

Behavior Analysis of Team Performance: A Case Study of Membership Replacement

Henry H. Emurian, Kip Canfield, & Joseph V. Brady

Abstract

A three-person team performance task (TPT) is described, and evaluative results are presented under conditions of individual fixed ratios required to complete a work component and a team fixed ratio required to complete a work component. After an initial team performed the task over four successive days, a member was replaced with a novice, and the newly formed team performed the task over four successive days thereafter. The results showed differences in performance metrics between the individual and team ratio conditions and between the original and the reformed teams. When communications among team members were permitted at the start of the last two sessions of the study, individual contributions by the three members to the team ratio requirement were equivalent during the final session. The results show the sensitivity of the task to individual and team performance requirements and to membership replacement. They also show the impact of tactical decision making on work distributions. The range of outcomes suggests the utility of this type of task to assess the status of a team and to act as a potential countermeasure to team fragmentation.

Keywords: Team performance, behavioral health

The need to develop tools to assess and support the behavioral health of space-dwelling crews continues to be acknowledged by NASA (Suedfeld, Bootzin, Harvey, Leon, Musson, Olthmanns, & Paulus, 2010). In that regard, the term “behavioral health” encompasses a broad range of affective, social, and skilled individual and crew performances that must be sustained under the obviously stressful circumstances of long-duration spaceflight (Brady, 2007; Emurian & Brady, 2007). The detection of impending performance degradation necessitates the consideration of innovative approaches to monitor and measure both individual and team performances that realistically relate to the operational status of a crew. The introduction of effective countermeasures to such degradation is complementary to detection, and potential solutions to these two challenges will benefit from a technology that can integrate both considerations within a common conceptual framework with respect to task performance.

A three-person team performance task (TPT) was proposed as a tool to diagnose the status of a crew (Emurian, Canfield, Roma, Gasior, Brinson, Hienz, Hursh, & Brady, 2009), and the rationale of its design, from the perspective of behavior analysis, and an evaluation of its effectiveness have been reported (Emurian, Canfield, Roma, Brinson, Gasior, Hienz, Hursh, & Brady, in press). The initial evaluations were based upon having subjects perform the task for fixed time periods (e.g., 12 min), with instructions to maximize performance effectiveness. Although providing important feedback regarding the properties of the task and performance metrics associated with individual and team performances, a more realistic diagnostic scenario would require a crew to complete a given task without regard to temporal constraints. Accordingly, the present extension of the task implements a fixed-ratio requirement on performance accuracy at the level of the individual team member and at the level of the team. The present report is a case study of the evaluation of such an extension under conditions of the replacement of an established team member with a novice. The context of this study includes analyses of group membership replacement previously undertaken within a continuously programmed environment (Emurian, Brady, Ray, Meyerhoff, & Mougey, 1984).

Method

Subjects

Four UMBC undergraduate students volunteered to participate in response to an announcement posted on the student listserv. Volunteers were directed to read the information posted on the web (<http://nasa1.ifsm.umbc.edu/tpt/>). The study was approved by UMBC's Institutional Review Board, and informed consent was obtained at the time of each daily session. Each participant was paid \$30 in cash at the completion of a session.

Table 1 presents demographic details about the four subjects collected before Session 1 for

Table 1

S#	Status	Major	Sex	Age	Game Experience	Computer Experience
1	Junior	Health Administration	F	19	8	8
2	Junior	Social Work	F	19	2	6
3	Junior	Health Administration	M	20	7	7
2*	Sophomore	Biology	F	18	8	9

subjects 1, 2, and 3 and before Session 5 for subject 2*. Two rating scales were administered to assess each subject's experience with computer games and overall computer experience. Each scale was a 10-point scale, where the anchors were *1 = No Experience (I am a novice.)* to *10 = Extensive Experience (I am an expert.)*. Notable, perhaps, are the comparatively low ratings by S2 for both game and computer experience. Subjects 1, 2, and 3 reported being acquainted prior to the study. Subject 2 was replaced at Session 5 of the study by S2*, and the replacement reported having no prior acquaintanceship with the other two subjects. The subjects were instructed not to discuss the task between sessions, and post-session debriefings always confirmed that practice.

Team Performance Task (TPT)

The TPT was designed for use by three-person groups, and the prototype has been described in detail elsewhere (Emurian et al., in press). **Figure 1** presents a screen shot of the display presented to a subject (in this case, User1, the designation for S1). The display was similar for all three subjects, who operated the task from three separate computer terminals. In the current configuration, communications were not permitted among the three subjects during a session, beyond task-related actions during a session, to be described. The server for the TPT, which is deployed on the Internet, was running on a port behind the UMBC firewall.

In brief, the task requires the subject to capture and drag a Resource block at the top and deposit it on the target without striking a barrier. For an accurate deposit, the color of the Resource block deposited must match the color of the target, which changes from trial to trial. Between the resource blocks and the target are nine rows. Each row contains a barrier, and the barrier changes position in a row between 10 –

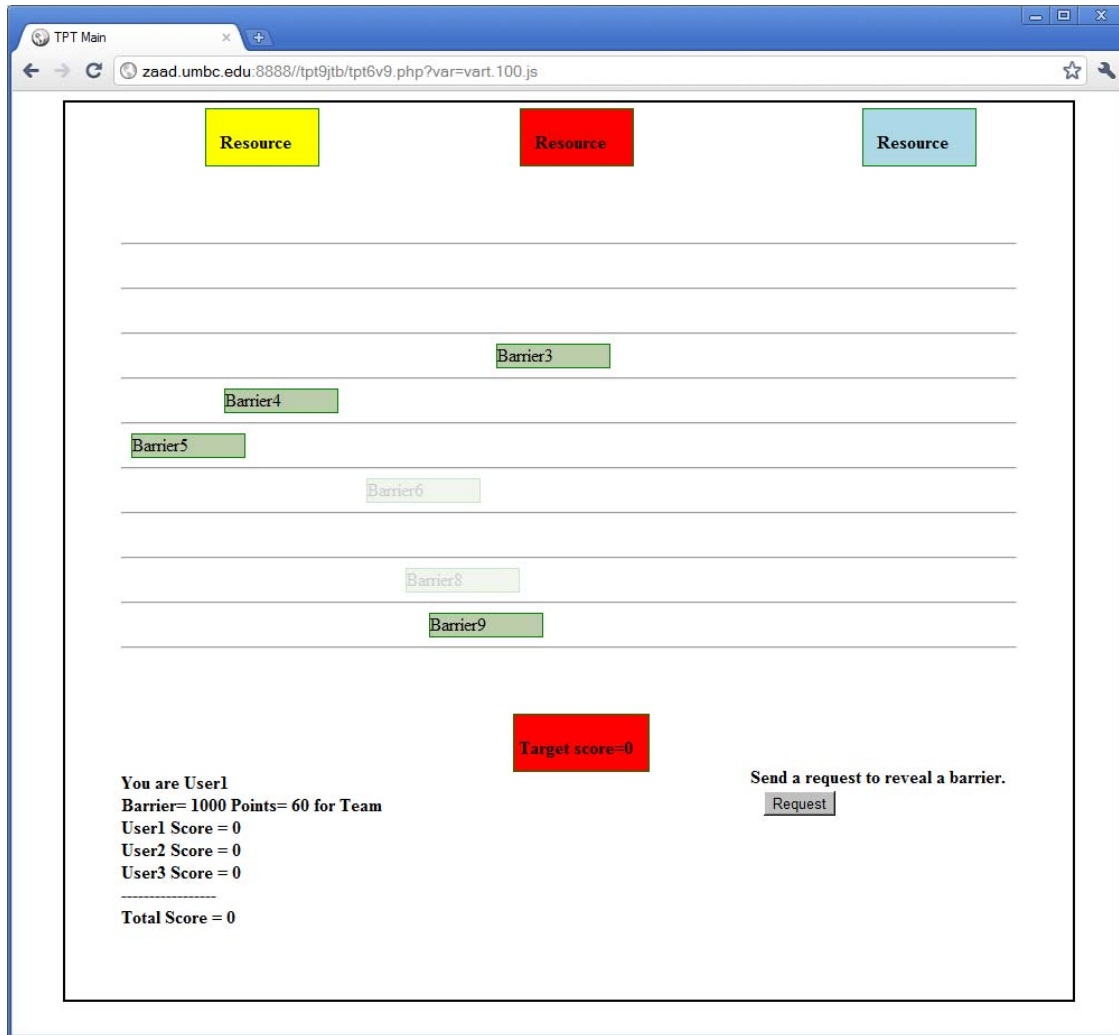


Figure 1. A screen shot of the TPT display for User1. This version of the task requires the team to accumulate 60 points to complete the component, irrespective of the relative contributions of the three team members to that ratio requirement. The hold time to reveal a barrier is 1 sec (1000 msec).

20 sec. Each subject controls the visibility of a barrier within three of the nine rows, and the rows assigned to each subject are determined randomly at the start of each task component, to be described, within a session. In **Figure 1**, four barriers are highlighted and visible to all subjects, and two barriers (Barrier6 and Barrier8) are dim. Barriers 6 and 8 are visible in that dim state only to User1. To make a barrier visible in the highlighted state to the other two subjects, the cursor must be positioned over the barrier for .25, 1, or 4 sec, depending on the component of the session, with the left mouse key held down. In **Figure 1**, Barrier = 1000 indicates that the cursor must be positioned over a barrier, with the mouse down, for 1 sec (i.e., 1000 msec) to highlight the barrier and make the barrier visible to all subjects.

For a correct deposit on the target, which involves dragging and depositing an identically colored Resource block without striking a barrier, one point is added to the Target score. The corresponding “scoreboard” of points is also updated. If a barrier is struck, whether highlighted, dim, or invisible, one point is subtracted from the score, which can become negative. Whenever such a barrier “hit” occurs, the associated Resource block being dragged is eliminated from further play, and a new Resource block at the top has to be dragged to the Target block. Optimal performance, then, requires cooperation among the

