michael D. Coovert, Dawn L. Riddle,
Thomas Gordon & Donald Miles
Institute of Human Performance, Decision Making & Cybernetics
University of South Florida, Dept. of Psychology, BEH 339
Tampa, Florida 33629-8200

Kimberly Hoffman & Thomas King
University of South Florida, Dept. of Psychology, BEH 339
Tampa, Florida 33629-8200

Linda Elliott & Mathieu Dalrymple
Veridian Engineering
2504 Gillingham Drive, Suite 201
Brooks AFB, Texas 78235

Samual Schiflett and Scott Chaiken
Air Force Research Laboratory
AFRL/HEAI, Brooks AFB, Texas 78235

Abstract

The U.S. Air Force's E-3 AWACS is a highly mobile early warning platform providing both surveillance and command and control functions for tactical and air defense forces. The weapons suite within the AWACS is occupied by multiple weapons directors (WD) and a Senior Director (SD), is critical to mission success. The cognitive demand placed on these team members is extremely high. Elliott, Stoyen and Chaiken (2000) note the limits of human cognitive processing relative to this demand and suggest agent technology has the potential to offer effective information and decision support the human operator. This paper examines relationships among experience, performance and the use of agent technology in a decision support role within a synthetic task. We first examine experience related factors leading to utilization of the agent, and subsequently explore the relationship between experience and performance, and how the presence of an automated agent might impact that relationship. Two interesting findings are presented: 1) differences in qualitative experience impact the degree to which weapons directors utilize agent assistance, and 2) the nature of the experience – performance relationship changes when agent assistance is available.

Introduction

The U.S. Air Force's E-3 AWACS is a highly mobile early warning platform that provides both surveillance and command and control functions for tactical and air defense forces. The AWACS mission is two-fold: its tactical role is to provide rapid deployment and quick-reaction surveillance by detecting, identifying, and controlling aircraft within its area of responsibility (AOR); while the aircraft in its defensive role is responsible for tracking and intercepting airborne threats. To accomplish this, AWACS platforms are strategically divided into weapons, surveillance, and self-defense suites. The weapons suite, occupied by a team composed of weapons directors (WD) and a Senior Director (SD), is critical to mission success by directing, coordinating, and communicating with aircraft within its AOR.

The cognitive demand placed on weapons directors is extremely high. For instance, a WD might simultaneously be directing a friendly fighter to intercept, sending an aircraft to refuel, tracking a hostile threat, and monitoring an aircraft launch – all while maintaining proper awareness of Rules of Engagement. Elliott, Stoyen and Chaiken (2000) note the limits of human cognitive processing relative to this demand and suggest agent technology has great potential in offering effective information and decision support. They also point out that much research is necessary to determine how to optimize the human–agent relationship.

The current investigation responds to the need for such research regarding the use of agent technology in support of the human operator. We first examine experience related factors leading to utilization of the agent, and subsequently explore the relationship between experience and performance, and how the presence of an automated agent might impact that relationship.

Prior work (Coovert & Riddle, 1999; Riddle, Coovert, Gordon, King, Hoffman, Miles, Elliott, & Schiflett, 2000) suggests an approach to *mining* complex relationships from complex data sets is appropriate for the type of data collected in this study; therefore we examine issues related to the use of agent technology within the synthetic task using a data mining tool based on Rough Sets Theory.

Rough Sets Theory

Rough set theory (Pawlak, 1982) provides the theoretical foundation for rule induction systems, which develop logical statements regarding relationships among variables that in turn are used to classify data. For example, rough set analysis may be used to define the causal relationship between specific symptoms and a diagnosis. Rules are derived from patterns of symptoms associated with a known diagnosis. These rules are then applied to classify or diagnose new patients. In addition to the medical field, rough set theory has been applied to problems in areas such as machine learning, knowledge acquisition, inductive reasoning and pattern recognition in artificial intelligence and cognitive science.

