Title
Cognitive Correlates of Team Performance: Variations Across Dimensions of Skill and Degrees of Expertise

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Banerjee, Sounak
Sims, Chris R.
Gray, Wayne D.

Publication Date
2026

Cognitive Correlates of Team Performance: Variations Across Dimensions of Skill and Degrees of Expertise

Sounak Banerjee (sounak.banerjee@nyu.edu)[1], Chris Sims[2] & Wayne Gray[2]
[1]Department of Psychology, New York University
[2]Cognitive Science Department, Rensselaer Polytechnic Institute

Abstract

While individual cognitive metrics have been shown to predict real-world task performance, their scalability to team behavior remains poorly understood. This study addresses the critical disconnect between individual cognitive abilities and team-based outcomes. We present an approach for studying this relationship using rich behavioral data collected from a virtual team-based task. Then, we further extend the analysis to show that these relationships can evolve with changes in team expertise. Regression models were used to predict performance across two dimensions of team skill, derived from an individual's ability to (1) perform task-specific actions (taskwork), and (2) coordinate their actions with others (teamwork). Individual performance across seven cognitive tasks were used as predictor variables. Results indicated significant associations between efficient information processing abilities and taskwork skill, in the game. Teamwork skill initially varied with visuospatial working memory. However, attention control predicted teamwork skill after sufficient practice in the task. Finally, independent measures of analytic intelligence reliably predicted both skills; however, task knowledge may have mediated that effect. Our results demonstrate that team performance is multifaceted, with distinct cognitive abilities contributing to various dimensions of team skill, at various stages of expertise.

Keywords: Cognitive Ability; Complex Task Performance; Expertise; Team Coordination; Multiplayer Games; Overcooked

Introduction

Traditionally, researchers have relied on one of two experimental designs to study the relationship between cognitive abilities and complex task performance (1) cross-sectional studies, where the relationship is studied at a particular point in time (Schmidt, 2002; Tucker & Warr, 1996; Voss et al., 2010), or (2) using baseline measures of cognitive abilities to predict future learning outcomes (Ackerman, 1988; Gonzalez et al., 2005; Murphy, 1989) . More recent research that focus on changes in cognitive abilities over time have found that training in complex tasks can lead to enhanced cognitive abilities related to attentional control, perception, spatial cognition, and problem solving abilities (Bavelier & Green, 2019; Bediou et al., 2018; Choi et al., 2020; Kowal et al., 2018). However, it is less clear how these changes in cognitive abilities interact with task skills and affect task performance, and if current results extend to cooperative multi-agent tasks. Here, we present an approach for investigating these complex relationships using behavioral data from a team-based task.

To study team behavior in complex dynamic tasks, we developed a cooperative cooking game based on the popular multiplayer game Overcooked, which is often used to study multi-agent coordination in both cognitive science (Banerjee & Gray, 2023; Bishop et al., 2020; Carroll et al., 2019) and reinforcement learning (Gessler et al., 2025). Our implementation of the task replicates the complex task dynamics of the original commercial game with tight experimental control, while also allowing us to collect behavioral data with high temporal and spatial precision.

Using a game-based experimental paradigm allowed us to better simulate the experience of a real-world task in a lab-based environment, as opposed to more traditional lab-based tasks (Lukosch et al., 2018). Indeed, computer video games have proven to be a powerful tool for studying cognitive abilities and their effect on complex task performance over the last decade (Bediou et al., 2018; C. S. Green & Bavelier, 2003; Hambrick et al., 2010; Valls-Serrano et al., 2022). In addition, game-based paradigms offer several advantages for studies that involve human expertise (Gray, 2017).

Further, training in action video games have shown potential for generalization of cognitive skills across domains (Bavelier et al., 2018). For instance, training in simulation games have been shown to improve communication (Chamberland et al., 2018), as well as decision-making and risk assessment skills (Reynaldo et al., 2021) among medical professionals. In fact, games can also promote general task learning for complex tasks (C. Green & Bavelier, 2012), and the transfer of positive collaborative learning experiences in paradigms dissimilar to the original training task (Peppler et al., 2013). Therefore, we believe that training in our task will lead to improved cognitive task performance among our participants.

For our experiments, we asked participating teams to play the cooperative game over a course of eight 1-hour sessions, spread out across several weeks. Since our objective was to study changes in the relationship between cognitive abilities and cooperative task skills resulting from task-specific training: we collected high quality behavioral data from all out participants for the entire duration of the study. We measured cooperative task skill using two distinct dimensions of individual performance: taskwork skill and teamwork skill (de Toledo et al., 2022; Morgan et al., 1993). Additionally, to track cognitive abilities, we had participants take seven task-relevant cognitive tests at different experimental phases.