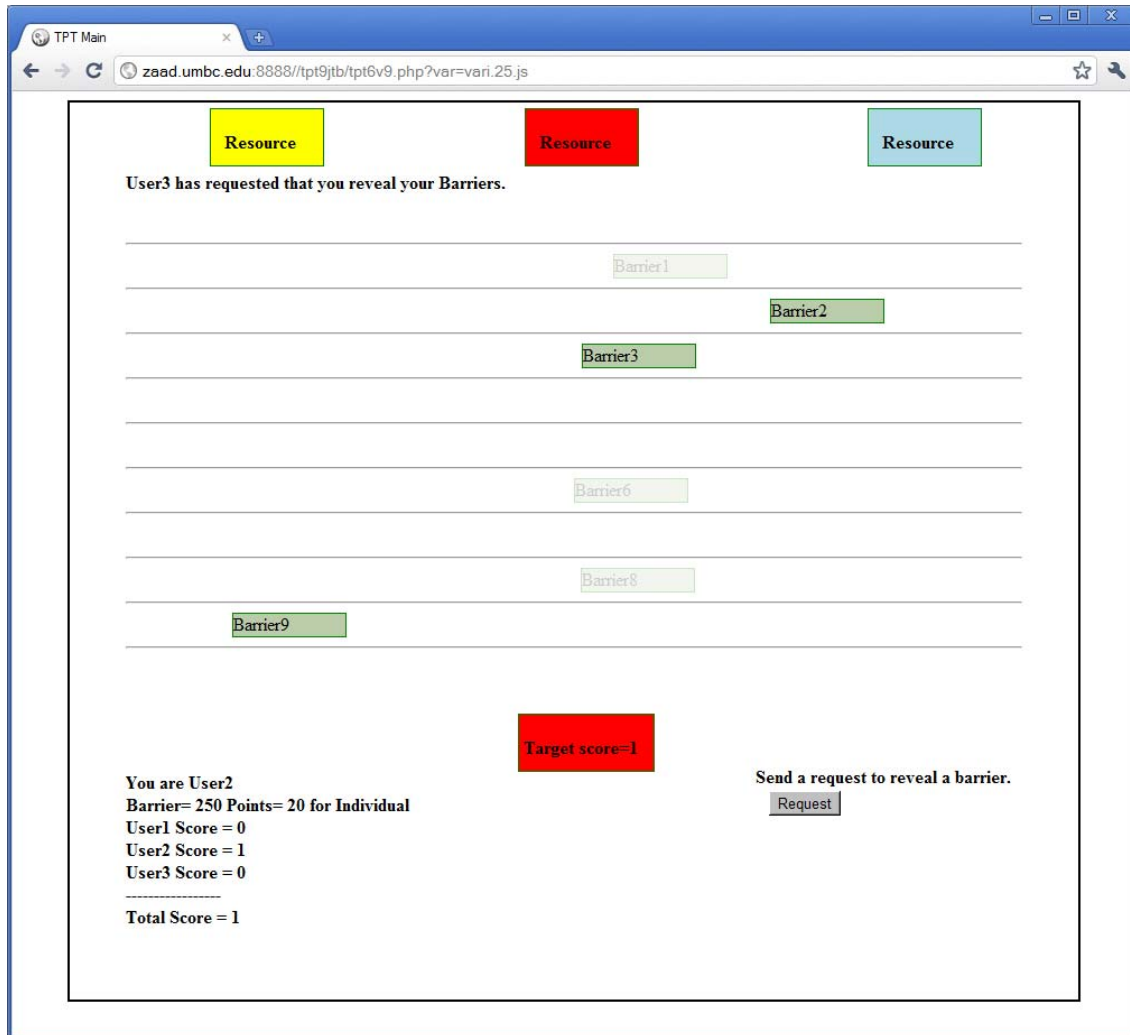


Figure 2. A screen shot of the TPT display for User2. In the upper left corner is a message indicating that User3 has requested that the other team members reveal their barriers. This example shows that User2 has accumulated 1 point. This version of the task requires each team member to accumulate 20 points to complete the component for all three members. The hold time to reveal a barrier is .25 sec (250 msec).

three team members to highlight their respective dim barriers so that Resource block movements by other team members can avoid hitting barriers, thereby permitting Target counts to be maximized.

At the bottom right of the display is a button labeled “Request.” When a subject clicks that button with the mouse, a text message is presented at the top of the displays of the other two teammates. For example, if User3 clicks that button, the following message appears to the other two subjects: “User3 has requested that you reveal your barriers.” Successive messages appear in a scrollable list, and all messages on a subject’s display are removed whenever that subject initiates movement of a Resource block.

Figure 2 presents a screen shot of the display for User2 after User3 has clicked the “Request” button. The message appears in the top left of the display. **Figure 2** also shows the scoreboard when there is a 20-point individual ratio in effect, and the hold time (“delay”) to reveal a barrier is .25 sec (250 msec).

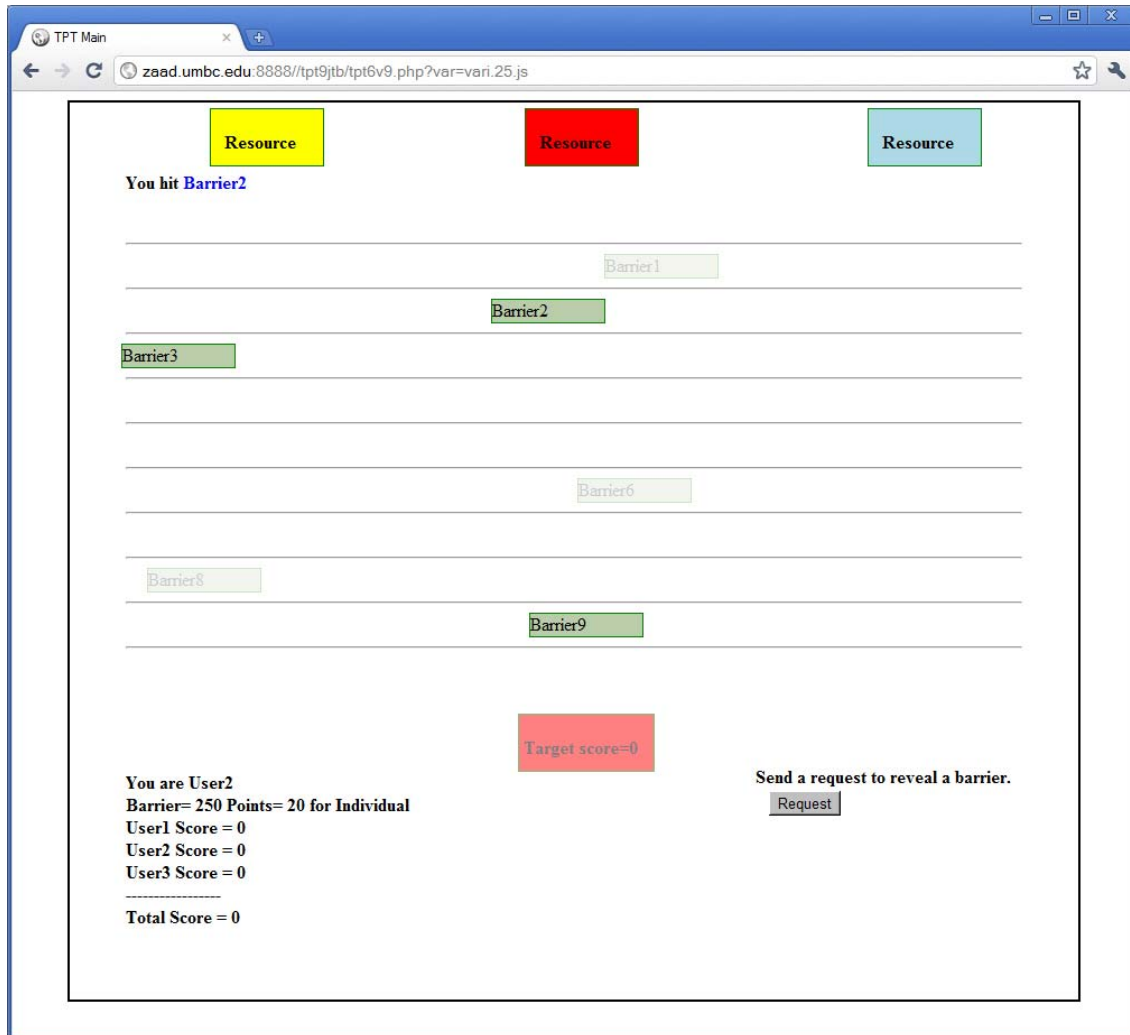


Figure 3. A screen shot of the display for User2 after a barrier was hit. A message appears in the upper left corner of the display. In this case, the barrier in row 2 (Barrier2) was hit. The target is dimmed until the next movement of a Resource block. In comparison to Figure 2, the display shows the results of a 1-point decrement from the Target score and from the scoreboard.

Figure 3 presents a screen shot of the display for User2 after a barrier was hit. A message appears in the upper left corner of the display. In this case, the barrier in row 2 (Barrier2) was hit. The Target block is dimmed until the next capture and movement of a Resource block. The display also shows the decrement of 1 point from the Target score and from the scoreboard for User2.

Procedure

The study took place within a small, rectangular, windowless laboratory containing four tables, each table holding a PC, with two PCs positioned back-to-back along two walls. The subjects were seated a few feet apart such that they did not face each other. The task was presented to each subject using a Dell Optiplex 745 PC having a 17-inch screen. The first author supervised the study and remained in the laboratory with the three subjects and the research assistant. The subjects were told not to speak to each other during performance on the task, and the only direct communications permitted were through requests to reveal barriers, a feature of the task described above.

Each daily session consisted of six components. The components differed in terms of the hold time required to reveal a barrier (“barrier reveal delay”). The following six barrier reveal delays were presented within successive components in the following order for all eight sessions of the study: (1) .25 sec, (2) 1 sec, (3) 4 sec, (4) .25 sec, (5) 1 sec, and (6) 4 sec. These durations were chosen to match the range of delays evaluated previously (Emurian et al., in press). For the individual condition (I), each subject was required to accumulate 20 points to complete a component. For the team condition (T), the subjects were required to accumulate 60 points to complete a component, irrespective of the relative contributions by the individual team members to that requirement. A counterbalanced order of the two conditions was in effect across the eight sessions, respectively: (1) I – T, (2) T – I, (3) I – T, (4) T – I, (5) I – T, (6) T – I, (7) I – T, and (8) T – I.

During each condition, there were three components, with each component having one of the three barrier reveal delays as presented. For example, during Session 1, the individual ratio condition was in effect first. There were three components in that condition. The first component required each subject to accumulate 20 points, and the barrier reveal delay was .25 sec. The second component also required each subject to accumulate 20 points, and the barrier reveal delay was 1 sec. The third component required each subject to accumulate 20 points, and the barrier reveal delay was 4 sec. Under the team condition, the sequence of barrier reveal delays was identical to the individual condition sequence, but the team was required to accumulate 60 points to complete each component, irrespective of the contributions of the individual team members to that requirement. For the individual ratio components, a task completion message appeared on the display after all subjects had accumulated 20 points. It was not possible for a subject to accumulate more than 20 points during the individual ratio components, but all other features of the task continued to function. During the team ratio components, the completion message appeared after 60 points had been accumulated by the team.

The study consisted of eight sessions, spaced a few days apart depending upon the schedules of the subjects. The first session was on 6/18/2010, and the eighth session was on 7/8/2010. Each session began sometime between 9 AM and 2 PM, and only one session was scheduled on a given day. Before the start of the first session, the task was explained to the subjects, and a practice session was administered with abbreviated parameters. Each subject was assigned a “user number” to be selected to start the task during each component across the conditions. As indicated above, the task automatically terminated when the ratio requirement in effect was completed.

At the conclusion of each component within a session, the subjects completed the 6-item Perceived Cohesion Scale (PCS) (Salisbury, Carte, & Chidambaram, 2006), which yielded ratings of group Belonging and Morale, and the NASA Task Load Index¹ (NASA-TLX), a measure of perceived workload (Cao, Chintamani, Pandya, & Ellis, 2009).

Following the fourth session with the original team, S2 was replaced by a new member. That subject (S2, Table 1) was replaced by convenience, because her schedule at the time did not allow further participation in the study. At the fifth session, the new member (S2*, Table 1) joined the team. She reported having no prior acquaintanceship with the other subjects. The subjects introduced themselves by name and major, and this was followed by the practice session to allow the new subject to be familiarized with the task. Thereafter, four sessions took place that exactly replicated the first four sessions in the study with respect to the components and the conditions. However, prior to the beginning of Session 7, the team was instructed to spend time together to discuss the task and to consider ways to optimize their performance. The meeting lasted about 10 min. A similar meeting occurred prior to the beginning of Session 8. Other than these discussions in the laboratory, the subjects did not otherwise discuss the task, and they reported having no contact with each other between sessions.

Footnote 1: <http://humansystems.arc.nasa.gov/groups/TLX/>

Results

Individual data records will be presented, together with summaries of outcomes augmented by statistical inferences of orderliness.

The analysis was undertaken with the Welch robust test for main effects, cellwise comparisons, and complex contrasts, with the presence of potential interaction effects being determined by graphical inspection of the outcomes. Where indicated, Dunnett's T3 method was used for post-hoc pairwise comparisons, and p for other multiple comparisons was corrected with $.05/a$, where a = number of comparisons. Non-significant effects are not reported, to include effects of subject, delay, condition, and part, and they are not presented in the figures. For the analysis, the original team is designated as Part 1, and the reformed team is designated as Part 2. Because of the likely influence of participants on each other in this design, statistical tests were undertaken using between-subjects techniques, which are conservative in rejecting a null hypothesis in comparison to within-subjects techniques. Maxwell and Delaney (2004) was the reference source for the analysis, which was undertaken with SPSS. The figures are labeled with sec or msec for the barrier reveal delay components, hereafter referred to as "delay components" or "components," depending upon available space on an axis.

