Leonard Adelman, Cedric Yeo, 
and Sheryl L. Miller

Understanding the Effects 
of Computer Displays and Time 
Pressure on the Performance 
of Distributed Teams

Introduction

Some studies (e.g., Urban et al., 1996) have shown that time pressure degrades team decision-making performance. In contrast, other studies have shown process changes but no effect on performance (e.g., Hollenbeck et al., 1997b). Using a longitudinal design, Adelman and colleagues (2003) found that teams adapted their processes in different ways, all of which were effective in maintaining high and constant performance until almost twice the initial time pressure level. Then different teams lost performance constancy as time pressure increased, depending on their process adaptations. The findings suggest that one should not necessarily expect to find a process–performance relationship with increasing time pressure; team processes change early and in different ways, but performance can be maintained under high time pressure levels depending on the effectiveness of the team’s process adaptations. These findings are consistent with Entin and Serfaty’s (1999) team adaptation model and, as discussed in Adelman, Henderson, and Miller (2001), Brunswik’s (1952) concepts of vicarious functioning (process intersubstitutability) and performance constancy.

Also consistent with Brunswikian theory (Hammond & Stewart, 2001), one can argue that one’s results depend on the task, and in the case of Adelman et al. (2003), the features of the human–computer interface. In a follow-up study, Miller and colleagues (2000) demonstrated this point, but only to a limited extent. They found that a “perceptual” interface designed to make it easier for team members to see what decisions they had made permitted teams to maintain the percentage of decisions they made as time pressure increased. However, contrary to prediction based on previous research (e.g., Balzer, Doherty, & O’Connor, 1989), a “cognitive” interface designed to provide feedback on team members’ decision processes had no effect on decision accuracy. As time pressure increased, team decision accuracy declined regardless of the type of human–computer interface.

We were unable to explain this finding fully at the time of the study. We can now do so using a Brunswikian theory of team decision making and the lens model equation. The next section presents the relevant theoretical concepts; then we describe the experiment and new analyses. The final section discusses the strength and limitations of the research.
Brunswikian Theory and Lens Model Equation

Brehmer and Hagafors (1986) extended Brunswik's lens model to a "multilevel lens model" to represent staff functioning. Figure 4.1 shows the multilevel lens model for our aircraft identification task. The "outcome" (correct answer) is an aircraft's true hostility level, which two subordinates and a leader are trying to judge using six pieces of information ("cues"), such as how the aircraft responded to an Identification Friend or Foe (IFF) query. All three team members received the three cues represented by the solid lines, plus one other cue. The unique cues are represented in Figure 4.1 by the dashed lines at the top, middle, and bottom of the figure. Conceptually, one subordinate received the cue shown at the top; the other subordinate received the cue shown at the bottom. Only the leader received the IFF cue, which was the most important and is shown in the middle of the figure. Team members had to send their unique cue information to each other to have all the information about an aircraft. The heavy, dotted lines going only from the subordinates to the leader represent the process of them sending their judgments ("recommendations") to the leader.

Using the multilevel lens model, Hollenbeck et al. (1995) and Ilgen et al. (1995) developed the multilevel theory of team decision making. They showed that team decision accuracy depends on the effective distribution of information among the team (team informity), the accuracy of subordinates' recommendations (staff validity), and the leader's ability to rely on the recommendations of the more accurate subordinates (hierarchical sensitivity). Team informity, staff validity, and hierarchical sensitivity have been called "core constructs." Multi-level theory has predicted that the effect of noncore variables, such as team members' experience or system feedback, on leaders' decision accuracy should be fully mediated by the core constructs.

**Figure 4.1.** The multilevel lens model representing the task.
Research (e.g., Hollenbeck et al., 1995, 1998) has shown this to be largely the case with minimal variance in the leaders' decision accuracy being attributable to noncore variables.

If team members sent their unique cues to each other all the time in our task, then team informity would be perfect because all team members would always have all six cues prior to the subordinates and leader making their recommendations and decisions. If the leader also has the subordinates' recommendations, then his or her decision can be based on the cues and/or recommendations. By correlating each subordinate's recommendations with the outcomes, one can calculate the accuracy of each staff member (staff validity, \( r_m \)), as shown pictorially in Figure 4.1. Similarly, by correlating the leader’s decisions with the outcomes, one can calculate the leader’s (and team’s) level of accuracy (\( r_s \)). Using hierarchical regression, one can assess the extent to which the leader is using the cues and recommendations. If the leader is using the recommendations of the more accurate staff member, his or her “hierarchical sensitivity” will be high.