To explore the relationship between cognitive abilities and cooperative task skills at various stages of expertise, we oper-

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

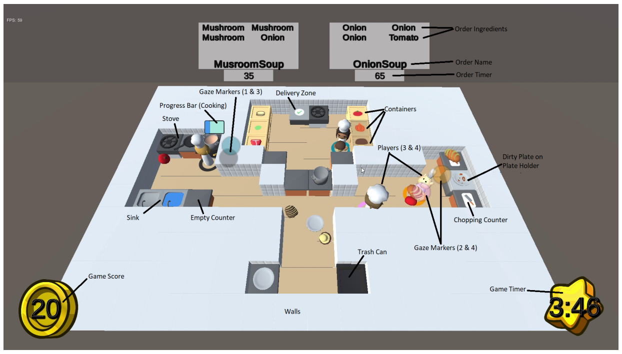

Figure 1: A (labeled) screenshot of a game in the ‘Clover’ kitchen layout.

ationalized the problem into three specific research questions:

1. RQ1: Do cognitive abilities improve after several hours of practice in a complex cooperative task?
2. RQ2: Do specific cognitive abilities support teamwork skill in our task, and do they change with expertise?
3. RQ3: Do specific cognitive abilities support taskwork skill in our task, and do they change with expertise?

Trends in our data reveal that distinct cognitive abilities support different dimensions of team skill at various stages of expertise. Our results further show that different cognitive abilities predicted taskwork and teamwork skills. For example, while abilities related to visual attention and memory were found to be associated with teamwork skills, cognitive information processing and attention control were more closely associated with taskwork skills.

Finally, participants also presented enhanced cognitive abilities after several hours of practice in the complex task, which is consistent with previous findings that show that practice in a complex task can help develop task-agnostic cognitive skills (Bavelier & Green, 2019; Bediou et al., 2018).

Methodology

The Cooperative Action Task

The Cooperative Action Task (CAT) is a game-based experimental paradigm developed to study changes in small-team coordination, under controlled lab environments. The CAT is a complex and dynamic task where four players must work in parallel and coordinate their actions with other team members to prepare and deliver orders in a virtual kitchen.

For each “kitchen”, orders appear at the top of the screen along with a timer indicating the time remaining to prepare the order. Players execute a series of actions to prepare each order as it becomes active. For example, Figure 1 shows two outstanding orders, a mushroom soup (expires in 35 seconds) and an onion soup (expires in 65 seconds). To prepare the mushroom soup, players would have to chop three mushrooms and one onion (at the chopping counters), cook them in a pot (on a stove), plate the soup, and carry it to the delivery zone. Progress bars are used to indicate the progress of different

Figure 2: The Divided kitchen layout.

actions. Teams are awarded points for each order they deliver, which is proportional the number of ingredients in the order.

Different kitchen layouts were developed to introduce unique challenges to team coordination. These challenges were accomplished by adding constraints in the kitchen design, which include counter/floor space restrictions, narrow corridors, isolated players, and kitchens, which are partitioned so that players are unable to move from one space to another. Sixteen unique kitchen layouts were used for the study.

For the current study, we focus on a single dimension of design constraint, which limits direct player interaction. The Divided kitchen layout (Figure 2) was designed to prevent team interaction by completely isolating all 4 players in their own mini-kitchens, forcing them to work alone. Therefore, player performance in these kitchens provided a reliable of measure individual (taskwork) skill, while, cooperative (teamwork) skill was measured from games in other kitchen layouts.

Cognitive Task Battery (CTB)

Seven cognitive tasks were selected for our experiment based on (1) the relevance of the measured cognitive ability to the current complex task, and (2) previous reports of statistically significant correlations with performance in similar complex tasks. Our rationales for choosing each task are also provided.

All tasks, except one, were administered as part of a Cognitive Task Battery (CTB), twice during the course of the study. Participants were asked to complete the CTB at the beginning and the end of the study, to test for changes in cognitive ability. The Raven’s Advanced Progressive Matrices (RAPM) task was only administered once (halfway through the study), because we did not expect any changes in intelligence score (Buschkuehl & Jaeggi, 2010; Haier, 2014).

A common concern with test-retest measures of cognitive abilities stems from the possibility of practice effects biasing performance during retest sessions (the second CTB session in this case), where familiarity with the task, as a result of repeated testing, could potentially improve performance. However, practice effects in such tasks have been shown to disappear after 4 weeks (Marina G. Falleti & Darby, 2006). Given that a 5-6 week gap was maintained between the two CTB sessions in the current study, we do not expect practice effects in the CTB to be a concern. In addition, 'practice' trials were added to the beginning of each task to allow participants to familiarize themselves with the task and further mitigate practice effects (De Cesarei et al., 2023; Zorowitz & Niv, 2023). Practice trial data was excluded from analysis.

Performance in all tasks except the Corsi Block-Tapping task was measured using an inverse efficiency metric similar to the one described by Large et al. (2019). To compute this metric, we divided the median reaction time of correct responses (lower is better for the numerator) in a task by the proportion of correct responses (higher is better for the denominator). Thus, a higher value of this metric was associated with lower performance in a task. It is a convenient measure of performance, that combines information about reaction time and response accuracy into a single composite variable.

Anti-saccade Task The anti-saccade task (Hallett, 1978) measures individual abilities for visual attention control (Munoz & Everling, 2004), which is crucial for the performance of complex visuospatial tasks (Arthur et al., 1995; Hubert-Wallander et al., 2011). In fact, studies on the relationship between performance in saccadic tasks and expertise in complex visuospatial tasks have found a direct correlation between the two (Chen et al., 2024; Di Russo et al., 2003). We therefore hypothesized that players with superior visual attention control may have an advantage over others, since it involves navigating and selectively attending to specific aspects of a complex visuospatial environment.