Points

Figure 4 presents total points accumulated at the completion of each barrier reveal delay component during the team condition by each subject across the eight sessions. As indicated in the Procedure section, S2 was replaced at Session 5, and the reformed team held a brief meeting (10 min)

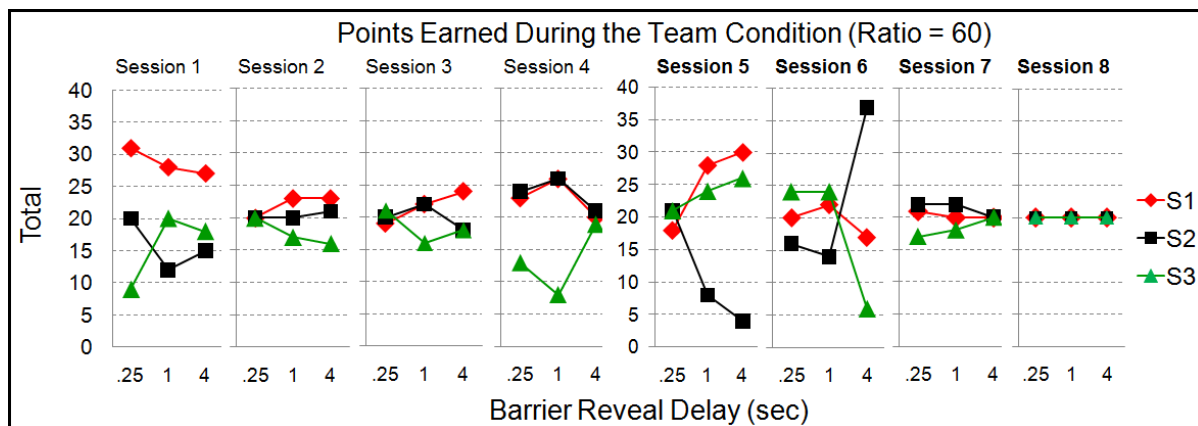


Figure 4. Points accumulated by each subject at the completion of the team ratio condition across the three delay components of the eight sessions. S2 was replaced at Session 5.

to discuss their performance tactics before Session 7 and Session 8.

The figure shows graphically that during Sessions 1 - 4, the three original team members did not contribute equally to the 60-point ratio in effect during the team condition. In that regard, the discrepancies among the subjects were greater during Sessions 1 and 4 in comparison to Sessions 2 and 3. It is notable, perhaps, that S3 showed the lowest point accumulation in seven of the 12 components in Sessions 1 - 4.

When S2 was replaced at Session 5, the impact on performance is graphically apparent. During the 1-sec and 4-sec delay components, S2's point accumulation differed markedly from the other two subjects, and S2 showed the lowest point accumulation of the study during the 4-sec delay component. During Session 6, although S2's point accumulation was lower than the other subjects during the .25-sec and 1-sec delay components, S2 showed the highest point accumulation of the study during the 4-sec component.

The impact of the tactical meeting is graphically evident in the point distributions during Session 7. Over the three delay components, the final component (i.e., the 4-sec barrier reveal delay) shows that each team member contributed 20 points to the 60-point team ratio requirement. That distribution persisted during Session 8, in which 20 points were contributed by all team members within each of the three components. Although performance on this particular metric stabilized over the last two sessions, ratings of team cohesion remained diminished, as presented below.

Durations of Session Components

Figure 5 presents durations for each team member to complete the individual ratio across the three delay components for the eight sessions. Also presented are the durations for the team to complete the team ratio. Time to complete the ratios generally decreased across the sessions for individual and team

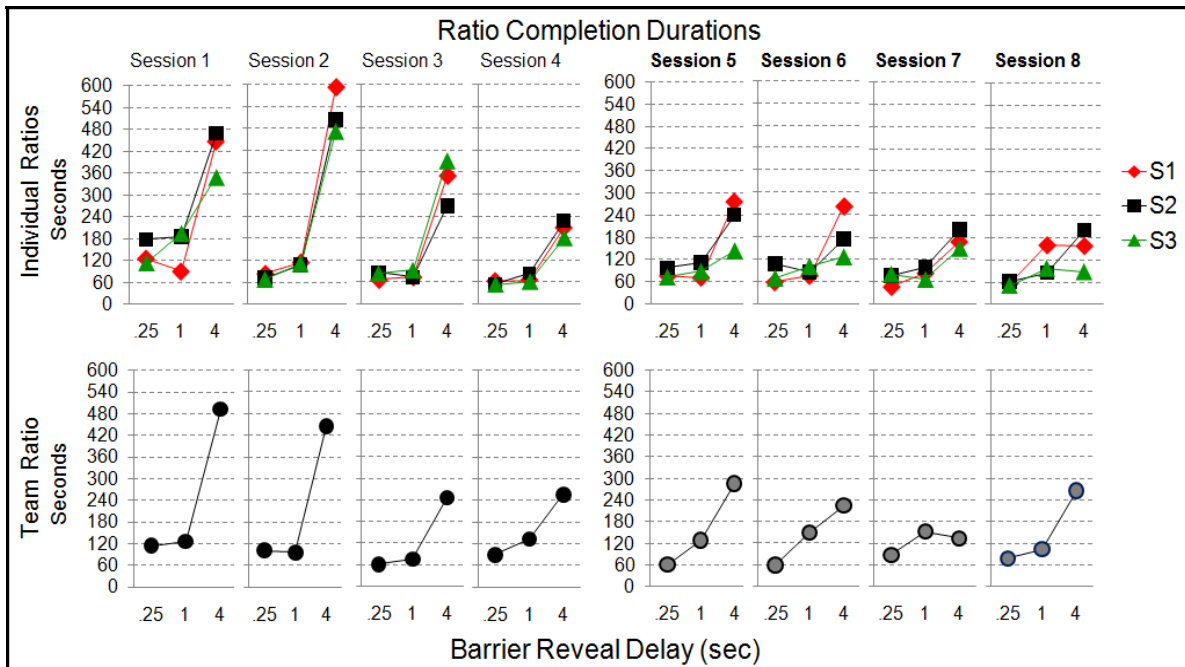


Figure 5. Durations to complete individual ratios across the three delay components for the eight sessions. Also presented are the durations for the team to complete the team ratio. S2 was replaced at Session 5. Black circle = original team, gray circle = reformed team.

ratios most notably in the 4-sec delay component. The shortest time to complete an individual ratio (i.e., 47 sec) was evidenced by S1 on Session 7 in the .25-sec delay component, and the longest time was also evidenced by S1 in the 4-sec delay component on Session 2 (592 sec). With respect to the team ratio, **Figure 5** shows graphically that the duration increased across the three delay components for six of the eight sessions. When S2 was replaced at Session 5, the gains acquired over the early sessions appeared to carry over when the team was reformed, at least in comparison to the first two sessions. The skill acquired by S1 and S3 may have compensated for the novitiate's inexperience with the task.

There was insufficient evidence to support differences among the subjects on the mean durations to complete individual ratios during the individual condition for Part 1 and Part 2. Accordingly, for the statistical analysis, the duration of a session component during the individual condition was taken to be the duration from the start to the completion of the component, without regard to the time when the first two individual ratios were completed.

Figure 6 presents mean durations and 95% confidence intervals across the three barrier reveal delay components for Part 1 and Part 2 of the study. The overall component durations are also presented. Notable in this figure is the comparatively higher durations observed during the 4-sec delay component, in comparison to the other delay components, and the higher durations observed there in Part 1 in comparison to Part 2.

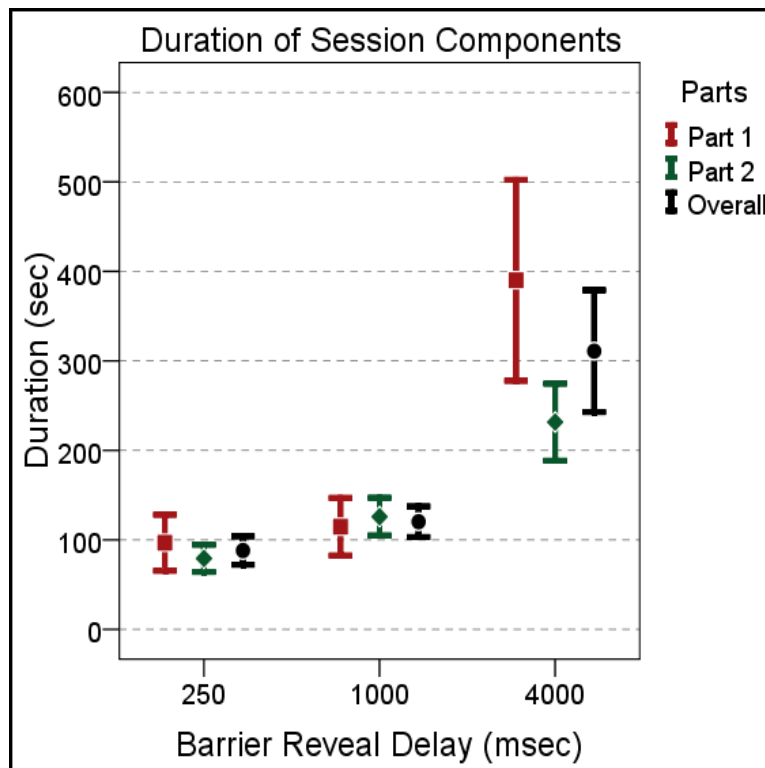


Figure 6. Means and 95% confidence intervals for durations of delay components for Part 1 and Part 2. The overall durations are also presented. Square = Part 1, diamond = Part 2, circle = Overall.

A Welch test was performed on the overall mean durations across the three delay components, and the outcome was significant ($W = 24.271$, $p = .000$). Using Dunnett's T3, all pairwise comparisons were significant ($p < .01$). With respect to Part 1 and Part 2, a delay by part interaction was suggested by a significant Welch test comparing the means for Part 1 and Part 2 at the 4000 msec delay ($W = 9.726$, $p < .05$, corrected for three comparisons). During that component, the mean for Part 1 is graphically higher than the mean for Part 2, suggesting that the skill acquired by S1 and S3 during the first four sessions carried over into the last four sessions of the study despite the participation of the novice, S2.

