Trust in Adaptive Automation:
The Role of Etiquette in Tuning Trust via Analogic and Affective Methods

Christopher A. Miller

Smart Information Flow Technologies
Minneapolis, MN
cmiller@sift.info

Abstract

In this paper, we begin by discussing a definition of trust and settle on one provided by Lee and See (2004). This definition emphasizes the nature of trust as an attitude toward the uncertain future actions of an agent. Some important implications of this definition for adaptive automation systems are discussed including (a) trust is not synonymous with user acceptance; in fact, trust should be tuned to result in accurate usage decisions by operators, (b) trust becomes more important with complex, adaptive automation precisely because it becomes less plausible for a human operator to fully understand what and how the automation will operate in all contents, and (c) as an attitude, trust is produced and affected by methods other than rational cognition. In fact, Lee and See provide a model with three methods of tuning trust—analytic, analogic and affective—with special emphasis on their roles trust for adaptive automation. Of these, we argue that the latter two will be more important in human interaction with adaptive automation than they are with traditional automation. We define and discuss a method of tuning analogic and affective trust: the "etiquette" of human interaction with automation. We provide examples from two recent projects, one involving a laboratory experiment and the other involving human interaction with automation in a realistic full mission simulation, which illustrate trust effects from etiquette-based design and system behavior manipulations.

1 Introduction

1.1 What is "Trust"?

Any discussion of trust, especially as it applies to human-automation interaction, should begin with a definition of terms. Trust is all-too-readily taken as synonymous with user acceptance, but repeated work over the past twenty (Lee and Moray, 1992, 1994; Lee and See, 2004; Parasurman and Riley, 1997) years has shown that both of these concepts are complex and intertwined. Users sometimes use automation that they are suspicious of (perhaps because they do not have the time or workload capacity to do otherwise) and sometimes do not use automation they believe is competent (perhaps because they enjoy doing the job themselves).

Lee and See (2004) in their recent comprehensive review of trust literature in the field of human factors, identify trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability." (Lee and See, 2004, p. 51). This definition is based on a careful analysis of a vast number of studies and experiments of human operators trust-related behaviors and I will use it in the remainder of this paper. It also conveys several subtle distinctions that are important to both the study of trust in human-machine interactions and in designing systems so that *accurate* trust results:

- Trust is an attitude, which means that it is a response to knowledge or belief about world states, but it is not, itself, those beliefs. Even less so is it a decision based on those world states. Many other factors may intervene to produce automation usage or non-usage decisions and, while trust is an important element in those decisions, it is far from the only one.
- Trust is an attitude, which means that, in part, it is affective. Trust and mistrust produce feelings about the agents they are directed at, and we base future trust in part on such feelings.
- Trust is ego-centric in that it is centered on an interpretation about an agent's ability and willingness to help me achieve my goals. In this sense, it is at least potentially separable from basic knowledge about

how and agent works or what it does in an abstract sense, though if I know these things, I can generally infer (at the cost of cognitive effort) whether or not the agent is likely to prove helpful to me in a specific set of circumstances and, hence, whether or not to "trust" it in those circumstances.

- Trust plays its largest role in 'situations characterized by uncertainty and vulnerability". Trust plays a smaller role to the degree that the situation is well understood and predictable. This distinction is critical when discussing trust in highly complex and adaptive automation systems and will be discussed in more depth below.

1.2 Why trust matters (especially) for Adaptive Automation

If trust is an attitude about the future, uncertain behavior of an entity in context, then it is very nearly irrelevant to speak of "trust" in a thoroughly understood and predictable system. I may say that I "trust" that an apple will fall to the ground if I drop it because, after all, there is always some uncertainty in the world, but this is hardly in the same league as saying that I trust that my shares of Company X will increase in value on the stock market. Trust is important precisely to the degree that knowledge and certainty about future behaviors is absent.