One main advantage of rough sets theory is that is does not make assumptions about the form or distribution of the data (e.g., such as the requirement of multivariate normal within and between variables, required by many traditional analyses). This makes it attractive for application to the problem domain identified in the introduction of this paper. Table 1 presents some key terminology from rough sets theory described within an Air Force context. These terms and concepts are elaborate on in the following example.

Table 1. Rough set theory terminology applied to Job Experience – Job Performance Relationship.

Term	Job Experience Example
Cases	US Air Force personnel; e1, e2, e3, …, eN
Attributes (predictors)	Job Experience variables 1) Number of times performed, 2) Months in present unit
Decision (outcome)	Composite Job Performance & Hands-on performance test scores.
Elementary Sets	Sets of AF personnel with equivalent attribute (experience) profiles
Concept Set	Set of AF personnel performing at the same outcome level (e.g., effective).
Lower Approximation (for a given concept)	Sets of AF personnel with equivalent experience profiles where all members are *consistently* performing at the same level (e.g., effective).
Upper Approximation (for a given concept)	Set of AF personnel with equivalent experience profiles that are performing at the specified level (e.g., effective), as well as, AF personnel with the same profile who are performing at different levels (e.g., ineffective).
Reduct	A set of attributes (e.g., set A) is the reduct of another set of attributes (e.g., set B) if A is minimal and the elementary sets defined by A and B are identical.
Rules	
Certain	Specific experience profiles are *without exception* predictive of a particular level of performance. (Certain rules are induced from the Lower Approximation)
Possible	Specific experience profiles are s*ometimes* predictive of the specified level of performance, and sometimes *not* predictive of the level of performance. The profiles are sometimes associated with a different level of performance as well. (Possible rules are induced from the Upper Approximation)

Rough Set Analysis: An example

Table 2 is an information or decision table involving experience profiles and performance ratings for six USAF personnel. The remainder of this section leads you through the application of the rough set procedure to the data in Table 2, concluding with a summary of the process shown in Table 3.

Table 2. Decision table specifying job experience profiles and job performance ratings.

Case	Attributes		Decision
Individual	Months in **Present Unit**	# of **Times Performed**	Rating of Performance
e1	0-12	5-9	effective
e2	0-12	0-4	ineffective
e3	13-36	5-9	ineffective
e4	37+	10+	effective
e5	13-36	5-9	ineffective
e6	0-12	0-4	effective

In this example, individuals e2 and e6 have identical profiles, each with 0-12 months in the present unit, and having performed the task 0-4 times. Similarly, individuals e3 and e5 have identical attribute profiles with 13-36 months in the present unit and having performed the task 5-9 times. Based on their identical attribute profiles, individuals e2 and e6 form one elementary set and individuals e3 and e5 form a second elementary set. Unique cases such as e1 and e4 are the sole members of separate elementary set and are termed singletons. These elementary sets reflect attribute patterns essential for differentiating among cases.

A similar process occurs relative to the outcome or decision variable. Cases are grouped according to decision value. For example, profiles for e1, e4, and e6 are associated with the concept effective, while e2, e3, and e5 are associated with the concept ineffective.

Next, relationships between elementary sets (predictor attribute profiles) and concepts (outcomes) are also identified. Two issues need to be noted here. First, keep in mind that a number of distinct profiles may be associated with a single outcome. Second, the same profile may be associated with a particular outcome in one instance, and a different outcome in another instance.

Rough set theory addresses these issues by generating two classes of rules for each outcome. The first type of rule is generated from the Lower Approximation. The Lower Approximation describes those profiles that are associated -- without exception -- with the particular outcome. That is, every time a particular profile is encountered in the data, it is associated with a single, specified outcome or concept. For example, the profile of individual e1 with 0-12 months in present unit and 5-9 times performed is always associated with an effective outcome (it is the only time this profile occurs). Similarly, the attribute profile of individuals e3 and e5 is always associated with the ineffective outcome.

The second type of rule is generated from the Upper Approximation and accounts for those cases where the profile is related to more than one outcome. The profile is associated with the outcome of interest, but it is also related to another outcome as well. This type of rule is much less definitive. It suggests that the specified outcome *might* be obtained when the profile is present. In our example, individuals e2 and e6 have identical profiles, but the outcome for e2 is ineffective while for e6 it is effective.