Barrier Reveals

Figure 7 presents total barrier reveals for each subject across the three delay components for the eight sessions for the individual and team ratio conditions. S2 was replaced at Session 5. The figure

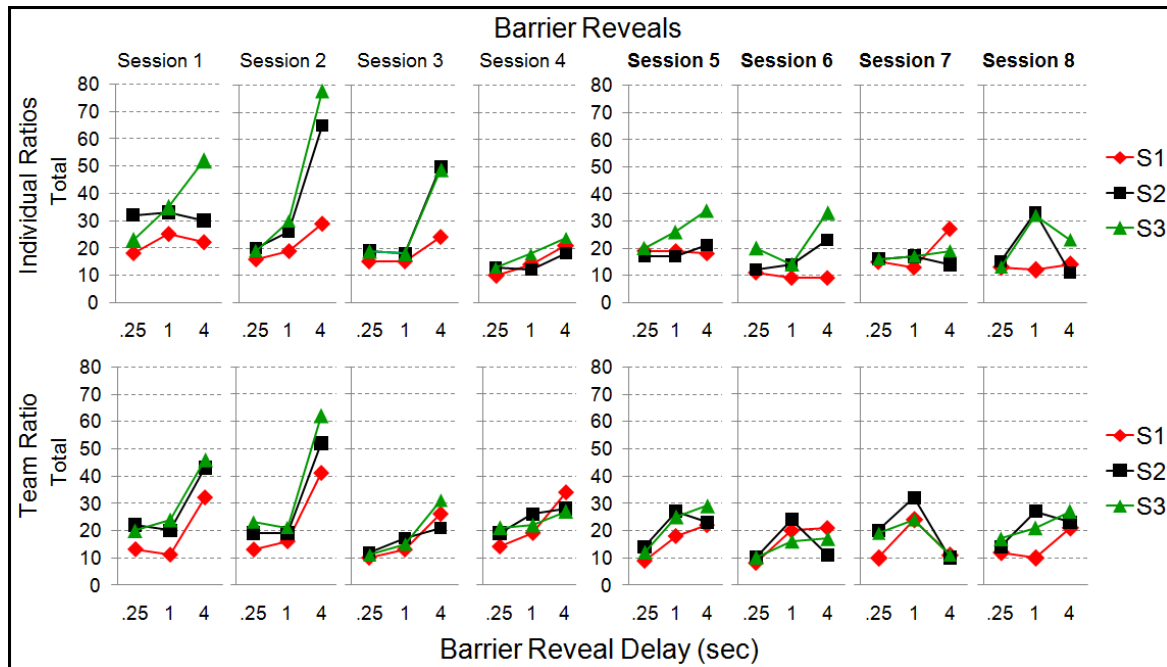


Figure 7. Total barrier reveals for each subject across the three delay components for the eight sessions for the Individual and Team ratio conditions. S2 was replaced at Session 5.

shows graphically that total barrier reveals for all subjects generally increased across the three delay components for many of the sessions. When S2 was replaced at Session 5, performance for all subjects seemed similar to the first four sessions. With respect to the three individual subjects, **Figure 7** suggests that S3 made the most barrier reveal actions and that S1 made the fewest such actions. The figure also suggests that total barrier reveals were higher in Part 1, in comparison to Part 2, especially with respect to the first two sessions.

Figure 8 presents means and 95% confidence intervals for the total number of barrier reveals exhibited by the subjects across the three delay components for the sessions within Part 1 and Part 2. The overall means are also presented. The figure shows graphically that the overall mean number of barrier reveals increased as the reveal delay increased, and the means for parts were higher during Part 1 for the 250 msec and 4000 msec delays.

A Welch test was performed on the overall mean barrier reveals across the three delay components, and the outcome was significant ($W = 20.089$, $p = .000$). Using Dunnett's T3, all pairwise comparisons were significant ($p < .01$). A Welch test comparing the mean reveals within all delay components across conditions between Part 1 (mean = 25.1) and Part 2 (mean = 18.1) was significant ($W = 14.590$, $p = .000$). With respect to Part 1 and Part 2, a delay by part interaction was suggested by a significant Welch test comparing the means at the 4000 msec delay ($W = 25.270$, $p < .001$, corrected for three comparisons).

Figure 9 presents means and 95% confidence intervals for barrier reveals by each subject across all delay components in the individual and team conditions for the four sessions in Part 1 and Part 2. A Welch test performed on the means across the subjects was significant for Part 1 ($W = 4.419$, $p = .018$). Using Dunnett's T3, pairwise comparisons showed a significant difference between S1 and S3. A Welch test performed on the means across the subjects was also significant for Part 2 ($W = 4.745$, $p = .013$). Using Dunnett's T3, pairwise comparisons showed a significant difference between S1 and S3. Welch tests comparing the means for Part 1 and Part 2 were significant for all three subjects ($p < .05$).

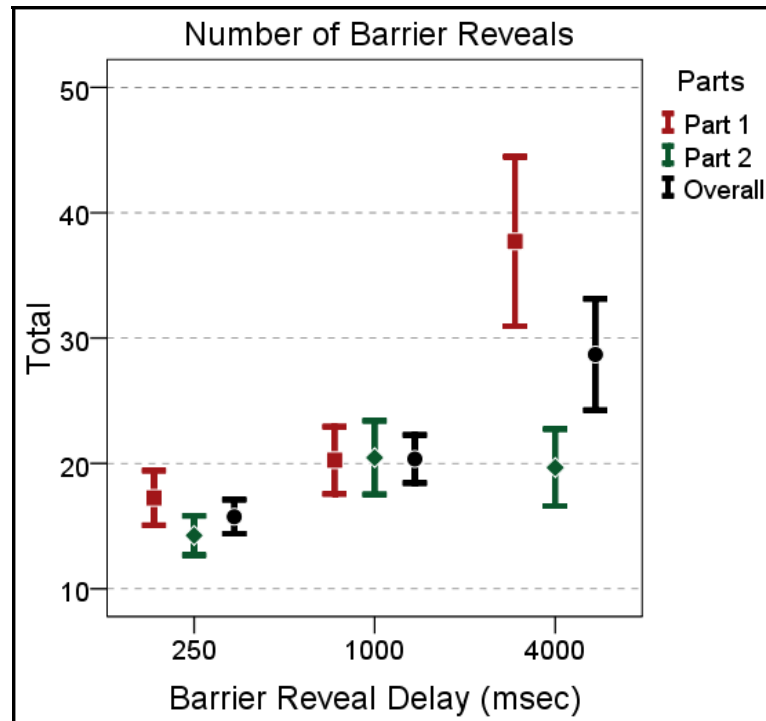


Figure 8. Means and 95% confidence intervals for the total number of barrier reveals exhibited by the subjects across the three barrier reveal delay components for the sessions within Part 1 and Part 2. The overall means are also presented. Square = Part 1, diamond = Part 2, circle = Overall.

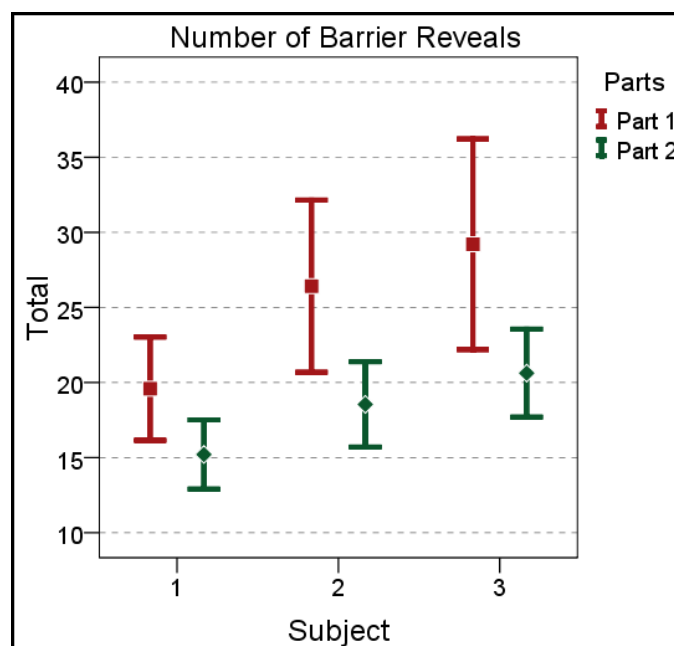


Figure 9. Means and 95% confidence intervals for total barrier reveals across all delay components by each subject during Part 1 and Part 2. Square = Part 1, diamond = Part 2.

Barrier Reveal Range

For analysis, a maximum range value was determined for the outcomes observed during each delay component for all sessions within the individual and team ratio conditions. This maximum range value was computed by subtracting the smallest from the largest number of barrier reveals within all delay components, yielding 48 such values, 24 within the individual condition and 24 within the team condition. For example, in **Figure 7**, the maximum range in Session 2 for the individual condition is graphically apparent in the data for S3 and S1.

Figure 10 presents means and 95% confidence intervals for the maximum range values in the individual and team conditions across the 250, 1000, and 4000 msec delays. The overall means are also presented. Most notable in this figure is the comparatively high mean value for the individual condition at the 4000 msec delay.

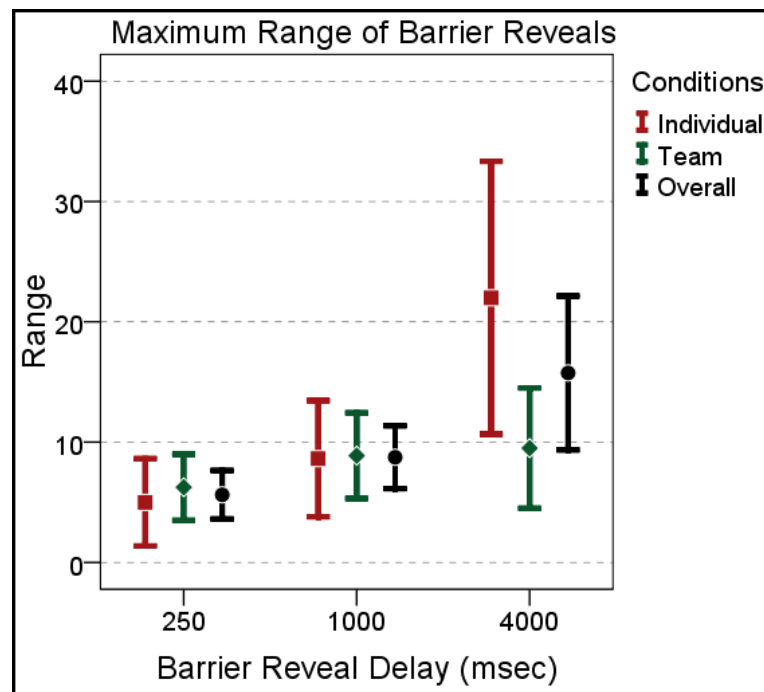


Figure 10. Means and 95% confidence intervals for the maximum range of barrier reveals for the three delay components across the Individual and Team conditions. Square = Individual condition, diamond = Team condition, circle = Overall.

A Welch test was performed on the overall mean range across the three delay components, and the outcome was significant ($W = 10.256$, $p = .000$). Using Dunnett's T3, significant differences were obtained between the means for the 250 and 4000 msec components and for the 1000 and 4000 msec components ($p < .05$). To compare the range between the individual and team conditions, a difference score was computed based upon the difference between the range in the individual condition minus the corresponding range in the team condition for each of the eight sessions. Under the assumption that the differences are all zero when no effect of condition is present (Maxwell & Delaney, 2004, p. 626), the outcome of the comparison of the observed differences with a population of zeros was significant ($F(1,7) = 5.70$, $p < .10^2$) for the 4000 msec component, corrected for three comparisons.

Footnote 2: Because this is an initial investigation, we chose to lower the threshold to reject the null hypothesis for this instance of a multiple comparison, rather than risk accepting a false hypothesis. Systematic replications will determine the robustness of this outcome.