Multilevel theory has not, however, used the lens model equation (LME) developed by Hursch et al., (1964) to decompose the leader’s accuracy into its distinct conceptual components. This omission may have occurred because multilevel theorists (e.g., Hollenbeck et al., 1995) usually defined accuracy as the mean absolute error of the team's decision. However, Brehmer and Hagaforis’s (1986) original formulation of the multilevel lens model defined the leader’s accuracy in terms of the achievement correlation, typically used in Brunswikian social judgment theory research (e.g., Brehmer & Joyce, 1988; Cooksey, 1996; Hammond et al., 1975). Although Hollenbeck et al. (1997a) used the achievement correlation, they did not decompose it into its component parts using the LME. Bisantz and coauthors (2000) did so using a dynamic aircraft identification task, but they focused on individual not team decision making.

Because both the correct answers and decisions were predicted by a linear equation in the current study, we adopted Tucker's (1964) version of the LME, presented in Equation 4.1:

\[
(4.1) \quad r_s = GR_e R_e
\]

where

- \( r_s \) = decision accuracy or “achievement,” defined operationally as the correlation (Pearson product-moment) between a person's judgments and the correct answers;
- \( G \) = model similarity or “knowledge,” defined operationally as the correlation between the predictions of two models, the best-fitting (least-square) linear model predicting the person’s judgments and the best-fitting (least-square) linear model predicting the correct answer, which the participants were trained to use;
- \( R_e \) = judgment model predictability or “cognitive control,” defined operationally as the correlation between a person’s actual judgments and their predicted judgments based on the best-fitting, linear model; and
- \( R_e \) = task model predictability, defined operationally as the correlation between correct answers and the predicted correct answers based on the best linear model.

This chapter shows how Brunswikian theory and the LME can explain the previously unexplained effects of computer displays and time pressure on the achievement (\( r_s \)) of distributed teams reported in Miller et al.’s (2000) experiment. For by using the LME and holding task predictability (\( R_e \)) constant, one can determine whether changes in leaders' achievement (\( r_s \)) was caused, on average, by corresponding changes in their knowledge of what to do (\( G \)) or their cognitive control in applying their knowledge (\( R_e \)). By using a path model to operationally define multilevel theory, one can determine mediating pathways between noncore variables (e.g., time pressure), core construct variables (e.g., informity), and the LME parameters (\( G \) and \( R_e \)) comprising leaders' achievement (\( r_s \)). An initial effort by Henderson (1999) suggested the potential viability of this approach but was limited by not using the steps described next to calculate \( G \) and \( R_e \) under the range of conditions of cue and recommendation availability.

**Experiment**

**Hypotheses**

The effectiveness of different human–computer interfaces was studied under increasing levels of time pressure. Because previous research (Adelman et al., 2003) found a main effect for time pressure, we again predicted a time pressure main effect. In
addition, we predicted an interface main effect and a time pressure by interface interaction. Previous research (e.g., Balzer, Doherty, & O'Connor, 1989; Hammond, 1971) showed that cognitive feedback about how one made judgments and about how one should according to task characteristics, improved achievement \( (r_a) \) because it improved knowledge \( (G) \) and cognitive control \( (R) \). Therefore, we predicted that teams using such an interface would have higher \( r_a, G, \) and \( R \), when averaged over all time pressure levels and be able to maintain higher achievement at higher time pressure than teams using other interfaces. (As noted earlier, Miller et al. [2000] found that teams using the perceptual interface described herein maintained decision quantity longer than teams using the cognitive interface. These results are not considered here. We only consider \( r_a \).

Participants

Twenty-four Army Reserve Officer Training Corps (ROTC) cadets from a suburban university volunteered to participate for seven two-hour sessions. The cadets were placed in eight three-person teams based on availability. An effort was made to equate teams by gender and cadet rank. By default, the highest-ranking cadet was the team leader. All team leaders were male because there were few female cadets in the ROTC program. Teams competed for prize money donated to the cadet’s fund-raising mission. (A ninth team was dropped when the leader was hospitalized due to an extended illness.)