As stated earlier, lower and upper approximations are generated for *each concept*. Although e1, e4 and e6 are members of the concept “effective”, only individuals e1 and e4 fall into the lower approximation for concept (effective) because only the profiles exhibited by those individuals are consistently associated with the “effective” outcome. The Upper Approximation (effective) consists of those individuals rated effective (e1, e4, e6), plus individuals with the same profile as those rated effective. Therefore, e2 which is characterized by the same profile as e6, falls into the Upper Approximation (effective) even though it is associated with an ineffective outcome.

Now that relationships among predictors and between predictors and outcomes have been identified, reducts are identified defining the relationships for each case in the decision table. If a subset of variables (a reduct) can provide the same discrimination between cases as the total set of variables, then those variables outside of the subset are unnecessary. That is, they do not contribute unique information to discriminate among cases.[1] Therefore, only those attributes providing unique information from which to differentiate cases are captured in reduct sets.

To keep our discussion manageable, only reducts for the Lower Approximation (effective) in our example are listed. Attributes important for discriminating e1 from all other examples, include *months in present unit* and *number of times performed* creating the [present unit, times performed] reduct set. In order to discriminate e4 from all other examples, only the *months in present unit* data is needed, resulting in a [present unit] reduct. This process is repeated for the concept ineffective (for example, e3 and e5 would fall into the Lower Approximation for “ineffective” because the profile presented by those individuals consistently yields the same “ineffective” outcome). Rules are then generated reflecting patterns of attribute -- outcome relationships identified in the lower and upper approximations. The entire rule set can then be used to classify (predict) job performance of new personnel based on his/her experience profile.

Table 3. Rule generation for experience-performance example from Table 2.

Concepts	Examples from Table 2
Elementary Sets	Elementary Set (Present Unit, Times Performed) = {e1}{e2,e6}{e3,e5}{e4}
Concept (effective)	Concept (effective) = {e1,e4,e6}
Lower Approximation	Lower Approximation (effective) = {e1,e4}
Upper Approximation	Upper Approximation (effective) = {e1,e4,e6,e2}
Reduct (e.g., Lower Approximation only)	[present unit], [present unit, times performed]
Rules	*CertainRule1*: [(Present Unit,37+)] → (effective) *CertainRule2*:[(Present Unit,0-12),(Times Performed,5-9)]→ (effective) *PossibleRule1*: [(Present Unit,0-12)] → (effective)

[1] This is similar to the situation in traditional statistical analyses when one variable is perfectly correlated with a second variable, or one has zero variance.

In addition to classification, we can investigate the nature of the relationships among predictors and outcome by closer examination of the overall reduct and rule sets. The frequency of occurrence of variables within the reduct set provides the relative importance of variables in discriminating among cases with respect to the decision or outcome variable. This provides a starting point from which to begin examining the rule set. Rule statistics offer unique information regarding the value(s) of those predictors that are contributing to discrimination. Finally, examination of the rules themselves provides insight into the specific nature of the relationships identified.

In the current effort, rough sets analysis is employed to examine experience related factors that discriminate between participants who are using the agent to a greater or lesser extent. In addition the technique is employed to explore the relationship between experience and performance, and how the availability of agent decision support might impact that relationship. In terms of the rough sets analysis we focus discussion on the reducts to identify the relative influence of attributes on classification and on detailed examination of the rule set to provide specific information regarding the nature of the relationships among predictors and outcomes.

Method

Task

The software used in this study provides an AWACS weapons director platform with automated agent capabilities. This software is a distributed, simulated, real-time team environment comprised of air, sea, and ground assets in a combat environment (Stoyen, Chiara, Schifflett, Elliott, Walrath, & Dalrymple, 1999), and is based primarily on the roles and responsibilities of AWACS WD team members. Decision support is supplied in the form of agent technology. Introduced as a "coach" or "adviser", an agent provides recommendations for action based on scenario events. The participant can either accept or ignore agent recommendations.