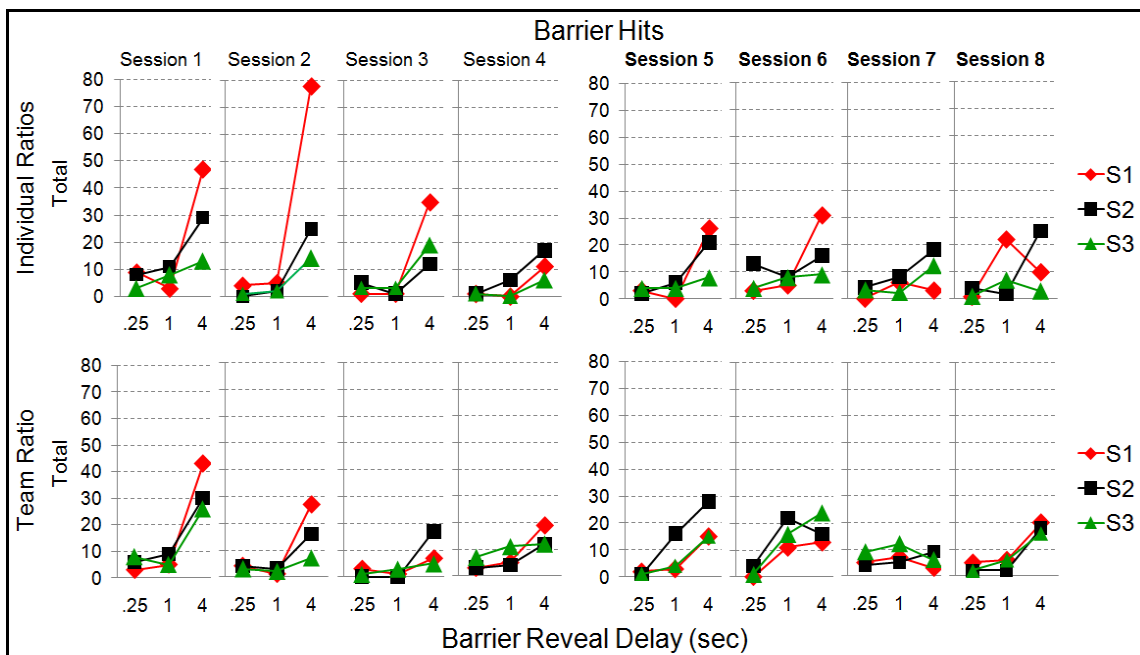


Figure 11. Total barrier hits for each subject across the three delay components for the eight sessions for the Individual and Team ratio conditions. S2 was replaced at Session 5.

Barrier Hits

Figure 11 presents total barrier hits for each subject across the three delay components for the eight sessions for the individual and team ratio conditions. The figure suggests graphically that barrier hits declined over the first four sessions and that within sessions, hits were highest within the 4000 msec delay

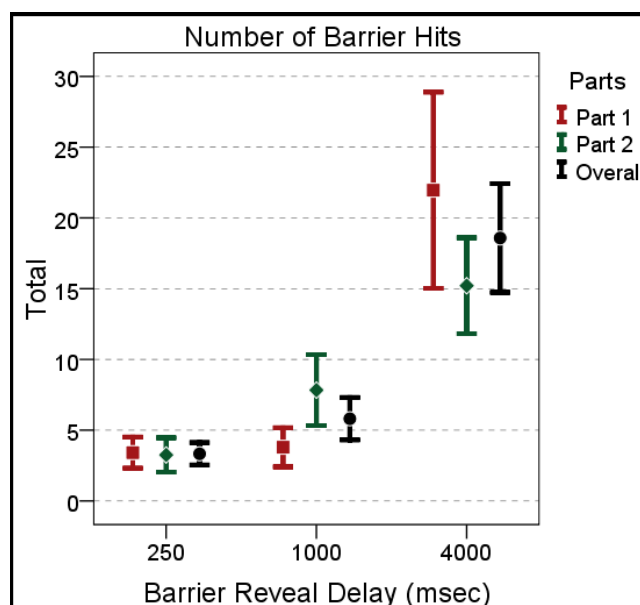


Figure 12. Means and 95% confidence intervals for the total number of barrier hits observed across the three delay components for Part 1 and Part 2. The overall means are also presented. Square = Part 1, diamond = Part 2, circle = Overall.

component. When S2 was replaced at Session 5, barrier hits thereafter were somewhat irregular during the individual condition, but were notably low and similar during Sessions 7 and 8 for the team condition. **Figure 11** also suggests that S1 evidenced the highest barrier hits during many of the sessions, especially in comparison to S3.

Figure 12 presents means and 95% confidence intervals for the total number of barrier hits observed across the three delay components for Part 1 and Part 2. The overall means are also presented. A Welch test performed on the overall means across the three delay components was significant ($W = 32.790$, $p = .000$). Using Dunnett's T3, all pairwise comparisons were significant ($p < .01$). A Welch test comparing the means between Part 1 and Part 2 was significant only at the 1000 msec delay ($W = 8.545$, $p < .01$, corrected for three comparisons).

Figure 13 presents means and 95% confidence intervals for barrier hits for each subject across the three delay components during Part 1 and Part 2. For Part 1, a Welch test performed on the means for the 4000 msec delay was significant ($W = 3.821$, $p = .050$). For Part 1, a complex contrast comparing the mean for S3 with the combined means for S1 and S2 was significant ($W = 7.211$, $p < .05$), suggesting a subject by part interaction in barrier hits at the 4000 msec component.

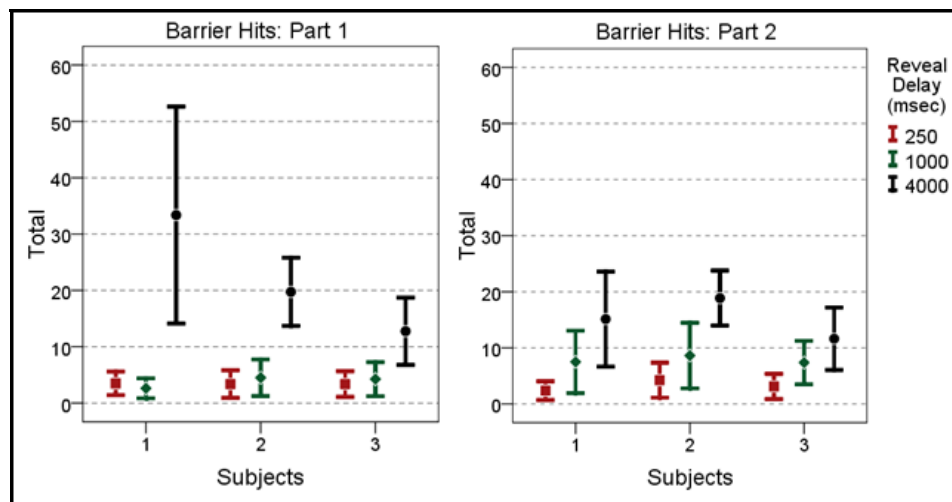


Figure 13. Means and 95% confidence intervals for barrier hits for each subject across the three delay components during Part 1 and Part 2. Square = 250 msec, diamond = 1000 msec, circle = 4000 msec.

Barrier Hits Range

Figure 14 presents means and 95% confidence intervals for the maximum range of barrier hits within each delay component across the individual and team conditions. The overall means are also presented. A Welch test was performed on the overall mean ranges across the three delay components, and the outcome was significant ($W = 10.256$, $p = .001$). Using Dunnett's T3, differences were significant between the 250 and 1000 msec delays and between the 250 and 4000 msec delays ($p < .01$). To compare the range between the individual and team conditions, a difference score was computed based upon the difference between the range in the individual condition minus the corresponding range in the team condition for each of the eight sessions. The outcome of the comparison of the observed differences with a population of zeros was significant ($F(1,7) = 12.39$, $p < .01$) for the 4000 msec component, corrected for three comparisons.

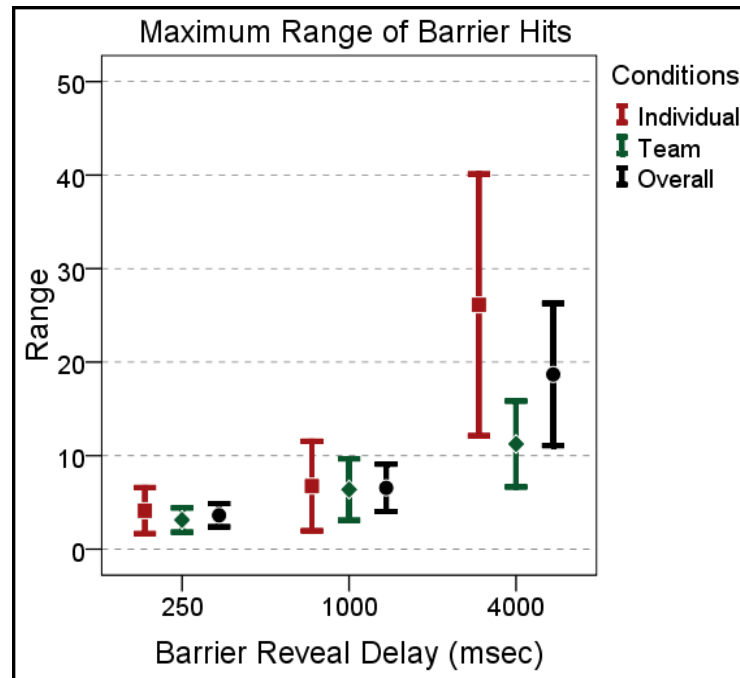


Figure 14. Means and 95% confidence intervals for the maximum range of barrier hits within each delay component across the Individual and Team conditions. The overall means are also presented. Square = Individual condition, diamond = Team condition, circle = Overall.

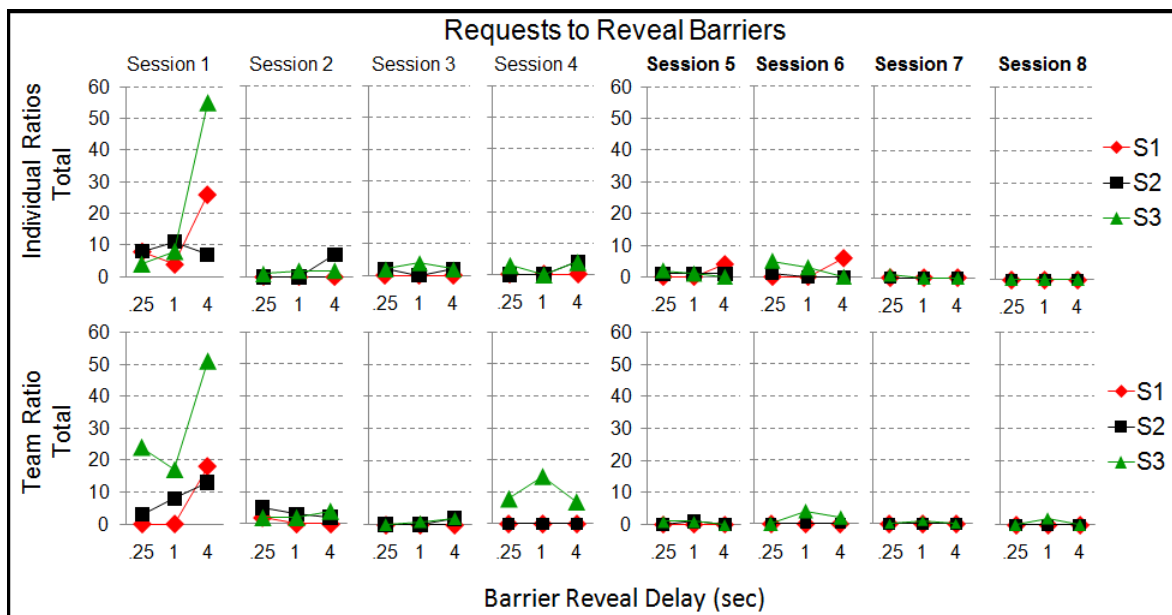


Figure 15. Total requests to reveal barriers for each subject across the three delay components for the eight sessions for the Individual and Team ratio conditions. S2 was replaced at Session 5.

Barrier Reveal Requests

Figure 15 presents total barrier reveal requests for each subject across the three delay components for the eight sessions for the individual and team ratio conditions. With the exception of the first session, requests were generally low, or even zero, thereafter. One other notable exception was S3 during Session 4 in the team condition.