Task

The task was a simulated air defense task representative of that by Hollenbeck et al. (1995), and modeled after the task of AWACS and Patriot air defense teams. It was designed for a leader and two subordinates dispersed geographically who communicate with each other only through their computer system. The leader and subordinates tracked multiple aircraft (called targets) on their radar screens, shared unique information about the target, and made decisions about targets’ threat level, as described generally in Figure 4.1. Although team members could request information, it was almost always sent without being requested. Therefore, information passing represented a “voluntary, cooperative act” (Adelman et al., 1986, p. 799) and a form of “implicit coordination” (e.g., Entin & Serfaty, 1999, p. 313). Although interaction was only possible through the system during the identification task, team members could talk face to face between trials.

Apparatus

The task was implemented using the Argus synthetic task environment (Schoelles & Gray, 2001). Figure 4.2 shows the basic interface. The first component is the radar screen on which aircraft appeared in different segments represented by concentric circles (left). The second component is the data display window, where cue values and recommendations appeared (upper right). The third and fourth components are the message inbox (middle right) and email system (lower right).

An ideal interaction involved the participant first “hooking” a target by clicking on its icon on the radar screen. The data associated with that target would appear in the upper right section of the display. All participants received the bearing and three cues (airspeed, course, and range) about each target. They also received a unique cue. Leaders received IFF, one subordinate received altitude, the other radar. Ideally, participants would send the unique cue to their teammates and receive cue information in return. Received information was presented in red in the data display window. Participants made a decision about the target’s threat level on a 1–7 (increasing threat) scale. For example, a 7 meant the target is certainly hostile; fire a missile to shoot it down.

The environmental model was determined by Equation 4.2, which participants were trained to use to calculate a target’s threat level. Participants were given target cue information in the appropriate metric and then transformed it into a three-point hostility-indicator scale (from 0 to 2) used for each of the cues in Equation 4.2. Participants then transformed the overall threat level, which could go from 0 to 14 using Equation 4.2, into the final seven-point decision scale. The task required two to three hours of training before reaching criterion, which was an achievement correlation \( (r_a) \) of 0.80 between a participant’s threat recommendations/decisions and the true threat levels for a set of targets.

\[
\text{Threat} = (2 \cdot \text{IFF}) + (1 \cdot \text{Speed}) + \\
(1 \cdot \text{Course}) + (1 \cdot \text{Range}) + \\
(1 \cdot \text{Altitude}) + (1 \cdot \text{Radar})
\]
**Experimental Conditions**

**Time Pressure ("Tempo")**

Time pressure was manipulated by increasing the tempo level, defined operationally as the number of new targets per minute appearing on the radar screen during a 15-minute scenario. Tempo was a within-subject variable; each team performed all seven tempo levels (0.8, 1.2, 1.6, 2.0, 2.4, 2.8, and 3.2). Only three levels (1.2, 2.4, and 3.2) were used in the analysis presented herein because of the slow, laborious effort required to develop the judgment models for calculating G and Rₜ, as will be described. Tempo 1.2 had 18 targets, tempo 2.4 had 36 targets, and tempo 3.2 had 48 targets. Two judgments were required per target, so the number of required judgments ranged from 36 to 96. (Note: There were two runs at each tempo level using isomorphs of the basic scenario. The analysis reported herein is always for the second run. Analysis of variance [ANOVA] results reported replicated those using both runs for all tempo levels.)

**Interface**

There were three human-computer interface conditions.

- Original: Targets always appeared as yellow circles on the radar screen.
- Perceptual: Used color to indicate a target's status and shape to indicate remaining time. Each target initially appeared on the screen in red. When the target was hooked, it turned yellow, and when a recommendation/decision was made, it turned blue. Thus, color enabled operators to determine which targets had been dealt with without having to rely on memory. Also, targets changed shape when they approached the concentric circles to visually indicate that a judgment was required before the target moved into the next sector. Targets turned into a rectangle when there was 30 seconds remaining for a judgment, they turned into a triangle when there were only 15 seconds left. Thus, shape...
served as a cue to help operators determine which targets to prioritize.

- Original plus cognitive feedback: Teams used the original interface to make judgments during each trial, and received cognitive feedback after it. The leaders’ feedback included the number of decisions and recommendations made by them and each of the two subordinates, and their (and subordinates’) decision accuracy (achievement correlations, r_a) and performance. (Note: The analysis reported herein focused only on decision accuracy, r_a. We note here for completeness that performance was defined as \[ \left( \frac{17 - \text{Absolute Value (Truth - Decision)}}{17} \right) \times \text{Decision} = 0.0 \] for all decision opportunities for which no decision was made, averaged over all decision opportunities.)