Participants

Thirty eight weapons director personnel stationed at Tinker Air Force Base in Oklahoma City, Oklahoma were used in this study. Of these, 31 were male, 4 female, and 3 did not complete the demographic portion of the biographical questionnaire.

Procedures

Participants engaged in a pre-mission brief outlining rules of engagement for the scenarios. This was followed by 2 hands-on training sessions during which participants were instructed on the functionality of the platform, as well as the use of the agent. Upon completion of training, participants engaged in 3 trials (parallel forms of an AWACS WD scenario) using the platform described earlier. Participants randomly received either 1 or 2 of 3 trials with access to agent recommendations, while the remaining trial(s) offered no decision support.

Measures

BioData Questionnaire

A BioData questionnaire was used to obtain information regarding participant demographics and experience. Of interest to this effort are both quantitative and qualitative measures of experience.

Specific measures include: years as a WD, simulator hours, E-3 flight hours, STK hours, CAP hours & HVA hours, occupation, flight status, number of evaluations, Combat Mission Ready / Inexperienced (CMR/I) versus Combat Mission Ready/ Inexperienced (CMR/E), instructor qualified, and Basic Mission Capable (BMC).

Agent Recommendations Accepted

Participants accepted an agent recommendation by either clicking a specified "accept" button, or by manual execution of the recommendation within a specified period of time. The number of accepted recommendations is summed across the scenario. For participants encountering 2 trials with the agent, counts of accepted recommendations were averaged across trials yielding one score.

Task Performance

Task performance is based on the individual's overall score for the scenario. Participant performance is derived from a scoring algorithm that focuses primarily on points gained by destroying hostile resources and points lost by losing friendly resources.

Continuous variables were pre-processed or discretinized based on frequency distributions (see Table 4) for use in rough sets analyses.

Table 4: Cutoff values for variables used in rough sets analyses.

Measures	Dichotomous Values	
Quantitative Experience	Low	High
CAP hours	0	>0
STK hours	0	>0
HVA hours	0	>0
Simulator hours	X<250	X=>250
Years of Experience	X<1.66	X=>1.66
E3-flight hours	X<400	X=>400
Performance	X<141	X=>141
Task Performance Score		
Agent Use	X<=5	X>5
Recommendations Accepted		
Qualitative Experience	Values	
Flight Status	MR	DNIF
Occupation	WD	AWO
Number of Evaluations Completed	Low (2 or fewer)	High (3 or more)
CMR/E	No	Yes
CMR/I	No	Yes
BMC	No	Yes
Instructor Qualified	No	Yes

Results

Experience and Agent Use

Exploration of potential differences in agent use due to experience was conducted using two rough sets analysis. The first analysis examines relationships between *quantitative* measures of experience (years as a WD, simulator hours, E-3 flight hours, STK hours, CAP hours & HVA hours) and agent use, while the second involves *qualitative* measures of experience (occupation, flight status, number of evaluations, Combat Mission Ready / Inexperienced (CMR/I) versus Combat Mission Ready/ Inexperienced (CMR/E), instructor qualified, and Basic Mission Capable (BMC)) and agent use.

Quantitative Experience and Agent Use

Using trials in which participants had access to the agent, a 70% development sample (n=26) was created to generate reducts and rules describing the relationship between quantitative experience and the number of recommendations accepted. The remaining 30% cross validation sample (n=11) was used to test the quality of the relationships identified.

Examination of the reducts will allow us to identify quantitative experience characteristics that discern between participants who accept a high number of agent recommendations and those who accept a low number of agent recommendations. The frequency of occurrence of each variable in the reduct sets reflects the relative contribution to discrimination. In this analysis, years of experience as a WD occurs most frequently followed closely by simulator hours; hence they have the greatest influence on predicting high versus low acceptors. With fewer occurrences E3- flight hours, STK hours, CAP hours, and HVA hours have relatively less influence on recommendation acceptance.