When we fail to comprehend all the causal mechanisms associated with an observed behavior, or fail to be able to track them in a timely fashion in context, we are increasingly forced away from an analytic understanding or prediction of an agent's behavior and instead are forced to rely more heavily on trust. As automated systems become ever more complex, their behaviors become ever less predictable and even comprehensible to their users. This is likely to be even more prevalent with adaptive automation systems whose very nature is not to exhibit the same behavior all the time and, therefore, which necessarily introduce another dimension of uncertainty into the human + machine system. Worse, is the fact that we want automation support (and, especially, augmented cognition support) in exactly those situations were our basic human cognitive capabilities are inadequate—either because they are limited in their range of knowledge or in the speed with which we can use them to analyze and decide, or both. Adaptive automation systems should, where possible, behave in a predictable fashion and users should be trained to understand how they will make their behavior decisions (Miller and Hannen, 1999), but this will not be possible to the degree that the automation reduces user workload by removing some elements of decision and execution authority from him or her (Miller and Parasuraman, submitted; Miller, Funk, Goldman, Meisner and Wu, this volume).

Trust matters especially for adaptive automation precisely because it will be largely unreasonable to expect human operators to know (by analytically reasoning through the same cognitive processes as the automation), in all circumstances, whether the adaptive automation is behaving in a correct manner or not. Instead, operators will need to have *accurately tuned* trust as to whether to accept or reject behaviors and recommendations provided by adaptive automation. If we are to assume that the human operator should remain in overall charge of operation with the adaptive automation in a subservient, aiding role, then this puts the designers, and ultimately, the operators, of adaptive automation systems in the position needing to understand how trust is developed and used.

2 Methods for Trust Tuning

Lee and See (2004) identify 3 alternate routes that humans use to develop and tune their trust: analytic, analogic and affective methods. In each case, the method affects or produces the 'attitude about the future, uncertain behavior of an entity' as described above, but the route to achieving that effect is different, making use of different cues, knowledge and different degrees of cognitive processing.

Analytic methods involve a detailed understanding and rational assessment of the mechanisms by which the entity produces its behaviors. Analytic methods assume rational decision making on the basis of what is known about the motivations, interests, capabilities and behaviors of the other party. In other words, I *reason*, taking into account uncertainty, about your likely future behaviors based on what I know about your motives and goals. I may trust you to help me write a paper because I believe that you are motivated to publish good work (and get credit for it).

Analogic methods involve using observable cues to infer broader category membership in a group or context and then to apply trust assumptions for the group as a whole to the newly encountered individual. Analogic trust can also be based on the word or endorsement of intermediaries whom we trust. We trust bankers to behave in responsibly with regards to our money in large part because they are a member of a category that is hired, trained, supervised and watched over to ensure that those behaviors are maintained. When we encounter an individual behind the counter in a bank, we are comfortable handing them our money largely on the basis of this analogic reasoning about their category membership, not because we deeply reason through their motivations with regards to our money.

Affective methods are based strictly on the affect generated by and toward the entity. This method accounts for the frequently recorded and observed finding (e.g., Cialdini, 1993) that we tend to trust those people and devices that please us more than those that do not. While affective trust tuning may be used as a shortcut for detailed reasoning in the analytic or analogic senses, it is nevertheless an effective method of managing our time and attention under conditions of cognitive overload. It is always unpleasant to have someone thwart our goals or plans, and the correlation of unpleasant feelings with thwarted plans is generally stronger than with outcomes that ended up being good for us. Thus, in conditions of uncertainty and especially when there is no time or attentional capacity available to perform more detailed reasoning about the situation, it may well be an effective heuristic to assume that, if something or someone is irritating, it does not have my best interests at heart and is, therefore, not to be trusted.

A interesting and important final element of Lee and See's model is the realization that there is a temporal element to trust building. It takes time to acquire, whether through experience, training or attending to hearsay and the experiences of others, the knowledge required to use either the analogic or analytic methods of trust tuning. When experiencing a new person or a novel system for the first time, with no background knowledge about the agent's motivations, behaviors, or group memberships, the only information a person may have about whether or not to trust it will be affective information. Furthermore, if the affect is negative enough to prompt a strong "do not trust" response, then no further information will be gathered. Analogic trust, in turn, requires less knowledge gathering and cognitive processing than does analytic trust and can, again, serve as a kind of hurdle to further experience. In other words, if a person or system does not do a reasonable job of providing appropriate cues to achieve at least moderate levels of affective and analogic trust, s/he/it may never have a chance to build analytic trust.