Corsi Block-Tapping Task The Corsi block-tapping task is used to measure visuospatial working memory capacity (Corsi, 1972), which has been shown to predict performance in multiple-object tracking tasks (Li et al., 2023; Trick et al., 2012). In the CAT, items are constantly spawned, moved, and despawned by other players. Therefore, a player's ability to quickly recall the type and position of various items in the kitchen could significantly affect game performance.

Performance in this task was measured by computing the average sequence length across all trials (meaning that a higher score for this task is indicative of better performance). The inverse efficiency metric was not used here because reaction time was not relevant to this task.

Hick-Hyman Task The Hick-Hyman task was developed to assess cognitive information processing capacity (Hyman, 1953). This ability has been linked to higher performance in complex decision-making tasks (Adams & Ericsson, 1992; Yang et al., 2021), where individuals have to quickly assimilate relevant information from rich information spaces and decide among multiple actions. Therefore, greater performance in this task may translate to higher performance in the CAT, since players need to be able to make quick and efficient action choices based on complex task-state information to ensure optimal performance.

Fitts' Reciprocal Tapping Task The reciprocal tapping task is often used to compare motor control ability across individuals (Smith & Claytor, 2018; Temprado et al., 2013; York & Biederman, 1990). For the current study, we asked participants to perform this task using the joystick on Xbox controllers, to determine the extent of their ability to control objects in the virtual world. We ensured that the controllers used in this experiment were the same ones used to control player movement in the CAT, for consistency.

Eriksen Flanker Task The Eriksen Flanker task (Eriksen & Eriksen, 1974) is a response interference task, where irrelevant stimuli have to be inhibited in order to attend and respond to relevant target stimuli. Such selective attention mechanisms play a key role in decision making in tasks with complex perceptual spaces (Lev-Ari et al., 2022; Rahimi et al., 2022). As a result, performance in this task may reflect a player's ability to filter out irrelevant elements from complex environments and only attend to relevant information, which is crucial for making effective choices in the CAT.

Go-NoGo Task The Go-NoGo task (Logan & Cowan, 1984) is a response inhibition task, that can capture individual differences in sustained attention and response control. These skills may be highly relevant to the process of assembling ingredients for orders in the CAT, as players with better response control may be able to better inhibit actions that lead to incorrect combinations of ingredients. this is consistent with previous research, which found significant associations between inhibitory control and action errors (Bravi et al., 2022), as well as decision-making accuracy (Cao et al., 2024), in highly dynamic tasks.

Raven's Advanced Progressive Matrices The Raven's Advanced Progressive Matrices is a cognitive assessment tool for measuring abstract reasoning ability and analytic intelligence (Carpenter et al., 1990). Individuals who score higher in this task may present superior critical thinking skills (Liu et al., 2024) and consequently choose more optimal playing strategies in the game, leading to superior performance.

Experimental Setup

Four acoustic pods (small sound-isolated rooms), one assigned to each team member, were used to prevent verbal communication during game sessions. Each pod was also fitted with a monitor and an Xbox game controller, that the players used to play the game. Monitors from all 4 pods were connected to a central computer (using an HDMI splitter), that ran the game. This allowed all four players to simultaneously receive the same video feed for the game from inside their pods. All 4 controllers were also connected to the central computer, via Bluetooth.

Procedure

The study required participants to come to the lab for 11 one-hour sessions, spread out over a period of 5-6 weeks. The

11 sessions were executed in the following order: (1) In the first session, participants completed a Cognitive Task Battery (CTB) which included 6 cognitive tasks; (2) in the next 4 sessions (sessions 2-5), participants played the game; (3) during the sixth session, participants completed the Raven's Advanced Progressive Matrices task; (4) this was followed by 4 more game sessions (sessions 7-10); (5) in the final session (session 11), the CTB from session 1 was repeated.

Each game session involved participants playing eight 5-minute games (40 minutes total). Each player played the game from inside their assigned (by the experimenter) pods. After playing the first 4 games, players were asked to step out of their pods and take a short break before returning to their pods to play the last 4 games of the session.

Participants

The participants were 64 university students (30 female, 29 male, 3 non-binary, and 2 participants who chose not to answer). Participant age ranged from 18-31 years (mean=20.3, SD=2.75), where 56 of the 64 participants were between 18 and 22. Only three of the 16 participating teams were homogeneous (all male/female) groups. Participants were financially compensated for their time. Each participant received $15 US for each session, along with a $50 US bonus on completion of the entire study. All experimental procedures were reviewed and approved by the University IRB.

Data

Data from each team consists of sixty-four 5-minute games (8 games played in each of the 8 sessions). All player action and game state data was recorded at 60Hz. The complete dataset contains 5 hours and 20 minutes (sixty-four 5-minute games) of gameplay data per team, for 16 teams. Due to a technical fault, data from one game was corrupted, leaving us with valuable data from 1023 games.