This information from the reducts guides our investigation of the rule set, suggesting that the most consistent, interpretable relationships will be found between years of WD experience and agent use, and between simulator hours and agent use. Upon examination of the rules we find that high years of WD experience (over 1.66 years) is consistently associated with accepting a high number of recommendations. Similarly, low years of WD experience yield a low number of accepted recommendations. In addition, high simulator hours are associated with high acceptance of recommendations and rarely with low acceptance; while low simulator hours are generally associated with low recommendation acceptance. As anticipated, experience in terms of E-3 flight hours does not differentiate between high and low frequency of recommendations accepted. The only consistent pattern emerging from the data regarding flight hours reveals that high flight hours are generally *not* associated with low recommendations accepted. It follows that if E-3 flight hours is not predictive of recommendation acceptance, than role specific measures of flight hours, including CAP, STK, and HVA hours are not discerning between high and low acceptance groups either. Applying rules defining these relationships to the cross validation sample yields correct classification in 82% of the cases.

In sum, results suggest that experience profiles associated with a high number of accepted agent recommendations (more than 5) are consistently characterized by high years of WD experience (1.66 years of more), and / or high simulator hours (250 hours or more). Conversely, participants accepting a low number of agent recommendations (5 or fewer) are characterized by low years of

WD experience (less than 1.66 years). The 82% classification accuracy of the holdout sample suggests these relationships hold when classifying unseen cases.

Qualitative Experience and Agent Use

The second rough sets analysis examines the relationship between variables that reflect competency or skill related experience and the number of recommendations accepted. Again, rule development involved 70% of the available cases (n=26), with the remaining 30% to be used or cross validation (n=11).

Examination of the reducts suggests occupation ranks highest in terms of importance in discriminating between participants accepting a high versus low number of recommendations. Occupation is followed by flight status, the number of evaluations completed, CMR/I, CMR/E, instructor qualified, and finally BMC. In terms of occupation, being an enlisted weapons director leads to accepting a low number of recommendations, while being an AWO is consistently associated with accepting a large number of agent recommendations. WDs accept a high number of recommendations only when they are also rated inexperienced combat mission ready. Flight status was useful in differentiating among participants only in that of the two values for flight status only one, mission ready, was related to recommendation acceptance. No consistent relationship between DNIF and accepting agent recommendations was identified. Rules also suggest that having completed more than two evaluations is consistently related to a low number of accepted agent recommendations, whereas having completed two or fewer evaluations leads to accepting a high number of recommendations. Finally, being CMR/I is consistently related to only high acceptance. The converse however does not hold, not being CMR/I is associated with both low and high acceptance of agent recommendations.

Using the rules defining the relationships described above, rough sets correctly classified 90% of the cross validation sample. Again this level of accuracy suggests the relationships described which underlie the rules being used to classify the holdout sample, are robust.

In sum, participant's accepting fewer recommendations are characterized as WDs, and individuals who have completed a high number of evaluations. Conversely, participants accepting fewer recommendations are characterized as AWOs, and individuals having completed fewer evaluations and who are CMR/I.

Experience, Performance, and Agent Condition

The results above identify experience related characteristics of individuals accepting a high number of recommendations and of those accepting a low number of recommendations. One goal of automated decision support being enhanced performance, the following analysis examines the impact of agent availability on the relationship between experience and performance. Two analyses were conducted to identify experience variables impacting performance, one using trials in which participants had access to the agent, and one in which participants did not have access to the agent.

Agent On Only

As in the previous analyses, the rank order of factors contributing to discrimination is determined by the frequency of occurrence of each variable in the reduct sets. Occupation occurs most frequently, appearing in nearly 46% of the reducts. Occupation is followed by number of evaluations (37.9%), E-3 flight hours (35.6%), simulator hours (35.6%), CMR/E (33.3%), years of experience as a WD (31.0%), instructor qualified status (29.9%), basic mission capable (25.3%), and CMR/I (25.3%), respectively.