3 Applying Trust Tuning Methods to Adaptive Automation

The three approaches to building or tuning trust that Lee and See (2004) lay out can be used to form a coarse taxonomy of approaches to achieving accurate trust in adaptive automation. Furthermore, as noted above, the framework emphasizes the importance of analogic and affective approachesto trust tuning—and, as we will see below, these methods take on particular importance in complex, high criticality systems.

In the case of complex automation, I will generally have a less than complete knowledge of the processes and knowledge by which the automation operates. This is particularly true of automation which is performing millions of computations in a highly time pressured environment, such as much decision support and control automation in high criticality domains. Indeed, one of the prime reasons automation has been incorporated into such systems is because it can either do things that a human simply cannot do, or can do them faster than the human can. Nevertheless, operators are still expected (and generally trained) to understand the basic "motivations" of the automation—that is, the reasons why it was built—and the primary factors on which it bases its conclusions. In this sense, then, analytic trust approaches may not demand a detailed thorough thinking through of the actually computational methods by which automation produces the behaviors it does, but instead an application of general heuristics or presumed intentions.

3.1 Analytic approaches

Simplified analytic approaches provide some method of regularizing, "chunking" and/or constraining (and therefore simplifying) the potential behaviors of the agent. A familiar, if far from simple, example of this is Azimov's "laws" of robotics—a simplified set of heuristics that, at least in Asimov's stories, nevertheless served as a highly abstract view of how his robots actually produced their behaviors. Robots were deemed to be motivated 'never to harm a human or, through inaction, allow harm to come to a human' and this, in turn, served as an element in explaining their behaviors. A human in Asimov's fictional world could "trust" a robot even without completely understanding its thought processes because s/he knew that it would, in turn reason and behave in accordance with that motivation.

A somewhat more practical and current example of simplifying analytic trust tuning is represented by Taylor's (2001) Pilot Authorisation and Control of Tasks (PACT) approach to human-machine interaction whereby the human constricts the set of alternative behaviors available to cockpit aids in an initial "PACT" or contract which coarsely predefines acceptable ranges of automation behavior. Later automation behavior can therefore much more readily be predicted and evaluated against the framework established in the PACT and, thus, analytic trust (positive or negative) is more easily produced even for a very complex and adaptive automation system. Furthermore, the human operator has an enhanced awareness of the system's ability to behavior in accordance with the PACT precisely because the operator provided the behavior constraints in the first place, they were not something s/he had to learn. A similar benefit is sought in our own delegation and Playbook™ approaches to human-automation interaction (Miller, Pelican and Goldman, 2000; Miller, Goldman, Funk, Wu & Pate 2004; Miller, Funk, Goldman, Meisner, Wu, this volume).

3.2 Analogic approaches

Analogic methods of engendering trust for automation may be as simple as the knowledge that the automation was produced by a "trustworthy" company or approved by an agency known to be familiar with the operator's domain. Similarly, user communities and their accumulated impressions and "lore" about how a complex piece of automation behaves can place a very large role in building operator trust or mistrust.

Another way in which automation can tune trust via analogic methods is via elements of the "persona" it assumes—and encourages the operator to perceive in it. If automation uses the professional "jargon" of a highly trained domain (for example, the cadence, syntax and particular vocabulary items of Air Traffic Control), then it encourages operators to assume, via analogic trust reasoning, that it has the competence and range of reasoning abilities of other agents (primarily human air traffic controllers) who use such language. Such an approach might result in over-trust in the system if, in fact, it does not have those capabilities.

Note that the "embodiment" of the system in an animated character (a la Cassell, Sullivan, Prevost and Churchill, 2000) is not required to achieve this effect—simple language usage or the presentation of information in styles unique to a specific domain might be sufficient to trigger analogic trust, whether warranted or not. On the other hand, embodiment and personification open many additional channels (i.e., voice, gesture, movement, appearance, facial expressions, etc.) via which the system's interface can either succeed or fail in conveying group membership and therefore triggering analogic trust mechanisms. Cassell and Bickmore (2004) report a particularly complex, embodied real estate sales agent that manipulates conversational small talk in order to build enough social trust to enable it to introduce personally threatening topics such as income levels and expected family size as a part of its attempts to suggest real estate properties an individual or couple might be interested in.