NASA-TLX

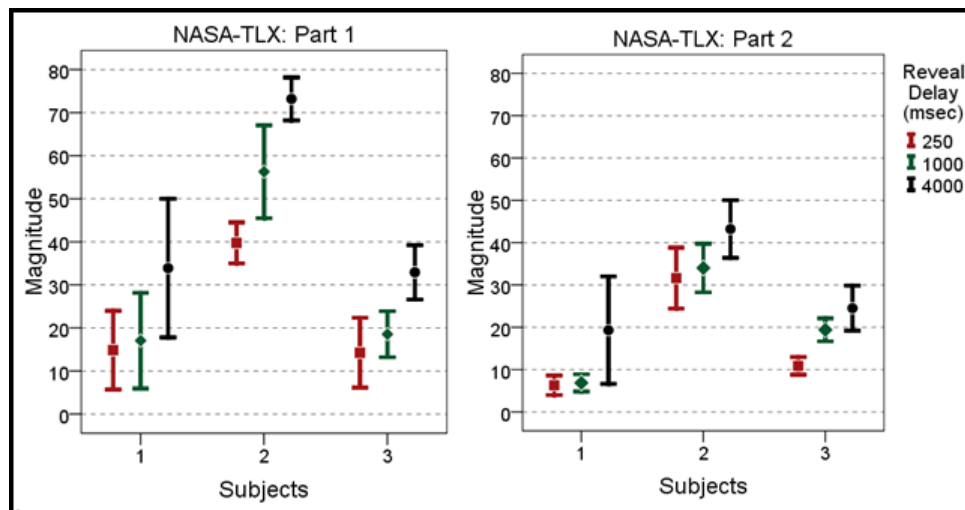


Figure 16. Means and 95% confidence intervals for the NASA-TLX for each subject during Part 1 and Part 2 across the three delay components. Square = 250 msec, diamond = 1000 msec, circle = 4000 msec.

Figure 16 presents means and 95% confidence intervals for the NASA-TLX for each subject during Part 1 and Part 2 across the three delay components. The figure shows graphically that the NASA-TLX ratings increased across the three delay components for all subjects in Part 1 and Part 2.

For Part 1, the means across the three delay components were as follows: 250 msec (mean = 22.9), 1000 msec (mean = 30.6), and 4000 msec (mean = 46.7). A Welch test comparing these means was significant ($W = 9.104$, $p = .000$). Using Dunnett's T3, pairwise comparisons were significant for differences between 250 and 4000 msec delays and between 1000 and 4000 msec delays ($p < .01$).

For Part 2, the means across the three delay components were as follows: 250 msec (mean = 16.3), 1000 msec (mean = 20.1), and 4000 msec (mean = 29.0). A Welch test comparing these means was significant ($W = 5.371$, $p = .008$). Using Dunnett's T3, pairwise comparisons were significant for differences between 250 and 4000 msec delays and between 1000 and 4000 msec delays ($p < .01$).

For S1, the means for the two parts were as follows: Part 1 (mean = 21.9) and Part 2 (mean = 10.8). A Welch test was significant ($W = 7.559$, $p = .009$). For S3, the means for the two parts were as follows: Part 1 (mean = 21.9) and Part 2 (mean = 18.3). A Welch test was not significant ($W = 1.803$, $p = .187$). A comparison for S2 was not undertaken because that participant differed between Part 1 and Part 2, and the data were based upon self-reports.

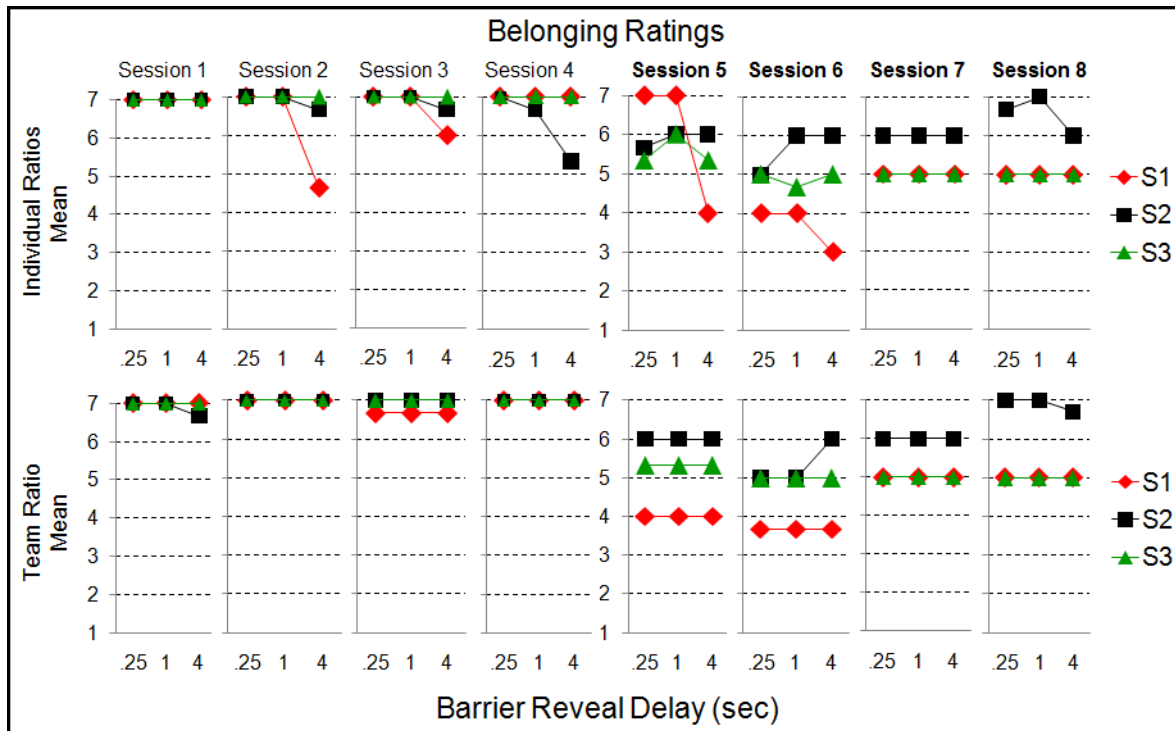


Figure 17a. Mean ratings for all subjects on the three PCS Belonging items across the three delay components for all sessions within the Individual and Team ratio conditions. S2 was replaced at Session 5.

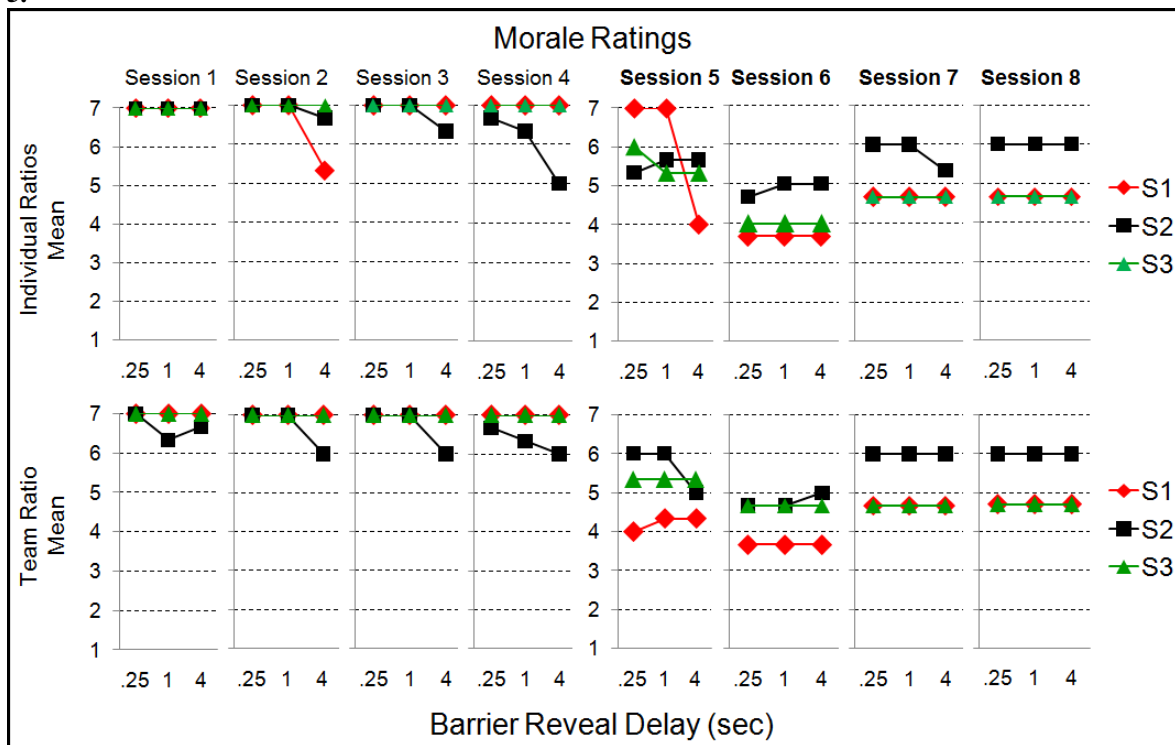


Figure 17b. Mean ratings for all subjects on the three PCS Morale items across the three delay components for all sessions within the Individual and Team ratio conditions. S2 was replaced at Session 5.

Group Cohesion Ratings

Figure 17(a,b) presents mean ratings on the PCS Belonging scale (a) and Morale scale (b), respectively, for all subjects across the delay components for all sessions within the individual and team ratio conditions. Inspection of the self-reports for S1 and S3 indicates that both of those subjects evidenced lower Belonging and Morale ratings during Part 2 of the study, in comparison to ratings during Part 1. For those subjects, ratings were stable on both scales over the three delay components during Sessions 7 and 8, but they were consistently low despite the fact that the team had agreed to equalize the work distributions during those sessions. Only S2, the novice, showed a modest increase in Belonging and Morale ratings during Session 8, in comparison to earlier sessions in Part 2 of the study.

Exemplar of a Behavioral Process

The detection and interpretation of orderliness in behavior often require summary indices, such as means, confidence intervals, and totals. The reason is that examining discrete responses over time offers unique challenges, although such a micro-level of analysis can provide insights to complement other measures. For example, **Figure 18** presents cumulative records of barrier reveals during the 4-sec delay component for all subjects across all eight sessions during the individual and team ratio conditions. That component was selected as the exemplar because it was the most challenging component, as evidenced by the NASA-TLX outcomes, and it was also associated with other performance changes. The measure of barrier reveals was selected because it is, perhaps, the most direct metric of a team member's disposition to assist other members in avoiding barriers. In the figure, the passage of time from the start of a session is presented on the y axis, and successive instances of a barrier reveal are indicated on the x axis.

The figure shows graphically that during the first three sessions, at least, S3 showed a high rate of reveals, and S1 showed a comparatively low rate of reveals. Furthermore, over those sessions in the individual condition, pausing is evident in the behavior of S1. These records suggest an inconsistency in performance tactics within and among subjects during a session that are not always evident in the summary indices.

The figure also shows the range of reveals among the subjects at the end of a session. With the exceptions of Session 4 and Session 7, the range is graphically greater for the individual condition, in comparison to the team condition. It is also notable that in the team condition for Session 7, all members showed the fewest reveals in the study, and the records virtually overlap.

Discussion

The results of this case study show that a behavior-analytic framework can be used to operationalize a three-person task that yields several indices of performance effectiveness. Accomplishing the session objectives under both individual and team ratio requirements required task integration at several levels. First, revealing barriers by a team member was essential to enable teammates to drag resource blocks to the target without striking a hidden barrier. Second, because the task had a completion requirement, a barrier hit by an individual member reduced the point accumulation, indirectly affecting all members of the team by extending the duration of a session. In that regard, a greater degree of task interdependency resulted, in comparison to the previous study that imposed a time limit on a session (Emurian et al., in press). For example, in the present scenario, barrier hits and degraded performance (e.g., a slow rate of point accumulation) by one team member would have a general effect on the team, not just on the individual team member exhibiting such performance, no matter the rate of barrier reveals exhibited by that member. Such a level of task integration is indicated when the task is intended to provide multidimensional indices of individual and team performance effectiveness.