For the Cognitive Task Battery, data from a single participant was lost for the Anti-saccade Task due to a technical fault. Additionally, for the Go-Nogo task, data for 3 participants was corrupted, and could not be recovered. Finally, data from one participant was also lost for the RAPM task. Missing data points were excluded from model fits during analysis.

Analysis

Paired samples t-tests were performed to test for changes in cognitive abilities. However, one of the assumptions of this test is that the data should have a normal distribution. Therefore, a Shapiro-Wilk test was carried out to check for normality in the distributions of performance values for the six cognitive tasks. Results indicated that cognitive task scores were either positively or negatively skewed. To normalize the data, the single parameter Box-Cox transformation (Box & Cox, 1964) was applied to each set of cognitive task scores. The resulting (lambda) parameter values for the transformations are 0.18 (Anti-saccade), -2 (Corsi, Hick-Hyman, and Flanker), 0.02 (Fitts'), -0.67 (Go-Nogo), 0.06 (RAPM).

Task	Summary Stats				Results
	CTB 1		CTB 2		
	M	SD	M	SD	
Anti-Saccade *	10.06	.9	9.44	.79	$t(62) = 6.97;\ p < .001$
Corsi	.47	4.04E-3	.47	4.05E-3	$t(63) = -1.81;\ p = .076$
Hick-Hyman	.5	6.8E-7	.5	5.97E-7	$t(63) = 1.16;\ p = .251$
Fitts' *	7.4	.27	7.17	.32	$t(63) = 5.18;\ p < .001$
Flanker *	.5	7.6E-7	.5	6.24E-7	$t(63) = 3.37;\ p = .001$
Go-Nogo	1.48	9.67E-4	1.48	9.18E-4	$t(60) = .717;\ p = .476$

Table 1: Result summary for paired samples t-tests, to test for significant changes in performance between the first and second CTB sessions, for all six tasks. Some of the variables have extremely low standard deviations. This is a result of the Box-Cox transformation, which caused significant narrowing of the distributions of these variables. Tasks marked with an '*' presented significant changes ($\alpha \leq .001$) across sessions.

Regression models were used to test for dependence between measured cognitive abilities and dimensions of task skill. Specifically, individual linear regression models were fit to predict each dimension of cooperative task skill based on individual measures of cognitive ability. Simple linear regression models were used here, as opposed multivariable models, in the interest of parsimony and interpretability.

Individual models were used for each cognitive ability, because, possible presence of partial overlaps in the executive processes involved in various cognitive tasks can lead to correlations between performance in these tasks (Burgoyne et al., 2022). These correlations can, in turn, lead to multicollinearity among independent variables, which can have unintended consequences on the covariance structure of regression models, creating challenges for reliable interpretation of model parameters (Morrison, 2003; Nye et al., 2022). Approaches for dealing with multicollinearity in multiple-regression models (Paul, 2008; Schroeder et al., 1990) also come at the cost of interpretability of the resulting models. Therefore, we follow Morrison's recommendation of fitting multiple simple regression models instead (Morrison, 2003).

Finally, simple regression models were also used here (as opposed to multilevel models) to avoid issues with estimation biases and misrepresentation of result significance often cause by small sample-sizes at level 2 (Maas & Hox, 2005; Stegmueller, 2013). Since we only have data from 16 teams, fitting multilevel models can lead to erroneous results.

Results

RQ1: Change in Cognitive Abilities

Paired samples t-tests were performed on the transformed CTB data to test for statistically significant changes in performance between the two CTB sessions (sessions 1 and 11) for the six cognitive tasks (results provided in Table 1). Additionally, since we are performing multiple comparisons, we run the risk of erroneous inferences due to sample bias. To mitigate this risk, the Bonferroni Correction (Dunn, 1961) was employed to adjust the α value for significance testing, with an initial α of 0.05. The corrected value of α was approximately 0.008.

Task	Test-Retest (CTB) Tasks : Fit Summary					
	Set 1			Set 2		
	Coeff.	p-Val.	Adj. R^2	Coeff.	p-Val.	Adj. R^2
Anti-Saccade	-0.42	.058	.041	-0.48	**.04***	.052
Corsi	125.84	**.009***	.089	-18.38	.693	-.013
Hick-Hyman	-3.94E+5	.181	.013	-4.44E+5	.156	.017
Fitts'	-0.9	.226	.008	0.599	.303	.001
Flanker	-3.08E+5	.242	.006	-6.14E+5	**.039***	.052
Go-Nogo	-223.1	.288	.002	-188.5	.359	-.002

Task	Single-Session Task : Fit Summary					
	Session 2			Session 10		
	Coeff.	p-Val.	Adj. R^2	Coeff.	p-Val.	Adj. R^2
RAPM	-1.08	**<.001***	.152	-0.057	.858	-.016

Table 2: Model summary (coefficient for task score, along with p-value and adjusted R^2 for the model) for all regression models fit to predict coordination performance in the CAT based on performance scores from all 7 cognitive tasks. Significant p-vales ($\alpha < .05$) are marked with an asterisk (*).