3.3 Affective approaches

Affective methods involve essentially anything which results in producing positive or negative affect toward the system. While we might like to believe that operators of highly complex equipment in high criticality domains are above being influenced by "interpersonal" annoyances (or, conversely, pleasing layout and design), there is increasing evidence that this is simply not the case (cf. Norman, 2004). Indeed, much of the philosophy behind techniques of Crew Resource Management (e.g., Weiner, Kanki and Helmreich, 1993) are based on understanding and managing these "etiquette" and "style" variables in human-human cockpit interactions.

4 The Importance of Automation Etiquette

As Lee and See (2004) themselves point out, one interesting approach to achieving analogic and affective trust in complex systems is the "etiquette" that they exhibit. By "etiquette", we (Miller, 2002; 2004) mean the largely unwritten codes that define roles and acceptable or unacceptable behaviors or interaction moves of each participant in a common 'social' setting, that is, one between actors which are presumed to be both complex and reasonably understood according to intentional assumptions (Dennett, 1989). Etiquette rules create an informal contract between participants allowing expectations and interpretations to be formed and used about the behavior of others. As such, they can serve an analogic function in human-automation interaction—for example, the use of familiar domain jargon in advice or reporting can serve as a indicator that the agent using it (whether human or automation) falls into the category of "those experienced in the domain" and should, therefore, be accorded the trust reserved for those in that category. Because etiquette behaviors are also intimately tied to what we traditionally interpret as "polite" or "rude" behaviors (cf. Brown and Levinson, 1987), automation etiquette also holds the power to produce affective reactions in users, both positive and negative.

As Nass has pointed out (Reeves and Nass, 1996), humans are highly prone to interpret computer behaviors according to the same scripts or schema that are commonly used for human-human interactions. As discussed above, the sheer complexity of adaptive automation systems provides reason to believe that the role of etiquette in tuning trust via analogic and affective methods may be even larger than it is in less complex systems. We are beginning to see some examples that trust, user acceptance and even human + machine system performance are affected by etiquette variables. We will briefly discuss two of these below.

4.1 Interaction expectations and their role in tuning trust: Rotorcraft Pilot's Associate example

The U.S. Army's Rotorcraft Pilot's Associate (cf. Colucci, 1995), a highly complex adaptive automation system designed to assist the two pilots of an advanced attack-scout helicopter, operated by inferring ongoing and necessary tasks and then adapting information presentations and automation behaviors to best support those tasks. In designing the information management and task allocation aspects of the RPA, we noted (Miller and Hannen, 1999) that rotorcraft crews are trained to communicate with each other about ongoing and future tasks. By some estimates, nearly a third of the crew members' time is spent in crew coordination activities—particularly in maintaining shared awareness and coordination over who is doing what task when. We suspected that for the associate not to have this capability—that is, not to behave in accordance with the expected etiquette for the task domain-- might make it seem less trustworthy (via analogic means). Partly in an effort adhere to the associate metaphor in this domain, to provide an 'associate' that would be a good team player, we attempted to give the RPA some of this capability as well (Miller and Funk, 2001). The RPA cockpit design included an innovative, task-based display to provide the crew with both insight into, as well as some degree of control over, RPA's understanding of their intent. This "Crew Coordination and Task Awareness" display consisted of four small LED buttons located in the upper portion of each pilot's main instrument panel. The buttons report, in textual form, (1) the current inferred high-level mission context, (2) the highest priority current pilot task, (3) the highest priority associate task, and (4) the highest priority copilot task. Pressing these buttons permitted either pilot to override RPA's current inferred tasks and assert new ones (from an automatically scrolled list of higher-level tasks from the overall task network) via a single push button input.

Figure 1 shows the RPA cockpit simulation created for evaluation trials at the Boeing Company in Mesa, Arizona. The location of the Crew Coordination display is circled in Figure 1 and an enlargement and interpretive sketch of the display is provided for clarity.