Leaders also were shown their relative agreement with (or dependency on) each subordinate’s recommendations. This was calculated by simultaneously regressing their decisions on each subordinate’s recommendations, and then dividing the absolute value of each beta weight by the sum of the absolute values of the two beta weights. Last, leaders were shown their relative cue weights, which was the ratio of the beta weight on each cue based on the best fitting judgment model for all six cues divided by the beta weight based on the task model they should have used. The ratio was 1.0 if they weighted a cue correctly, greater than 1.0 if they overweighted it, and less than 1.0 if they underweighted it. Except for learning the team’s overall performance, subordinates only received feedback about themselves. Teams usually finished examining their cognitive feedback within five minutes. (Note: Feedback assumed that team members had all six cues when they made their judgments. The analysis presented herein will indicate that was a mistake.)

Two teams used the original interface, three teams used the perceptual interface, and three teams used the original interface with cognitive feedback. (The team dropped from the study because of the leader’s illness used the original interface.)

**Experimental Design**

The experimental design was 3 Interface (Original, Perceptual, and Original + Cognitive Feedback) x 3 Time Pressure (1.2, 2.4, and 3.2 tempo levels) factor design. Interface was a between-subject variable; time pressure was a within-subject variable.

**Procedures**

Training was incremental over a two-week period. During the first week, the participants read a detailed task outline that described the aircraft cues and rules for making recommendations and decisions. Written instructions were supplemented with verbal instruction. Quizzes were administered at appropriate intervals to assess task learning; the participants were required to memorize all task cues and rules. When sufficient task proficiency was achieved, teams were shown a computer demonstration of the task.

Participants refamiliarized themselves with the instructions and interface during the second week of training. Each one sat at an individual workstation and practiced performing the task. They were seated in separate rooms and instructed not to talk during the task to simulate a distributed team. After individuals were sufficiently trained, teams were presented with low tempo scenarios (e.g., tempo 0.4 = 6 targets/15 minutes). Teams were required to achieve an achievement (r_a) level of 0.80 on these scenarios before they were allowed to proceed to the experimental trials.

Teams performed the experimental trials over the following five weeks. Tempos progressed each week, beginning with tempo 0.80 on the first week and concluding with tempo 3.2 on the last. Teams performed two scenarios at each tempo, each of which presented the same aircraft and flight paths, but in different ordered sequences so participants could not remember them. Teams could not talk during a scenario, but they could talk after them. They knew tempo would be increased, but they did not know future levels. Teams also knew they were working with different interfaces but not the differences and were not shown the three interfaces until after the study. Last, consistent with what they were told at the outset, the three best teams (one for each interface) were identified for commendation at the year-end military ball, which clearly motivated them.

**Dependent Variables**

There were two sets of dependent variables: those defined by multilevel theory, which are presented first, and those defined by the LME.
• Informity: Team Informity was the average amount of unique cue information (sent via email messages) to the three teammates. For example, if the three teammates always sent their unique cue to each other, then team informity would be 100%; if they sent two messages, on average, it would be 66%; and so forth. We also calculated the Leaders Informity, which was the average amount of information sent to the leader, and Staff Informity, which was the average amount of information sent to the two subordinates.

• Staff Validity: The achievement (or validity) of each subordinate’s (or staff member’s) recommendations was illustrated in Figure 4.1 and defined as the Pearson product-moment correlation between their recommendations and the correct answers (r). Fisher z-transforms are routinely taken in lens model research to normalize the correlation distribution, which is skewed at its upper end. Consequently, (overall) staff validity was defined operationally as the mean Fisher z-transforms of the r coefficients for the two subordinates.

• Hierarchical Sensitivity: measures how effectively leaders weight the recommendations of their more valid staff members. Hierarchical sensitivity was calculated using a six-step process. First, using an ordinary least squares regression, we separately regressed the leader’s decisions on each subordinate’s recommendations to obtain the beta weights indicating how much the leader weighted each subordinate’s recommendations. Second, we regressed the correct answers on the recommendations to obtain the beta weights indicating how much they should have weighted each subordinate’s recommendations. Third, we calculated the absolute value of the difference between the leader’s and the correct answer’s beta weights for each subordinate. Fourth, we multiplied this difference by the number of recommendations made by each subordinate to ensure that the more active subordinate received more weight in the measure. Fifth, we summed the weighted differences for the two subordinates and divided it by the total number of recommendations to obtain the hierarchical sensitivity score. Sixth, we took the logarithm of the hierarchical sensitivities because they were heavily skewed. (The skewness value was 2.62 times its standard error [1.259/0.481].) Lower values meant greater hierarchical sensitivity because there was less difference between the leader’s and the task’s beta weights.