Task	Test-Retest (CTB) Tasks : Fit Summary					
	CTB 1			CTB 2		
	Coeff.	p-Val.	Adj. R^2	Coeff.	p-Val.	Adj. R^2
Anti-Saccade	-0.195	.236	.007	-0.324	.077	.035
Corsi	59.77	.096	.029	90.52	**.01***	.087
Hick-Hyman	-4.43E+5	**.037***	.053	-6.42E+5	**.007***	.097
Fitts'	-0.675	.214	.009	-0.362	.426	-.006
Flanker	-2.76E+5	.148	.018	-2.22E+5	.343	-.001
Go-Nogo	-167.4	.286	.003	-133	.427	-.006

Task	Single-Session Task : Fit Summary		
	Coeff.	p-Val.	Adj. R^2
RAPM	-0.56	**.022***	.068

Table 3: Model summary (coefficient for task score, along with p-value and adjusted R^2 for the model) for all regression models fit to predict task performance in the CAT based on performance scores from all 7 cognitive tasks. Significant p-vales ($\alpha < .05$) are marked with an asterisk (*).

Significant changes were observed for performances in the Anti-saccade, Fitts', and Flanker tasks, with $\alpha < 0.008$.

RQ2: Cognitive Abilities and Teamwork Skill

Teamwork skill in the CAT was derived from the average number of orders (per game) a player contributed to, in a single session, which consists of 8 games. This is distinct from an individual's task skill, as it necessitates anticipating task requirements based on other's actions (in real time), and adapting one's own actions in response. Since cognitive abilities were measured twice during the study (sessions 1 and 11), we measured teamwork performance in the two (temporally) nearest game sessions (sessions 2 and 10) to test for dependence between cognitive abilities and teamwork skill.

Two sets of six regression models (two models per CTB task) were fit to the data: (1) Set 1: performance scores from the first CTB session (session 1) were used to predict teamwork performance for the first game session (session 2). (2) Set 2: performance scores from the second CTB session (session 11) were used to predict teamwork performance in the last game session (session 10). Additionally, since the RAPM task was only performed once (in session 6), the performance score from that session was used to predict teamwork performance for both sessions 2 and 10. All model fit results are summarized in Table 2.

Results of the first set of regression models suggest that player performance in the Corsi Block-Tapping task in the first CTB session and the RAPM task (from session 6) significantly predicted teamwork performance in the first game session. However, the model fits from set 2 imply that scores for Anti-Saccade and Flanker tasks from the second CTB session predicted teamwork performance in the last game session. This indicates a potential shift in reliance on specific cognitive abilities (from analytical intelligence and memory to attention control) at different levels of teamwork skill. In addition, the small effect sizes ($R^2 \sim 0.05$) of significant fits in the second CTB session also suggest that other factors (such as team familiarity (Banerjee & Gray, 2023) or cognitive abilities not included in the current analysis) may have supported teamwork skill at later stages of practice.

Further, coefficients for regression models fitted to the Anti-Saccade, RAPM, and Flanker task data were negative, while the coefficient for the Corsi task was positive. In each case, the coefficients suggest that performance in the task was positively linked to teamwork performance in the game. The coefficients for the three tasks, except the Corsi task, are negative because lower values of the inverse efficiency metric indicate greater performance in the tasks.

RQ3: Cognitive Abilities and Taskwork Skill

Individual task performance in the game was measured in the divided kitchen layout. Taskwork skill was calculated as the average number of orders prepared by a player across all 4 games played in this kitchen.

The divided kitchen was incorporated in games from the 5th and 6th game sessions (Sessions 7 and 8). Therefore, by virtue of temporal proximity, it would be rational to rely solely on data from the second CTB session (Session 11) to test for dependence between taskwork skill in the CAT and cognitive abilities of players. However, to ensure scientific rigor, regression models were fit to data from both CTB sessions (Sessions 1 and 11). This way, any changes in predictive significance between the two CTB sessions for the current analysis can be cross-referenced with significant changes in performance across the sessions reported in Table 1. Any inconsistencies between the two could be an indication of spurious results. The results of this analysis are summarized in Table 3.

Regression fit results suggest that RAPM task scores did predict taskwork performance in the game. Further, performance in the Hick-Hyman task for both CTB sessions displayed significant predictive power. Finally, performance in the Corsi Block-Tapping task for the second CTB session was also a significant predictor of taskwork performance. However, given that there was no significant change in performance in the Corsi task between the two sessions (Table 1), the predictive significance of the task for taskwork performance should have also remained the same for both regression models in Table 3. This implies that the model significance for the two regression models for the Corsi task may be unreliable. The coefficients for regression models fitted to the Hick-Hyman and RAPM task data are negative, suggesting a direct relationship between taskwork performance and performance in these tasks.

Discussion and Concluding Remarks

We performed statistical analyses to identify associations between cognitive abilities and performance in a complex cooperative task. A cooperative cooking game called the 'Cooperative Action Task' (CAT) was used for the study. Teamwork performance in the CAT was measured at the beginning and end of the study (sessions 2 and 10), while taskwork performance was measured once towards the middle of the study (sessions 5 and 6). A Cognitive Task Battery (CTB) of 6 tasks was also administered at the beginning and end of the study (sessions 1 and 11). Finally, participants were also administered the Raven's Advanced Progressive Matrices test during the 6$^{\text{th}}$ session. Our analyses identified significant direct relationships between cognitive abilities and task performance.