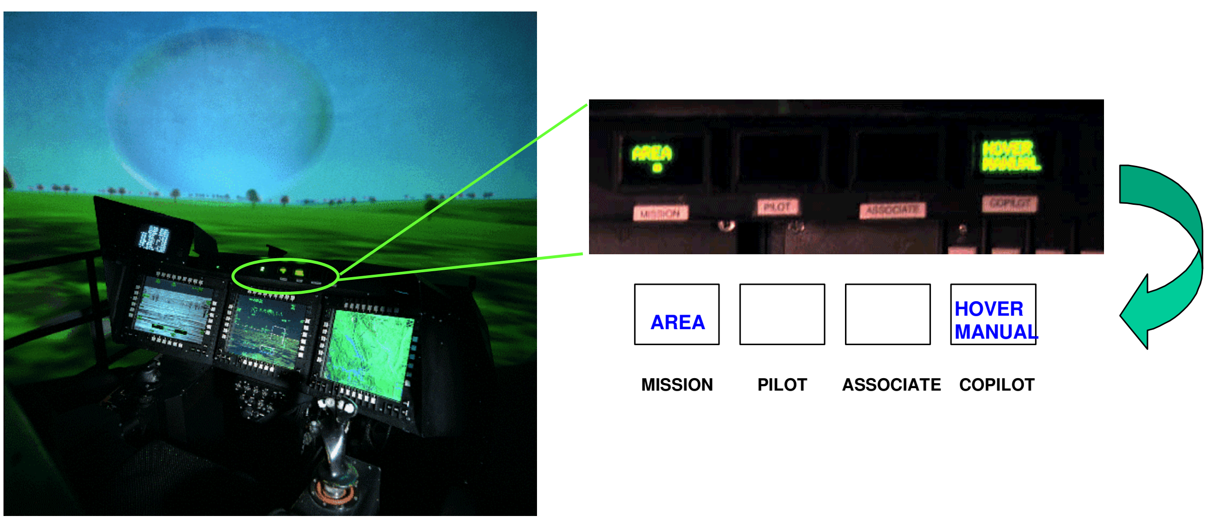

Figure 1. Location and appearance of the Crew Coordination and Task Awareness display.

While pilots were not asked explicitly about their trust in the RPA, pilot's acceptance of the LED Crew Coordination and Task Awareness display was high, as shown in Table 1. In spite of the perceived occasional inaccuracies in RPA's task inferencing mentioned above, and in spite of some pilot complaints about inadequate training in their use, most pilots found the individual task reports of the LED buttons 'Of Considerable Use' or 'Extremely Useful'. This provides some supporting evidence that our inclusion of a capability for the crew to interact directly with the associate's assumptions about active tasks was a capability that pilots welcomed—and that may have served to improve their overall impressions of CIM's capabilities and usefulness.

Table 1. Perceived usefulness of the LED Task Awareness Display (where 4.0='Of Considerable Use' and 5.0='Extremely Useful'.

LED Button for:	Score
Mission Task	4.4
Pilot Task	4.3
Copilot Task	4.3
Associate Task	4.0

The subjective pilot response data obtained from a series of full mission simulation trials (cf. Miller, 1999; Miller and Funk, 2001) suggest that the RPA system we designed and implemented generally met pilot expectations, contributed to perceived pilot effectiveness, reduced workload and gained pilot acceptance. Pilot's clearly felt that RPA provided benefit. Perfection in aiding and tracking pilot intent was not a prerequisite to the levels of acceptance we gained. To say what aspects of RPA design were ultimately responsible for these levels of acceptance is somewhat difficult, but we suspect that two attributes played a significant role. First, RPA was designed to provide high degrees of predictability, even at the cost of flexibility in managing displays in some instances. In some ways, RPA was behaving like a new member of the crew—trying to avoid doing unexpected things and making mistakes until it gained some acceptance. Second, the addition of an effective Crew Coordination and Task Awareness display may have contributed to pilots' willingness to tolerate these occasional 'mistakes' on RPA's part. Again, from the crew's perspective, RPA was adhering to expected crew etiquette and behaving like an acceptable junior member of the crew. Instead of jumping ahead and taking an action when it thought it knew what was going on, RPA generally reported what it thought was happening, followed the pilots' lead, and gave the crew a chance to correct it before it made major mistakes.

4.2 Experimental manipulation of etiquette behaviors: Trust, usage and performance impacts

While the RPA example cited above was for a realistic system tested in a full mission simulation, the results of adapting an interface to support analogic trust tuning are not conclusive. In this section, we summarize the results of a laboratory experiment (Parasuraman and Miller, 2004) in which etiquette variables were directly manipulated and the resulting impact on trust, automation usage decisions and performance we observed. We were interested in whether "good" (i.e., pleasing) etiquette would compensate for poor automation reliability and result in increased usage decisions and whether "poor" (i.e., displeasing) etiquette would wipe out the benefits of good automation reliability.