The three terms of the LME used in the data analysis were achievement (r), knowledge (G), and control (R). Task predictability (G) was constant, approaching 1.0, except for rounding error. It is important to note here that the procedure used to calculate the best-fitting judgment models had to account for the fact that the leaders and subordinates did not always have all the information that was to be sent to them when they made their judgments. This was accomplished in three steps.

The first step was to determine what information each participant had when they made a judgment. For example, when a leader made each of his decisions, did he have a recommendation from each of the two staff members? Did he have six, five, or four cue values for that target? (They would have six if both staff members sent the values for their unique cues.)

The second step was to perform a simultaneous regression for each data set to calculate the best-fitting judgment model for that data set. Depending on the team, tempo, and the amount of data per partition, a leader might have two or more best-fitting models. For example, a leader might have a best-fitting judgment model for cases (i.e., targets) where there were two recommendations and six cue values (i.e., eight independent variables), and another model for cases where there were no recommendations and only four cue values (i.e., four independent variables). Subordinates only had models based on six, five, or four cues.

The third step was to calculate G and R. G was calculated by correlating the predictions of the task model with the predictions for the best-fitting judgment model for all cases for which there was a judgment. The task model was always based on the six cues. The judgment model was the one appropriate for a particular case (i.e., target). So, for example, a leader might have made decisions for some cases when he had two recommendations and six cues, other decisions when he only had four cues and no recommendations, and so forth. Consequently, in the first example, we used the best-fitting judgment
model based on two recommendations and six cues; in the second example, we used the best-fitting model based on only four cues, and so on. \( R \) was calculated by correlating all the predicted judgments (regardless of the model used to generate them) with the person's actual judgments for those cases. Fisher z-transforms were taken of \( G \) and \( R \) to normalize both correlation coefficients. (It is important to again note that \( r_a \) for the leader and \( r_m \) for each subordinate is fully decomposed into \( G \) and \( R \), given that \( R \) approached one.)

Results

First, we present the ANOVA results for leader achievement (\( r_a \)) and multilevel theory core constructs. Then, we present the ANOVA results for the leaders' LME parameters (\( G_L \) and \( R_m \)) and the path model linking them to the multilevel theory variables and the independent variables.

**Achievement and Multilevel Theory Core Constructs**

A 3 Interface (Original, Perceptual, Original + Cognitive Feedback) 3 Tempo (1.2, 2.4, 3.2) ANOVA was performed for each dependent variable. Interface was a between-subject factor and Tempo was a within-subject factor.

- **Leader Achievement (\( r_a \)):** The only significant effect was a Tempo main effect \( \left[ F(2, 10) = 8.07, p = 0.008 \right] \). Mean \( r_a \) values were 1.22, 0.92, and 0.82 for Tempos 1.2, 2.4, and 3.2, respectively.

- **Hierarchical Sensitivity (log):** There were no significant main effects, although the main effect for Tempo approached the traditional \( p = 0.05 \) level \( \left[ F(2, 10) = 3.29, p = 0.09 \right] \), with poorer hierarchical sensitivity with increasing tempo.

- **Staff Validity (\( r_m \)):** There was a significant main effect for Tempo \( \left[ F(2, 10) = 9.80, p = 0.006 \right] \). As tempo increased, staff validity (\( r_m \)) decreased (1.22, 0.91, and 0.87).

- **Informity: Results for team, leader, and staff informity are presented, in turn.**

- **Team Informity:** There were significant main effects for Tempo \( \left[ F(2, 10) = 14.90, p = 0.002 \right] \) and Interface \( \left[ F(2, 5) = 26.45, p = 0.002 \right] \). As tempo increased, mean team informity decreased (0.89, 0.70, and 0.51).

Mean team informity was highest for the original interface (0.87), and then the perceptual (0.74), and cognitive feedback (and original) interfaces (0.50).

- **Leaders' Informity:** There was only a significant Tempo main effect \( \left[ F(2, 10) = 4.96, p = 0.032 \right] \), means were 0.87, 0.71, and 0.56.