Associations between teamwork skill (in the CAT) and cognitive abilities (Table 2) suggest that players may have relied on visuospatial working memory to coordinate with their peers when they first started playing the game, but by the end of the study, attentional control abilities became the relevant discriminator for teamwork performance in the game. The implication being that players may have initially found it challenging to keep track of items in the kitchen, but with familiarity with the task, they found strategies that allowed them to do so. However, increased skill with time may have led to higher gameplay activities, which placed greater loads on attentional control abilities for efficient coordination.

The increased dependence on attentional control was further supported by marked improvements in performance for the Anti-saccade and Flanker tasks in the second CTB session (Table 1). While previous research does provide evidence which suggests that both visuospatial working memory (Blacker et al., 2014; Toril et al., 2016) and attentional control abilities (Dale et al., 2020; Valls-Serrano et al., 2022) support performance in complex tasks such as video games, the current results further indicate that these relationships may change with expertise in a task.

Results of regression fits for taskwork skill in the CAT (Table 3) suggest that only information processing capacity predicted performance in the CAT. Higher information processing capacity may have allowed some players to develop more efficient game strategies in this case, as complex tasks usually require making decisions based on many contextual factors. This result is consistent with reports from previous work showing large effects of game skill on the efficiency of information processing, with increased decision complexity (Jiang et al., 2020; Powers et al., 2013).

Analytic intelligence (measured by the RAPM task) was found to be a reliable predictor of both taskwork and teamwork performance (Tables 2 and 3). However, it only predicted teamwork performance in the first game session. This may be because players with superior analytic intelligence were able to fully understand the game mechanics sooner than their peers, however, intelligence was no longer beneficial once players sufficiently learned the task. In fact, it has been suggested that prior knowledge about a complex task mediates the relationship between intelligence and performance in the task (Weise et al., 2020).

In the case of the Fitts' task, although a considerable improvement in performance was observed in the second CTB session (Table 1), it did not predict teamwork or taskwork performance in the CAT. No significant changes were observed in performance for the Go-Nogo task between the CTB sessions, and performance in the task did not predict teamwork or taskwork performance in the CAT. Indeed, earlier investigations also found no significant association between response inhibition and performance in action (Colzato et al., 2012) and strategy games (Leong et al., 2022).

The results can be summarized as follows: (1) certain cognitive abilities may be improved with training in complex tasks; (2) distinct cognitive abilities support taskwork and teamwork performance in complex cooperative tasks; and (3) humans likely rely on different cognitive abilities to engage the same behavior at different levels of expertise. Although, numerous studies have confirmed the beneficial effects of complex task training on cognitive abilities in the past (Bavelier & Green, 2019; Bediou et al., 2018; Choi et al., 2020; Kowal et al., 2018), to our knowledge, there is no previous research on the variations in cognitive abilities that support team performance across the two dimensions of team skill (taskwork and teamwork), at different stages of expertise.

Limitations Finally, given the novelty of the adopted research methodology and sample size limitations, caution is advised about drawing strong conclusions from these results. The present study makes no claims regarding the cross-task generalizability of the observed relationships between specific cognitive abilities and individual cooperative skills. Instead, the purpose of this study is to hopefully demonstrate that these relationships do in fact exist, and that investigations into such relationships in various tasks may help us better understand team performance in cooperative tasks. In addition, team tasks are more dynamic and complex than individual tasks as a result of inherent variabilities in team structure (Cooke et al., 2023). This property presents an additional dimension of complexity, which may be exploited to test the robustness of these results. Follow-up studies might also consider a greater range of cognitive abilities, maybe ones that account for social skills. In summary, this study may be considered a demonstration of an approach and an early exploration of the relationship between cognitive abilities and cooperative task skills. Future endeavors can build and expand on the current approach.

Acknowledgement

This work was supported by grant N00014-22-1-2109 to Wayne D. Gray from the Office of Naval Research.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*(3), 288–318. https://doi.org/10.1037/0096-3445.117.3.288

Adams, R. J., & Ericsson, A. E. (1992, June). Introduction to Cognitive Processes of Expert Pilots [[Online; accessed 7. Aug. 2025]]. https://apps.dtic.mil/sti/html/tr/ADA254760

Arthur, W., Strong, M. H., Jordan, J. A., Williamson, J. E., Shebilske, W. L., & Regian, J. W. (1995). Visual attention: Individual differences in training and predicting complex task performance. *Acta Psychol.*, *88*(1), 3–23. https://doi.org/10.1016/0001-6918(94)E0055-K

Banerjee, S., & Gray, W. (2023). Role stability and team performance in a 4-player cooperative cooking game. *Proceedings of the 21st International Conference on Cognitive Modelling (ICCM 2023)*, 38–43. https://mathpsych.org/presentation/1226

Bavelier, D., & Green, C. S. (2019). Enhancing attentional control: Lessons from action video games. *Neuron*, *104*(1), 147–163. https://doi.org/10.1016/j.neuron.2019.09.031