From the vast range of human etiquette behaviors, we chose to concentrate on a single dimension we label communication style. Communication style referred not to the specific wording of communications (which was held constant across conditions), but to the "interruptiveness" and "impatience" of delivering text messages. This was chosen as an etiquette dimension available to even fairly primitive and non-personified automation interfaces.

We tested 16 participants (both general aviation pilots and non-pilots) on a flight simulation task, the Multi-Attribute Task (MAT) Battery, which has been used extensively in prior high-criticality automation research (Parasuraman & Riley, 1997). The MAT incorporates primary flight (i.e., joystick) maneuvering, fuel management, and engine monitoring/diagnosis task. Participants always performed the first two tasks manually to simulate the busy operator of a high-criticality system. Intelligent automation support, modeled after the Engine Indicator and Crew Alerting System (EICAS) common in modern automated aircraft, was provided for the engine monitoring/diagnosis task. Automation monitored engine parameters, detecting potential engine faults, and advised participants when and what to examine to diagnose the faults. For example one advisory message was: "The Engine Pressure Ratio (EPR) is approaching Yellow Zone. Please check. Also, cross-check Exhaust Gas Temperature (EGT). There is a possible flame out of Engine 1."

"Good" automation etiquette involved a communication style that was "non-interruptive" and "patient" while poor etiquette was the opposite. In the non-interruptive case, advice was provided after a 5-second warning and not at all when the operator was already doing the requested action. Non-interruptive automation was "patient" in that it would not issue a new query until the user had finished the current one. Interruptive/impatient automation, by contrast, provided advice without warning and came on when the user was already querying EICAS . Automation also "impatiently" urged the next query before the user was finished with the current one.

These good and poor communication styles were crossed with the effects of automation reliability. Two levels of reliability were chosen: low, in which automation provided correct advice 60% of the time and high (80%).

We were primarily interested in the effects of etiquette and reliability on users' performance and on their rated trust in the automation. The percentage of correct diagnoses of engine malfunctions in all four conditions are shown on the left side of Figure 2. As expected, user diagnostic performance was significantly ($p < .01$) better when automation reliability was high (80%) than low (60%). Less obviously, good automation etiquette significantly ($p < .05$) enhanced diagnostic performance, regardless of automation reliability. Perhaps most interestingly, the effects of automation etiquette were powerful enough to overcome low reliability ($p < .05$). As the dotted line on the graph on the left side of Figure 2 indicates, performance in the low reliability/good etiquette condition was almost as good as (and not significantly different from) that in the high reliability/poor etiquette condition. These findings on diagnostic performance were mirrored in the results for user trust, shown in the center of Figure 2. High reliability increased trust ratings significantly, but so did good automation etiquette—to an even greater degree. Finally, good communication etiquette seems to have had a consistent effect in reducing perceived user workload in both high and low reliability conditions—as illustrated in the graph on the right side of Figure 2.

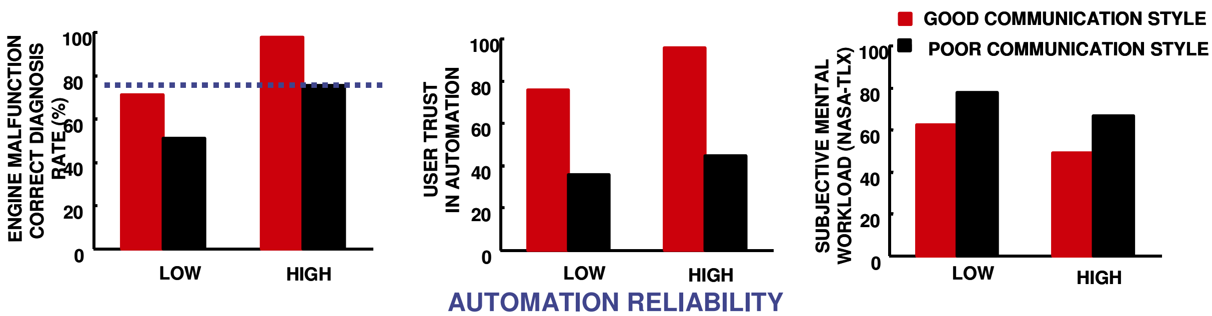

Figure 2. Results of etiquette manipulations in automation diagnostic recommendations presentations on correct performance (left), user trust (center) and subjective workload (right).