Examination of the rule set found that being a WD versus an AWO was consistently related to high performance, while the AWO occupation was generally associated with low performance. It is important to note that our sample consisted of AWOs who were generally less experienced then their enlisted counterparts. Furthermore, participants having completed a relatively low number of evaluations generally preformed more poorly on the task. Having completed a high number of evaluations was consistently related to high performance.

Low simulator hours did not inhibit task performance as long as participants had high experience on other measures. When associated with low hands on experience, low performance was observed even when simulator hours were high. Simulator hours could not compensate for a lack of hands on experience. Being CMR/I was associated with low performance, while not being CMR/I was associated with high performance. Finally, not being Combat Mission ready / Experienced routinely yields low performance, while being Combat Mission ready / Experienced is frequently associated with high performance, and in this sample, never associated with low performance. Overall, the rule set defining these relationships correctly classified 80% of the holdout sample.

In sum, when individuals have access to agent recommendations high performance is anticipated if participants are WDs, have completed a high number of evaluations, and have few simulator hours. Conversely, low performance is predicated by few evaluations, high simulator time and being an AWO.

Agent Off

A similar analysis was performed examining the experience – performance relationship in the agent off conditions.

Results show the most frequently occurring variables in the reduct set are simulator hours (37%), occupation (36%), years as a WD (36%), number of evaluations (34%), and CMR/E (31%). Hence, these variables will be examined for consistent relationships with performance within in the rule set.

For the agent off condition, rules demonstrate that low simulator experience yields low performance. Even those participants who have a high number of years as a WD or who are instructor qualified score low if they do not have high simulator hours. High simulator hours is consistently related to high performance. Even when other experience indicators are low, individuals with high simulator hours demonstrate high performance. This suggests that relative to performance on the synthetic task, real world experience does not compensate for low simulator hours, however high simulator hours can compensate for low experience.

WDs with little experienced typically performed low on the task. Those WDs performing high on the task, were generally either CMR/E or instructor qualified. Being an AWO, on the other hand, was consistently associated with high performance.

In addition, it was found that low years of experience was generally related to low performance, while high years of experience lead to high performance, except when the participant had low simulator hours. In addition, fewer completed evaluations was generally associated with low performance. A high number of evaluations was only associated with low performance when occurring with low simulator hours, or when the individual was not CMR/E. A high number of evaluations is related to high performance when other experience variables are high as well. Participants with a low number of evaluations only performed high if they also had acquired a high number of simulator hours.

Finally, Not being CMR/E often results in low performance, whereas being CMR/E is associated with high performance. Not being Combat Mission ready / Experienced was related to high performance only when participants were AWOs.

Table 5 summarizes the experience–performance relationships outlined above for both the agent on and off conditions. When considering only those trials where the agent was not made available, high performers are characterized as being an AWO (or a WD with CMR/experienced classification), high simulator hours, high years of experience and high number of evaluations. Low performance is characterized by being a WD, having low simulator time, years of experience, and evaluations. The accuracy of classification when using rules based on the relationships described above to predict performance in the agent off conditions is 90%.

Table 5. Experience characteristics of low and high performers by agent condition.

Performance - Agent On	
Low Performers	**High Performers**
AWO	WD
Low Evaluations	High Evaluations
High Simulator Hours	Low Simulator Hours
CMR/ Inexperienced	CMR/Experienced
Performance - Agent Off	
Low Performers	**High Performers**
WD	AWO
Low Simulator Hours	High Simulator Hours
Low Years of Experience	High Years of Experience
Low Number of Evaluations	High Number of Evaluations
	CMR/Experienced