Bavelier, D., Bediou, B., & Green, C. S. (2018). Expertise and generalization: Lessons from action video games. *Current Opinion in Behavioral Sciences*, *20*, 169–173. https://doi.org/10.1016/j.cobeha.2018.01.012

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin*, *144*(1), 77–110. https://doi.org/10.1037/bul0000130

Bishop, J., Burgess, J., Ramos, C., Driggs, J. B., Williams, T., Tossell, C. C., Phillips, E., Shaw, T. H., & Visser, E. J. d. (2020). Chaopt: A testbed for evaluating human-autonomy team collaboration using the video game overcooked!2. *2020 Systems and Information Engineering Design Symposium (SIEDS)*, 1–6. https://doi.org/10.1109/sieds49339.2020.9106686

Blacker, K. J., Curby, K. M., Klobusicky, E., & Chein, J. M. (2014). Effects of action video game training on visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1992–2004. https://doi.org/10.1037/a0037556

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Bravi, R., Gavazzi, G., Benedetti, V., Giovannelli, F., Grasso, S., Panconi, G., Viggiano, M. P., & Minciacchi, D. (2022). Effect of different sport environments on proactive and reactive motor inhibition: A study on open- and closed-skilled athletes via mouse-tracking procedure. *Front. Psychol.*, *13*, 1042705. https://doi.org/10.3389/fpsyg.2022.1042705

Burgoyne, A. P., Mashburn, C. A., Tsukahara, J. S., & Engle, R. W. (2022). Attention control and process overlap theory: Searching for cognitive processes underpinning the positive manifold. *Intelligence*, *91*, 101629. https://doi.org/10.1016/j.intell.2022.101629

Buschkuehl, M., & Jaeggi, S. (2010). Improving intelligence: A literature review. *Swiss Medical Weekly*. https://doi.org/10.4414/smw.2010.12852

Cao, L.-Z., He, H., Miao, X., & Chi, L. (2024). The contributions of executive functions to decision-making in sport. *International Journal of Sport and Exercise Psychology*. https://www.tandfonline.com/doi/full/10.1080/1612197X.2024.2371483#d1e362

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the raven progressive matrices test. *Psychological Review*, *97*(3), 404–431. https://doi.org/10.1037/0033-295X.97.3.404

Carroll, M., Shah, R., Ho, M. K., Griffiths, T. L., Seshia, S. A., Abbeel, P., & Dragan, A. (2019). On the utility of learning about humans for human-ai coordination. In *Proceedings of the 33rd international conference on neural information processing systems*. Curran Associates Inc.

Chamberland, C., Hodgetts, H. M., Kramer, C., Breton, E., Chiniara, G., & Tremblay, S. (2018). The critical nature of debriefing in high-fidelity simulation-based training for improving team communication in emergency resuscitation. *Appl. Cognit. Psychol.*, *32*(6), 727–738. https://doi.org/10.1002/acp.3450

Chen, J.-T., Kan, N.-W., Barquero, C., Teo, M. M. J., & Wang, C.-A. (2024). Saccade Latency and Metrics in the Interleaved Pro- and Anti-Saccade Task in Open Skill Sports Athletes. *Scand. J. Med. Sci. Sports*, *34*(8), e14713. https://doi.org/10.1111/sms.14713

Choi, E., Shin, S.-H., Ryu, J.-K., Jung, K.-I., Kim, S.-Y., & Park, M.-H. (2020). Commercial video games and cognitive functions: Video game genres and modulating factors of cognitive enhancement. *Behavioral and Brain Functions*, *16*(1). https://doi.org/10.1186/s12993-020-0165-z

Colzato, L. S., van den Wildenberg, W. P. M., Zmigrod, S., & Hommel, B. (2012). Action video gaming and cognitive control: Playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychological Research*, *77*(2), 234–239. https://doi.org/10.1007/s00426-012-0415-2

Cooke, N. J., Cohen, M. C., Fazio, W. C., Inderberg, L. H., Johnson, C. J., Lematta, G. J., Peel, M., & Teo, A. (2023). From teams to teamness: Future directions in the science of team cognition [PMID: 36946439]. *Human Factors*, 00187208231162449. https://doi.org/10.1177/00187208231162449

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain* [PhD thesis]. McGill University.

Dale, G., Kattner, F., Bavelier, D., & Green, C. S. (2020). Cognitive abilities of action video game and role-playing video game players: Data from a massive open online course. *Psychology of Popular Media*, *9*(3), 347–358. https://doi.org/10.1037/ppm0000237

De Cesarei, A., D'Ascenzo, S., Nicoletti, R., & Codispoti, M. (2023). Novelty and learning in cognitive control: Evidence from the simon task. *Psychological Research*, *87*(8), 2390–2406. https://doi.org/10.1007/s00426-023-01813-z

de Toledo, T. F. N., Benvenuti, M. F. L., Marques, N. S., & Glenn, S. S. (2022). Schedule performance as a baseline for the experimental analysis of coordinated behavior: Same or different units of analysis? *The Psychological Record*, *72*(2), 185–195. https://doi.org/10.1007/s40732-022-00510-4

Di Russo, F., Pitzalis, S., & Spinelli, D. (2003). Fixation stability and saccadic latency in élite shooters. *Vision Res.*, *43*(17), 1837–1845. https://doi.org/10.1016/S0042-6989(03)00299-2