- **Staff Informity:** There were significant main effects for Tempo \( \left[ F(2, 10) = 11.05, p = 0.007 \right] \) and Interface \( \left[ F(2, 5) = 23.15, p = 0.003 \right] \). Mean values for tempo were 0.89, 0.70, and 0.50, respectively. The mean was highest for the original interface (0.88), then the perceptual (0.76), and finally the cognitive feedback interface (0.45).

**Lens Model Equation Parameters**

We first present the results of the 3 Interfaces (3 Tempo ANOVAs for leaders' \( G_L \) and \( R_m \) values). The ANOVA results for the subordinates were almost identical to those of the leaders and are not presented here.

The only significant ANOVA effects were Tempo main effects. The mean values for Tempos 1.2, 2.4, and 3.2, respectively, and the ANOVA statistics for the leaders' \( G_L \) and \( R_m \) parameters were:

- **Mean \( G_L \):** 1.30, 0.96, and 0.86, \( F(2, 10) = 5.883, p = 0.02 \).
- **Mean \( R_m \):** 3.24, 2.33, and 2.17, \( F(2, 10) = 8.65, p = 0.007 \).

Figure 4.3 presents the path model for core and noncore variables and the leaders' \( G_L \) and \( R_m \) values. (Because subordinates' ANOVA results for \( G_L \) and \( R_m \) were almost identical to the leaders', they were not added to provide the most parsimonious model [fewest variables] predicting leader achievement given the small sample size.) The model was created by performing a series of hierarchical regressions predicting the leaders' achievement (\( r_a \)) by moving from the figure's right to left-hand side, consistent with the causal pathways proposed by multilevel theory (Hollenbeck et al., 1995). In particular, we first regressed \( r_a \) on all eight variables shown in Figure 4.3, but only leader knowledge (\( G_L \)) and cognitive control (\( R_m \)) significantly predicted \( r_a \). We then regressed \( G_L \) on the remaining seven variables, but only leader informity, hierarchical sensitivity, and staff validity significantly predicted \( G_L \). Then we regressed the remaining six...
variables on $R_{ac}$ and so forth. Only significant relationships ($p < 0.05$) and beta weights are shown in Figure 4.3. The total $R^2$ was 0.988 (adjusted $R^2$ was 0.981). This occurred because $r_x$ is fully decomposable into $G$ and $R_w$ when $R_w$ is held constant and approaches 1.0, as in our study.

We make five points. First, the effect of all non-core and core constructs on leader achievement ($r_a$) is mediated through its effect on leaders' knowledge ($G_a$) and cognitive control ($R_{ac}$). One needs these LME parameters to understand why time pressure (tempo) affected leader achievement, and why cognitive feedback did not.

Second, leaders' informity affected leaders' knowledge ($G_a$), their ability to use the judgment model they were trained to use Equation 4.2. Achievement ($r_a$) for targets without staff recommendations appears to have decreased with increasing tempo because of a decrease in the amount of information sent to the leaders, which caused them to adopt less accurate judgment strategies for combining information, as shown by the decrease in $G_a$.

In contrast, cognitive feedback had no effect on leader informity; consequently, it had no effect on their knowledge ($G_a$) or achievement ($r_a$) without staff recommendations. Cognitive feedback did affect staff informity and, in turn, leader achievement through the relationships shown in Figure 4.3, but these mediated effects were not sufficient to result in a significant interface main effect or interface by tempo interaction on leader achievement.

Fourth, as tempo increased, staff informity decreased, which decreased staff validity. Decreased staff validity caused a decrease in the leaders' knowledge of what to do ($G_a$), their sensitivity to their subordinates' accuracy (hierarchical sensitivity), and their control of the judgment strategies they tried to use ($R_{ac}$).

Fifth, decreased hierarchical sensitivity only affected $G_a$, not $R_{ac}$. Leaders' decreased sensitivity to the validity of their subordinates' recommendations significantly caused them to modify their judgment strategies. It had no effect on their ability to apply whatever strategies they chose.

**Discussion**

Brunswikian multilevel theory, as operationally defined by the LME and path modeling, showed that the decrease in leader achievement with increasing time pressure (tempo) was caused by a
breakdown in the flow of information among team members. Once informity fell, all the other constructs fell like dominos. As tempo increased, the amount of information passed among the team decreased, as did staff validity. With less information and less accurate recommendations, leaders were less able to use the decision model they were trained to use ($G_2$), less sensitive to which staff member was more accurate (hierarchical sensitivity), and less consistent ($R_e$) when making their decisions.