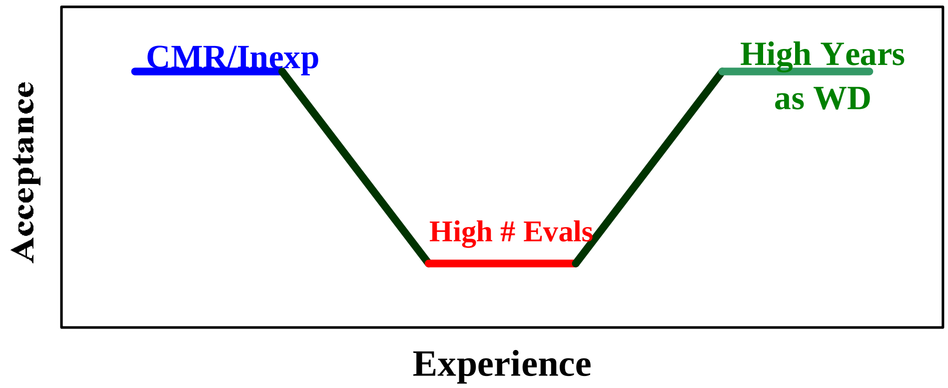
Discussion

Two intriguing findings have been presented here. First, differences in specific aspects of experience impact the degree to which weapons directors utilize agent assistance; and second the nature of the experience – performance relationship changes when agent assistance is available.

Experience and Agent Use

Results from the two analyses examining the relationship between experience and agent use offer unique and complementary perspectives. The first analysis suggests that participants with limited quantitative job experience accept a lower number of recommendations, than those with greater experience. The second analysis however, suggests participants with limited qualitative experience rely more on the agent than individuals with higher qualitative experience. The strength of the relationships identified in analysis 2, as illustrated by the 90% cross validation sample classification accuracy, suggest examination of qualitative variables in addition to quantitative variables can provide important information regarding the experience – agent use relationship.

Analyses examining quantitative and qualitative experience variables as predictors of agent use suggest something of a transformation, or evolution, taking place in the use of agent technology relative to experience (see Figure 1).



Results suggest that participants tend to rely on the agent when their experience level suggests they are in the early stages of skill acquisition (e.g., having completed fewer evaluations). That is WDs rely on the agent as a coach or trainer, demonstrating what *should* be done. In more advanced stages of skill and knowledge acquisition (having completed more evaluations) WDs rely less on the agent, because the WD “knows what to do”. Finally, participants who have been WDs for a long time, are confident in their abilities and use agent technology to augment performance. When agent recommendations are consistent with a WD’s own plan, accepting recommendations helps execute actions quickly and efficiently.

These findings have implications for the implementation of automated decision support in synthetic training environments. The study suggests that it may be useful to think of the agent as coach leading the participant to the correct action earlier in training, while the agent may play more of a information facilitator later in training.

Experience and Performance with Agent Assistance Versus Without Agent Assistance

The second set of analyses focus on the impact of agent availability on the experience – performance relationship. Findings suggest that the relationship between experience and performance changes somewhat when the agent is introduced into the simulation.

The most interesting finding of these two analyses involves the role of simulator time on differentiating between low and high performers. When no agent is available, high simulator time has a positive influence on performance. Conversely, in the agent on condition, high simulator time actually yields low performance, while low simulator time is associated with high performance. This suggests that the introduction of an automated agent poses a novel simulation situation in which previous simulator experience does not contribute to performance in the same way as it does when no agent exists. In the agent conditions, participants tend to rely more on tenure than simulator time.

Future longitudinal studies can tell us if this "novelty" effect diminishes over time and exposure to agent technology to the point where typical experience – performance relationship resume.

References

Coovert, D. L., & Riddle, D. L. (1999, November). The Application of Rough Set Theory to AWACS Weapons Director C3STARS Data. Final Technical Report. University of South Florida, Tampa, FL.

Riddle, D. L., Coovert, M. D., Gordon, T., King, T., Hoffman, K., Miles, D., Elliott, L., & Schiflett, S. (2000, April). Exploiting the data!: A rough set approach. Poster presented at the 15th Annual Society of Industrial and Organizational Psychology Conference, New Orleans, LA.

Pawlak, Z. (1982). Rough sets. International Journal of Computer Information Science, 11, 341–356.

Stoyen, A., Chiara, J., Schifflett, S., Elliott, L., Walrath, L., & Dalrymple, M. (June, 1999). Situational awareness measurement in AWACS weapons director teams. Paper resented at the Annual Conference on Situation Awareness, Pax River, MD.