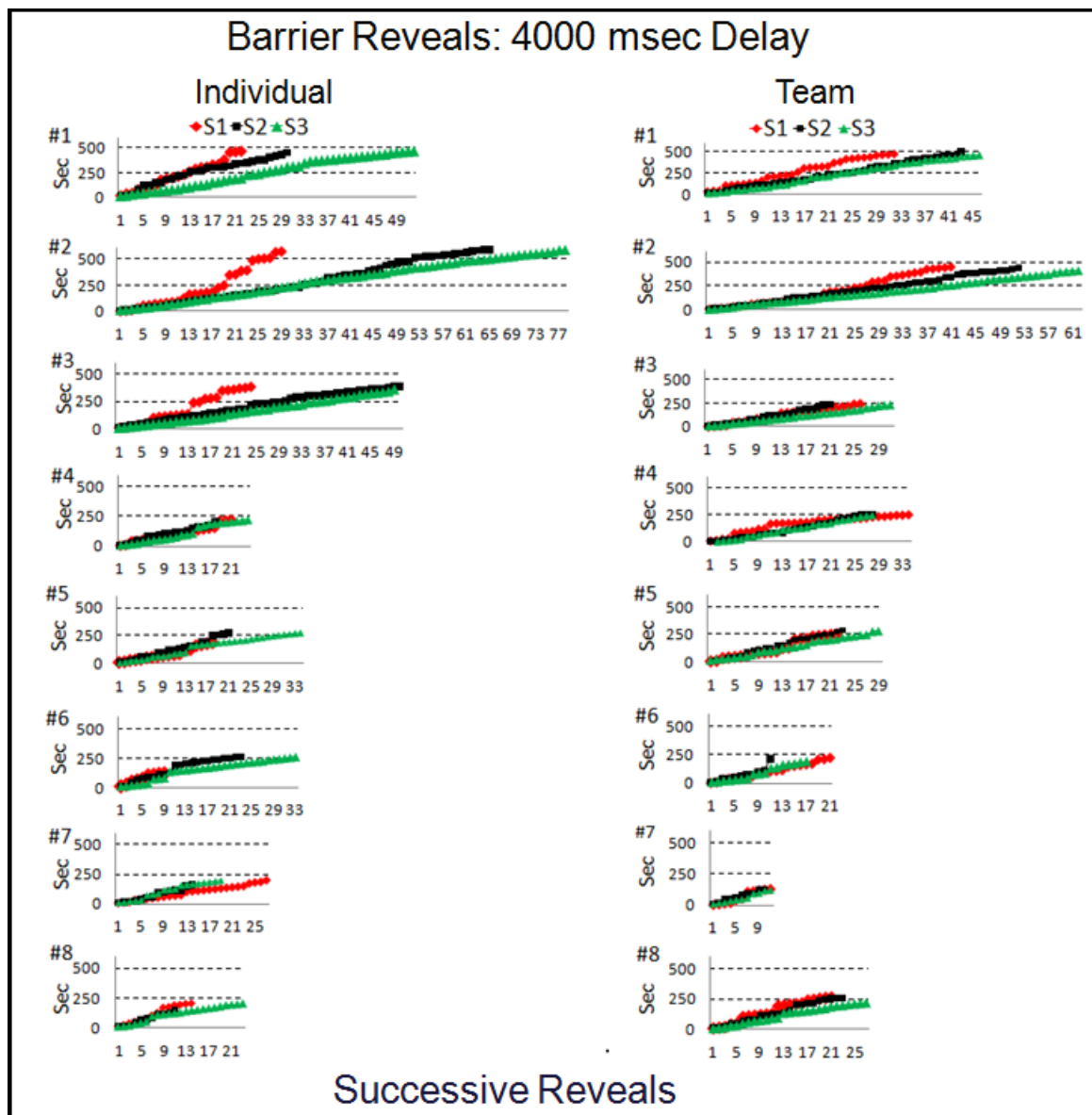


Figure 18. Cumulative records of barrier reveals for all subjects for the Individual and Team ratio conditions across the eight sessions.

Performance differences observed between the individual and team ratio conditions were evident in the range, or variability, of barrier reveals and barrier hits. During the 4-sec component, the mean range of barrier reveals was higher in the individual condition in comparison to the team condition (Figure 10), and the mean range of barrier hits was higher in the individual condition in comparison to the team condition (Figure 14). The debriefing statement by S1 after Session 8 (Appendix A) suggests that the individual and team ratio conditions differentially affected the way that team members perceived the task scenario and undertook to accomplish the ratio requirements under the two conditions.

When S2 was replaced at Session 5, the following differences were observed between Part 1 and Part 2 of the study. A wide range of points accumulated to the team ratio requirement was observed

during Session 5 and Session 6 (Figure 4). The mean duration to complete the 4-sec component was higher in Part 1 in comparison to Part 2 (Figure 6). The mean for barrier reveals during the 4-sec component was higher in Part 1 in comparison to Part 2 (Figure 8). The mean for barrier hits during the 1-sec component was higher in Part 2 in comparison to Part 1 (Figure 12). For S1, the mean magnitude of the NASA-TLX was higher in Part 1 in comparison to Part 2 (Figure 16). With respect to Belonging and Morale ratings, S1 and S3 showed higher ratings in Part 1 in comparison to Part 2 (Figure 17a,b).

Examination of the original team's performance over the first four sessions shows a modest "learning effect" for all subjects with respect to session durations, barrier reveals, and barrier hits, especially in the 4-sec component. Requests to reveal barriers were most frequent during Session 1, with few occasions of such responses after that session, although S3 did show reveal requests during all components of the team condition in Session 4. The data also show a wide range of individual differences in individual ratio completion durations, barrier reveals, and barrier hits over the first four sessions. And the range of such individual differences for barrier reveals and barrier hits was typically higher within the 4-sec component during the individual ratio condition, in comparison to the corresponding team ratio condition within and across sessions. The performances during Session 4, then, might be taken to be a "baseline" against which to assess the impact of replacing S2 at Session 5. It should be evident, however, that a "steady state" was apparently not reached, as evidenced by the variability in points contributed to the team ratio criterion by the subjects during Session 4.

When S2 was replaced on Session 5, the most immediate impact was observed in the points earned by the subjects during the team condition (Figure 4). In comparison to the .25-sec component, the discrepancies in accumulations are notable during the 1-sec and 4-sec components, when the performance by S1 seemed similar to Session 1, and the performance of S2 fell precipitously, especially during the 4-sec component. This outcome is related to the number of barrier hits by S2 during the 1-sec and 4-sec components (Figure 11) because each barrier hit reduced the point tally by 1 point. Barrier reveals exhibited by S1 during the team condition (Figure 7) were lowest among the subjects across the three delay components, although barrier hits by that subject were low and similar to those exhibited by S3. During Session 6, although the point accumulations by the three subjects across the .25-sec and 1-sec components were somewhat similar, they were widely disparate during the 4-sec component and in a direction opposite to that observed during Session 5. During the 4-sec component in Session 6, S3 showed the lowest point accumulation by that subject in the study, and the replacement subject, S2, showed the highest point accumulation of the study in comparison to all subjects. The high point accumulation on Session 6 by S2 may be attributable to the other two subjects deferring the movement of resource blocks, letting S2 perform most of that work during the 4-sec component. Taken together, these data show that even with two experienced team members, the reformed team struggled to accommodate the new member. It should be noted, however, that the duration for the reformed team to complete the team ratio requirement during the 4-sec component was never as high as the durations observed by the original team during Session 1 and Session 2. The transition to a new team member, although challenging and disruptive, did not prevent the team from completing its objective ("mission") in a timely fashion.

As indicated in the Procedure section, subjects were told not to discuss the task outside of the sessions, and communications were not permitted during the sessions. The outcomes observed during Session 5 and Session 6 provided the occasion to assess the impact of allowing the team to meet to discuss the members' tactics in reaching the ratio objective during the individual and team conditions. Accordingly, at the beginning of Session 7, the team members were instructed to discuss together their performance tactics in working on the task. The investigators left the laboratory and waited to be notified by the team that the meeting had finished. The meeting took no longer than ten minutes. The impact of the meeting is graphically evident in the points earned during the team condition during Session 7 (Figure 4). Variability in points earned decreased over the .25-sec and 1-sec components, and for the first time in the study, points were equivalent during the 4-sec component. Each team member contributed 20 points

to the 60-point team ratio requirement. Before Session 8, the team was allowed to meet again. The effect is again obvious in Figure 4, showing that during each delay component, each team member contributed 20 points to the team ratio requirement.

In that latter regard, the several performance metrics associated with the 4-sec delay component during the team condition on Session 7 all suggest an “optimally” performing team on that particular occasion. First, the team completed that component in the shortest time (134 sec), when compared to the other three sessions in Part 2 and to the four sessions in Part 1. Second, all team members showed the fewest barrier reveals during that component. Third, all team members showed the fewest barrier hits during that component. Fourth, no team member made a request to reveal barriers during that component. Although the performances were somewhat different during Session 8, all team members contributed 20 points to the team ratio requirement during the 4-sec component on that terminal occasion. The importance of identifying such a state of performance effectiveness is to be understood in terms of how team degradation or group fragmentation might be evidenced within the set of metrics made available in this version of the TPT.

The outcomes observed during Session 7 and Session 8, together with the debriefing statements (Appendix A), suggest that a shared “mental model” (DeChurch & Mesmer-Magnus, 2010; Zhou & Wang, 2010) of the task did not emerge from the feedback presented on the displays. Only after the team met as a group and discussed tactics to operate the task was a shared understanding of respective roles accomplished. This was the case even though the task required similar, if not identical, performances by all members. The sharing of expectations before Session 7 and Session 8 set the occasion for adopting optimal tactics, with respect to completing the task requirements with the least effort in the least amount of time. Such a meeting also provided the opportunity for the team members to reinterpret the explicit information being represented on the display screens in a more goal-directed fashion (Shah & Breazeal, 2010), rather than to exhibit competition in earning points, as suggested in the debriefing statements regarding the team condition ratio objective. Finally, the meetings may have occasioned team proactive performance, an emergent property of teams that reflects and shapes team interactions (Williams, Parker, & Turner, 2010).

The individual and team ratio conditions bear similarities to individualism-collectivism approaches to the analysis of teams (Hofstede, 1980), where a collective orientation was defined “...as the propensity to work in a collective manner in team settings” (Driskell, Salas, & Hughes, 2010, p. 317). Although the present task required component completion under individual and team ratio conditions, member interdependencies existed in both conditions. In fact, the debriefing statements by the subjects suggested that competitive factors came into play during the team condition, where a subject was inclined to attempt to be the team member with the highest point accumulation. There was a cost to such competitiveness, however, because degraded performance, in terms of barrier hits and failure to reveal barriers, indirectly impacted the team by requiring more work and time to reach the objective. Even though face-to-face interactions appeared to overcome at least some uncertainty regarding expectations of performance among the subjects, diminished ratings for Belonging and Morale continued to be evidenced by the two original subjects throughout Part 2 of the study. Individual differences in skill and collective orientation continued to manifest themselves throughout the study. With continued practice and communications, however, a steady-state performance might be anticipated to be reached, even with the presence of such individual differences, which would also include barrier reveals (Figure 9) and barrier hits (Figure 13) over the long term.