A possible objection to these findings is that any interruption might be expected to degrade user performance. Where the findings above due to the "rudeness" of this automation—which interrupted to "nag" (that is, to provide directives that users were working on, thereby effectively telling the operator's to 'hurry up')—or due to the simple interruption itself? To answer this question, Parasurman ran a control group of four participants using interruptions which were non-specific in content—for example, "Maintaining primary flight performance is important, but do not forget to check engine parameters for possible malfunction." These interruptions were varied in their intrusiveness as above—they were either preceded by a warning and not offered if the user was engaged in diagnosis (non-intrusive) or were given with no warning regardless of user activity (intrusive). Under these conditions, intrusiveness had no significant effect on correct diagnosis of engine malfunctions or user trust ratings. Thus, in contrast to the main experiment, less rude, non-specific interruptions were more easily ignored and did not adversely affect user-system performance or trust, implying that the adverse effects on trust and performance observed in the first experiment were the result of the rude "nagging" reminders to do work already being performed rather than the intrusions per se.

These results provide strong, if preliminary, evidence for the impact of automation etiquette on both user performance and trust in using an intelligent fault management system to diagnose engine malfunctions. While we did not collect any measures to indicate whether the trust (and, arguably, the performance) impacts were due to analogic or affective mechanisms, it would seem that affective methods are the most likely. We can speculate that analogic cues in the instructions that were provided would, if anything, have tended to *increase* trust since messages were given in the form and using the language of existing (and, arguably, trusted) cockpit displays and vernacular. On the other hand, interruptions and "nagging" are frustrating and unpleasant and, by the simplified connections which exist at the analogic level, those who do such things are unlikely to have my best interests at heart and are, therefore, not to be trusted or heeded.

These results also clearly show that building reliable automation may not be enough for overall human + machine system efficiency: both user diagnostic performance and trust were lowered by poor automation etiquette even when the reliability of automation advice was high. Furthermore, there is some hint that good etiquette can compensate for low automation reliability since users did better with well behaving automation (good etiquette) even if the automation was unreliable. Some may find this result disturbing, since it suggests that developing robust, sensitive, and accurate algorithms for automation, a challenging task at the best of times, may not be necessary so long as the automation "puts on a nice face" for the user. This would clearly be a violation of the dictum that trust should be tuned to reflect the actual level of performance of the automation, but it might nevertheless be tempting, especially in a commercial products setting. While some "sins" might be papered over in this manner, it was clear that the *best* user performance (and the highest trust) was obtained in the high reliability condition in which the automation also communicated its advice to the user in an etiquette-friendly way. Hence, good automation that behaves well should always be the highest goal. We have shown, however, that automation trust can be tuned to some extent by the etiquette that the automation exhibits and this may prove a useful tool in crafting automation behaviors for optimal use.

5 Conclusions

We have argued above that trust is not identical to user acceptance, but instead should be regarded as a process of tuning the user's attitudes toward automation (and, therefore, willingness to accept it's advice and behaviors) to comply with the actual performance of the automation. As designers of complex, high criticality systems, we must be aware of the factors that influence that trust for better and for worse. The goal is not to get users to accept our systems, but rather to get them to accept and make use of those systems when they will be helpful and to reject them when they will not.

The conceptual framework of Lee and See (2004) provides us with a taxonomy of methods for influencing trust and a model of how they interact. For complex, high criticality automation, and especially for adaptive automation which operates using augmented cognition technologies, there are many reasons to believe that analogic and affective methods of trust tuning will play a conceptually greater role than do more memory, knowledge and experience intensive analytic methods. A promising line of research and design adaptations to tune trust via affective and analogic methods is via the etiquette exhibited by the automation. We have provided two examples, one a laboratory experiment and one a high fidelity simulation of a realistic work environment, of instances in which we believe that the etiquette influenced the trust of operators of advanced automation equipment. Much future work remains to be done to refine our understanding of etiquette and its impact on trust, but this work lays the foundations.