The provided cognitive feedback could not compensate for this domino effect, because it was not designed to support information flow. Moreover, it was not as accurate as it should have been, for the regression equation showing team members the weights they placed on the cues assumed that they had all six cues when they made their decisions and therefore did not account for the breakdown in information flow.

There are ways to maintain information flow. For example, Hedlund, Ilgen, and Hollenbeck (1998) showed that face-to-face teams have higher informity levels than computer-mediated teams like those in our study. Hollenbeck et al. (1998) and Adelman et al. (2004) have demonstrated that information flow can be maintained by providing informity feedback or implementing a simple modification to the human–computer interface, respectively. However, providing rapid and accurate cognitive feedback may be more difficult because of the complexity of the steps to calculate judgment models accurately when all the cue information and subordinates' recommendations are not available. Of course, that is an issue for future research. Nevertheless, the presented analysis indicates that cognitive feedback will not matter much if the flow of information is not maintained.

It is important to emphasize that knowledge ($G$) and control ($R_e$) are conceptually independent (Hammond & Summers, 1972). Noncore and core constructs can affect one without affecting the other. For example, tempo's negative effect on leader informity only affected their knowledge ($G$), not their control ($R_e$) in applying it. In contrast, tempo's negative effect on staff informity negatively affected staff validity, which negatively affected leaders' knowledge and control, as well as their hierarchical sensitivity.

The path analysis was post hoc, because the experiment was not designed to manipulate specific pathways. However, the results fulfilled all three conditions for establishing mediation between tempo and achievement (Hedlund, Ilgen, & Hollenbeck, 1998). First, tempo (a noncore, independent variable) had an effect ($p < 0.05$) on the mediating core variables (leader informity, staff informity, staff validity, and hierarchical sensitivity, $p = 0.09$), and leader knowledge and cognitive control. Second, tempo significantly affected leader achievement ($r_\text{L}$), the dependent variable. Third, all mediating variables had a significant relationship with $r_\text{L}$, although informity dropped out of the regression when staff validity and hierarchical sensitivity entered it. Given the small sample size (8 teams by 3 temps for 24 data points), the effect size was large for all variables in the path model. Nevertheless, the post hoc nature of the analysis and the small sample size are limitations. Although the obtained pathways have substantial face validity in explaining the results in Miller et al. (2000) and meet the conditions for mediation, research needs to test them versus other pathways (Hollenbeck et al., 1998) to assess their generality.

Gigone and Hastie (1997) argued that mean square error (MSE) is the best measure of group accuracy because it not only incorporates $r_e$, but measures of mean bias and standard deviation bias in one's decisions. In an analysis not reported here, ANOVAs generated the same results for MSE as for $r_e$ because there were no effects for mean or standard deviation bias for either the leaders' or subordinates' decisions. Consequently, we used $r_e$ in the path modeling, not MSE, to have fewer variables given the small sample size.

In closing, the research showed how Brunswikian theory and the LME can be used to study the effect of computer displays on team (or individual) decision making. Moreover, we have built on this initial study. Adelman, Miller, and Yeo (2004) used the same task as the one described herein, but they controlled the flow and therefore amount and type of information received by individuals. They found that an icon telling individuals when they had received information from their (simulated) teammates about a target led to higher achievement under low tempo. However, achievement with the icon significantly decreased as tempo increased, in part, because of a significant decrease in cognitive control, not knowledge. In contrast, they found that individuals with higher working memory capacity had higher achievement, whether or not they had the icon, because they had higher knowledge (abil-
Understanding the Effects of Computer Displays

It is to use Equation 4.2) with more information, not because of any effect on their cognitive control (or consistency) in doing so. That research, like the one described here, shows that the sensitiviity provided by the LME and scope of the Brunswikian theory in which it is embedded has substantial power to improve our ability to understand the effect of technology on the user's mind.

Acknowledgments. We thank Michael Schoelles for modifying the Team Argus interface for the study and for implementing the data collection scheme required for data analysis. The research was supported under Grant no. F49620-97-1-0353 from the Air Force Office of Scientific Research to George Mason University. The views, opinions, and findings herein are those of the authors and not a Department of the Air Force position.

References