The use of games as a countermeasure to the psychosocial and cognitive effects of isolation and confinement has been suggested by Hauptlik-Meusburger, Aguzzi, and Peldszus (2010). Additionally, Voynarovskaya, Gorbunov, Barakova, and Rauterberg (2010) are evaluating the effectiveness of a three-person multi-player game to monitor the status of an isolated crew, with particular reference to its

applications to alleviate stress, and as an unobtrusive tool to monitor the mental capacity of astronauts and the development of different social interaction patterns within the crew. The current task may also be suitable as a rapid assessment tool given the range of metrics provided and the simplicity of its administration. During Session 7, for example, the team completed the team ratio in 134 sec during the 4-sec component (Figure 5), the most challenging of the three components as evidenced by the performance metrics and the NASA-TLX outcomes. Additional work is indicated to determine the extent to which the TPT can be used to detect emerging problems within a crew, once a steady state has been reached, and even to foster crew cohesion.

Behavior analysis continues to this day to be challenged in applying its foundational principles to the study of human behavior, individual and social. In that regard, previous research most closely related to the present scenario and published within the *Journal of the Experimental Analysis of Behavior* includes stimulus control of cooperation (Schmitt & Marwell, 1968), preferences for cooperation (Shimoff & Matthews, 1975), cooperative exchange (Matthews, 1977), altruistic responding (Weiner, 1977), and trusting behavior (Hake & Schmid, 1981), among others. Most relevant to the present scenario, perhaps, is the study reported by Burnstein and Wolff (1964) that involved the shaping of three-man teams on a multiple DRL-DRH schedule using collective reinforcement. Although behavior analysis appears to have much to offer in this domain of work, the final sentence in a review of human operant research by Buskist and Miller (1982, p. 141) published from 1958-1981 seems applicable today:

It would appear from the present census that the experimental analysis of human behavior has thus far fallen short of Skinner's 'active prosecution of a science of behavior.'

Hopefully, the next half-century will bring a different outcome.

References

- Brady, J.V. (2007). Behavior analysis in the space age. *The Behavior Analyst Today*, 8(4), 398-413.
URL: <http://www.baojournal.com/>
- Burnstein, D.D., & Wolff, P.C. (1964). Shaping of three-man teams on a multiple DRL-DRH schedule using collective reinforcement. *Journal of the Experimental Analysis of Behavior*, 7(2), 191-197.
- Buskist, W.F., & Miller, H.L. (1982). The analysis of human operant behavior: A brief census of the literature: 1958-1981. *The Behavior Analyst*, 5, 137-141.
- Cao, A., Chintamani, K.K., Pandya, A.K., & Ellis, R.D. (2009). NASA TLX: Software for assessing subjective mental workload. *Behavior Research Methods*, 41(1), 113-117.
- DeChurch, L.A., & Mesmer-Magnus, J.R. (2010). Measuring shared team mental models: A meta-analysis. *Group Dynamics*, 14(1), 1-14.
- Driskell, J.D., Salas, E., & Hughes, S. (2010). Collective orientation and team performance: Development of an individual differences measure. *Human Factors*, 52(2), 316-328.
- Emurian, H.H., & Brady, J.V. (2007). Behavioral health management of space dwelling groups: Safe passage beyond earth orbit. *The Behavior Analyst Today*, 8(2), 113-135.
URL: <http://www.baojournal.com/>

- Emurian, H.H., Brady, J.V., Ray, R.L., Meyerhoff, J.L., & Mougey, E.H. (1984). Experimental analysis of team performance. *Naval Research Reviews*, 36(1), 3-19.
URL: <http://nasa1.ifsm.umbc.edu/cv/NRR1984.pdf>
- Emurian, H.H., Canfield, G.C., Roma, P.G., Brinson, Z.S., Gasior, E.D., Hienz, R.D., Hursh, S.R., & Brady, J.V. (in press). A multi-player team performance task: Design and evaluation. In M.M. Cruz-Cunha, V.H. Carvalho, & P. Tavares (Eds.), *Business, Technological and Social Dimensions of Computer Games: Multidisciplinary Developments*, IGI Global.
URL: <http://www.igi-global.com/bookstore/TitleDetails.aspx?TitleId=46177>
- Emurian, H.H., Canfield, G.C., Roma, P.G., Gasior, E.D., Brinson, Z.S., Hienz, R.D., Hursh, S.R., & Brady, J.V. (2009). Behavioral systems management of confined microsocieties: An agenda for research and applications. *Proceedings of the 39th International Conference on Environmental Systems* (Paper number: 2009-01-2423), Warrendale, PA: SAE International, 2009.
URL: <http://papers.sae.org/2009-01-2423/>
- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Newbury Park, CA: Sage.
- Hake, D.A., & Schmid, T.L. (1981). Acquisition and maintenance of trusting behavior. *Journal of the Experimental Analysis of Behavior*, 35(1), 109-124.
- Hauptlik-Meusburger, S., Aguzzi, M., & Peldszus, R. (2010). A game for space. *Acta Astronautica*, 66, 605-609.
- Matthews, B.A. (1977). Magnitudes of score differences produced within sessions in a cooperative exchange procedure. *Journal of the Experimental Analysis of Behavior*, 27(2), 331-340.
- Maxwell, S.E., & Delaney, H.D. (2004). *Designing Experiments and Analyzing Data: Second Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Salisbury, W.W., Carte, T.A., & Chidambaram, L. (2006). Cohesion in virtual teams: Validating the perceived cohesion scales in a distributed setting. *The DATA BASE for Advances in Information Systems*, 37(2 & 3), 147-155.
- Schmitt, D.R., & Marwell, G. (1968). Stimulus control in the experimental study of cooperation. *Journal of the Experimental Analysis of Behavior*, 11(5), 571-574.
- Shah, J., & Breazeal, C. (2010). An empirical analysis of team coordination behaviors and action planning with application to human-robot teaming. *Human Factors*, 52(2), 234-245.
- Shimoff, E., & Matthews, B.A. (1975). Unequal reinforcer magnitudes and relative preference for cooperation in the dyad. *Journal of the Experimental Analysis of Behavior*, 24(1), 1-16.
- Suedfeld, P., Bootzin, R., Harvey, A., Leon, G., Musson, D., Oltmanns, T., & Paulus, M. (2010). Behavioral Health and Performance (BHP) Standing Review Panel (SRP) Final Report. NASA Center: Johnson Space Center; Publication Year: 2010; Added to NTRS: 2010-02-22; Document ID: 20100004763; Report Number: JSC-CN-19726.
URL: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20100004763_2010003527.pdf

- Voynarovskaya, N., Gorbunov, R., Barakova, E., & Rauterberg, M. (2010). Automatic mentalassistant: Monitoring and measuring nonverbal behavior of the crew during long-term missions. *Proceedings of Measuring Behavior*, Eindhoven, The Netherlands, August 24-27, 2010, 77-81. URL: <http://www.amha.id.tue.nl/>
- Weiner, H. (1977). An operant analysis of human altruistic responding. *Journal of the Experimental Analysis of Behavior*, 27(3), 515-528.
- Williams, H.M., Parker, S.K., & Turner, N. (2010). Proactively performing teams: The role of work design, transformational leadership, and team composition. *Journal of Occupational and Organizational Psychology*, 83, 301-324.
- Zhou, Y., & Wang, E. (2010). Shared mental models as moderators of team process-performance relationships. *Social Behavior and Personality: An International Journal*, 38(4), 433-444.

Author Note

This study was supported in part by the National Space Biomedical Research Institute through NASA NCC 9-58-NBPF01602. The authors acknowledge the assistance of Christian E. Demeke, an undergraduate major in Information Systems at UMBC, in testing the TPT and in conducting this study. We also acknowledge the contributions of Emily Toy and Oana Tibu, UMBC students, for their previous assistance in this stream of task development and research.

Contact Information

Henry H. Emurian
Information Systems Department
College of Engineering and Information Technology
UMBC
1000 Hilltop Circle
Baltimore, Maryland 21250
emurian@umbc.edu

Kip Canfield
Information Systems Department
College of Engineering and Information Technology
UMBC
1000 Hilltop Circle
Baltimore, Maryland 21250
canfield@umbc.edu

Joseph V. Brady
Behavioral Biology Research Center
Johns Hopkins University School of Medicine
5510 Nathan Shock Drive
Baltimore, Maryland 21224
jvb@jhmi.edu

APPENDIX, NEXT PAGE!

Appendix A

Debriefing Statements

Subjects were invited to submit email comments after each of the four sessions in Part .

S1 after Session 5: "I noticed that because we did not converse with each other too much before the experiment, such as what our majors were and a bit more background on each other, I did not feel as part of the group after the first run. People are different and I see that when a connection between the subjects is established, it allows for better teamwork, and allows me to feel more a part of the group."

S3 after Session 6: "I feel that before a team attempts to complete a task, there should be a sense of camaraderie even if it is just a little bit. Although we are not allowed to communicate with each other, if we were to get to know each other a bit more, figure out each other's intentions and be able to anticipate our team member's strategy, it might make the task run more smoothly."

The past 2 days, at times it has felt very competitive at times, even during the team trials. If everyone felt like a unit/team, we might work as a unit/team."

S1 after Session 6: "When we participated in the study again, I noticed one major problem and issue I had. This was that the more points my teammates were earning and the less I had I became more frustrated. After getting frustrated I would rush and try to earn more points, but the result was only negative points. I became more competitive when my teammates earned more points. Also I became greedy when my teammates did not reveal their barriers. I did not as well because as a team we were supposed to help each other but, during the team round for some reason it felt like everybody for themselves."

S1 after Session 7: "Yesterday when participating in the study, I believe when planning our strategy it was more organized, and we finished much faster when we discussed and carried through our plan. It affected the team because we all came with a single idea together, so that each one of us could help the team and finish at a faster and more efficient time. The meeting we were able to have helped us a lot and gave us a chance to get to know each other's technique to dragging however many blocks."

S2 after Session 7: "The team meeting definitely helped me, and I think all of us, feel more like a team and like we had a real game plan. We realized we all had a basic idea of how things should work, but new ideas were introduced, like for example, [S3] put out there that we should click our barriers as soon as they disappeared for the 25 and 100 levels, and just do your best on the 400. I contributed the idea of only going to 20 on the team round and once you reached that number, to help the others by revealing the barriers. We all agreed on those terms, and it definitely increased our efficiency, which then in turn increased our patience and feeling of a team. Prior to talking to one another, none of us knew each other and it was more frustrating when the others didn't quite do things the way you were doing them or wanted things to run."

S3 after Session 7: "Today when we had a moment to discuss some techniques and plans on how to approach the task, it helped to be more efficient at time. S2 suggested that during the team task, once each player reached 20 points, we stop and just reveal the barriers so we all get about the same number of points. This made the task less competitive than it has been the past few times. Before it felt like everyone was trying to get the most points, but this time when we each had a SPECIFIC number of points to gain, we worked more as a team. This was the biggest change I noticed after we discussed our game plan a bit." S1 after Session 8: "Since we had a chance to discuss with the group again we were able to refresh what plan we had the previous day. The plan we had worked so well and made each of us I think, very successful. We discussed that for the team trials we all would reach a limit of 20, after reaching 20 we would then continue to reveal and let the other members earn points. The result of this was finishing at a faster time, and less frustration. We did not discuss a plan for the individual trials because we all did not have a major issue or frustration with that portion of the task."

S1 after Session 8: "Since we had a chance to discuss with the group again we were able to refresh what plan we had the previous day. The plan we had worked so well and made each of us I think, very successful. We discussed that for the team trials we all would reach a limit of 20, after reaching 20 we would then continue to reveal and let the other members earn points. The result of this was finishing at a faster time, and less frustration. We did not discuss a plan for the individual trials because we all did not have a major issue or frustration with that portion of the task."

Copyright of Behavior Analyst Today is the property of Joseph D. Cautilli and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Posted with permission of the publisher.

[www.baojournal.com/BAT Journal/BATissues.html](http://www.baojournal.com/BAT%20Journal/BATissues.html)