6 Acknowledgements

Work on the Rotorcraft Pilot's Associate was funded by the U.S. Army AATD, contract number DAAJ02-93-C-0008. The author would like to thank Raja Parasuraman for his work in conducting and analyzing the etiquette experiment described above, and John Lee and Katrina See for their comments on earlier drafts of this paper.

7 References

Brown, P. & Levinson, S. (1987). *Politeness: Some Universals in Language Usage.* Cambridge,UK; Cambridge Univ. Press.

Cassell, J. and Bickmore, T. (2004). Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents. *User Modeling and User-Adapted Interaction. 13*(1): 89-132

Cassell, J., Sullivan, J., Prevost, S. and Churchill, E. (2000). *Embodied Conversational Agents.* Cambridge, MA; MIT Press.

Cialdini, R.B. (1993). *Influence: Science and Practice. 3rd Ed.* New York; Harper Collins.

Colucci, F. (1995). Rotorcraft Pilots' Associate update: The Army's largest science and technology program, *Vertiflite*, March/April, 16-20.

Dennet, D. (1989). The Intentional Stance, MIT Press; Cambridge, MA.

Lee, J., & Moray, N. (1992). Trust, control strategies, and allocation of function in human-machine systems. *Ergonomics*, *35*. 1243-1270.

Lee, J., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies, 40,* 153-184.

Lee, J. & See, K. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors, 46* (1), 50-80.

Miller, C. (1999). Bridging the Information Transfer Gap: Measuring Goodness of Information Fit. *Journal of Visual Language and Computation, 10,* 523-558.

Miller, C. A. (Ed.) (2002). *Working Notes of the AAAI Fall Symposium on Etiquette for Human-Computer Work.* Technical Report FS-02-02. Menlo Park, CA: American Association for Artificial Intelligence.

Miller, C. (Ed.), (2004). Human-Computer Etiquette: Managing Expectations with an Intentional Agent. Special section in *Communications of the ACM*, April, 2004.

Miller, C. and Funk, H. (2001). Associates with Etiquette: Meta-Communication to Make Human-Automation Interaction more Natural, Productive and Polite. In *Proceedings of the 8th European Conference on Cognitive Science Approaches to Process Control*. September 24-26, 2001; Munich.

Miller, C., Funk, H., Goldman, R., Meisner, J. and Wu, P. (this volume). Implications of Adaptive vs. Adaptable UIs on Decision Making: Why "Automated Adaptiveness" is Not Always the Right Answer.

Miller, C., Goldman, R., Funk, H., Wu, P. and Pate, B. (2004). A Playbook Approach to Variable Autonomy Control: Application for Control of Multiple, Heterogeneous Unmanned Air Vehicles. In *Proceedings of FORUM 60, the Annual Meeting of the American Helicopter Society*. Baltimore, MD; June 7-10.

Miller, C. & Hannen, M. (1999). The Rotorcraft Pilot's Associate: Design and evaluation of an intelligent user interface for cockpit information management. *Knowledge Based Systems, 12*, 443-456.

Miller, C. & Parasuraman, R. (submitted). Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control. Submitted for publication in *Human Factors*.

Miller, C., Pelican, M. and Goldman, R. (2000). "Tasking" Interfaces for Flexible Interaction with Automation: Keeping the Operator in Control. In *Proceedings of the Conference on Human Interaction with Complex Systems*. Urbana-Champaign, Ill. May.

Norman, D. (2004). *Emotional Design: Why we love (or hate) everyday things.* New York; Basic Books.

Parasuraman, R. and Miller, C. (2004). "Trust and Etiquette in High-Criticality Automated Systems". In C. Miller (Guest Ed.), special section on "Human-Computer Etiquette". *Communications of the ACM. 47*(4), April. 51-55.

Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39,* 230-253.

Reeves, B. and Nass, C. (1996). *The Media Equation.* Cambridge, UK; Cambridge University Press.

Taylor, R.M. (2001). Cognitive Cockpit Systems Engineering: Pilot Authorisation and Control of Tasks. In R. Onken (Ed), CSAPC'01. *Proceedings of the 8th Conferences on Cognitive Sciences Approaches to process Control*, Neubiberg, Germany, September 2001. University of the German Armed Forces, Neubiberg, Germany.

Weiner, E., Kanki, B. and Helmreich, R. (1993). *Cockpit Resource Management*. London; Academic Press.