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, *56*(293), 52–64. Retrieved September 5, 2024, from http://www.jstor.org/stable/2282330

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/bf03203267

Gessler, T., Dizdarevic, T., Calinescu, A., Ellis, B., Lupu, A., & Foerster, J. N. (2025). Overcookedv2: Rethinking overcooked for zero-shot coordination. https://arxiv.org/abs/2503.17821

Gonzalez, C., Thomas, R. P., & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence*, *33*(2), 169–186. https://doi.org/10.1016/j.intell.2004.10.002

Gray, W. D. (2017). Game-XP: Action Games as Experimental Paradigms for Cognitive Science. *Topics in Cognitive Science*, *9*(2), 289–307. https://doi.org/10.1111/tops.12260

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534–537. https://doi.org/10.1038/nature01647

Green, C., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, *22*(6), R197–R206. https://doi.org/10.1016/j.cub.2012.02.012

Haier, R. (2014). Increased intelligence is a myth (so far). *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00034

Hallett, P. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, *18*(10), 1279–1296. https://doi.org/10.1016/0042-6989(78)90218-3

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Appl. Cognit. Psychol.*, *24*(8), 1149–1167. https://doi.org/10.1002/acp.1624

Hubert-Wallander, B., Green, C. S., & Bavelier, D. (2011). Stretching the limits of visual attention: the case of action video games. *WIREs Cognit. Sci.*, *2*(2), 222–230. https://doi.org/10.1002/wcs.116

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*(3), 188–196. https://doi.org/10.1037/h0056940

Jiang, C., Kundu, A., & Claypool, M. (2020). Game player response times versus task dexterity and decision complexity. *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, 277–281. https://doi.org/10.1145/3383668.3419891

Kowal, M., Toth, A. J., Exton, C., & Campbell, M. J. (2018). Different cognitive abilities displayed by action video gamers and non-gamers. *Computers in Human Behavior*, *88*, 255–262. https://doi.org/10.1016/j.chb.2018.07.010

Large, A. M., Bediou, B., Cekic, S., Hart, Y., Bavelier, D., & Green, C. S. (2019). Cognitive and Behavioral Correlates of Achievement in a Complex Multi-Player Video Game. *Media and Communication*, *7*(4), 198–212. https://doi.org/10.17645/mac.v7i4.2314

Leong, A. Y. C., Yong, M. H., & Lin, M.-H. (2022). The effect of strategy game types on inhibition. *Psychological Research*, *86*(7), 2115–2127. https://doi.org/10.1007/s00426-021-01632-0

Lev-Ari, T., Beeri, H., & Gutfreund, Y. (2022). The Ecological View of Selective Attention. *Front. Integr. Neurosci.*, *16*, 856207. https://doi.org/10.3389/fnint.2022.856207

Li, H., Hu, L., Wei, L., He, H., & Zhang, X. (2023). Disentangling working memory from multiple-object tracking: Evidence from dual-task interferences. *Scand. J. Psychol.*, *64*(4), 437–452. https://doi.org/10.1111/sjop.12901

Liu, X., Han, X., Wang, T., & Ren, X. (2024). An account of the relationship between critical thinking and fluid intelligence in considering executive functions. *Thinking Skills and Creativity*, *52*, 101538. https://doi.org/10.1016/j.tsc.2024.101538

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Lukosch, H. K., Bekebrede, G., Kurapati, S., & Lukosch, S. G. (2018). A scientific foundation of simulation games for the analysis and design of complex systems [PMID: 30369775]. *Simulation & Gaming*, *49*(3), 279–314. https://doi.org/10.1177/1046878118768858

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

Marina G. Falleti, A. C., Paul Maruff, & Darby, D. G. (2006). Practice effects associated with the repeated assessment of cognitive function using the cogstate battery at 10-minute, one week and one month test-retest intervals [PMID: 16840238]. *Journal of Clinical and Experimental Neuropsychology*, *28*(7), 1095–1112. https://doi.org/10.1080/13803390500205718

Morgan, B. B., Salas, E., & Glickman, A. S. (1993). An analysis of team evolution and maturation. *The Journal of General Psychology*, *120*(3), 277–291. https://doi.org/10.1080/00221309.1993.9711148

Morrison, C. M. (2003). Interpret with caution: Multicollinearity in multiple regression of cognitive data. *Perceptual and Motor Skills*, *97*(1), 80–82. https://doi.org/10.2466/pms.2003.97.1.80

Munoz, D. P., & Everling, S. (2004). Look away: The antisaccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*(3), 218–228. https://doi.org/10.1038/nrn1345

Murphy, K. R. (1989). Is the relationship between cognitive ability and job performance stable over time? *Human Performance*, *2*(3), 183–200. https://doi.org/10.1207/s15327043hup0203_3

Nye, C. D., Ma, J., & Wee, S. (2022). Cognitive ability and job performance: Meta-analytic evidence for the validity of narrow cognitive abilities. *Journal of Business and Psychology*, *37*(6), 1119–1139. https://doi.org/10.1007/s10869-022-09796-1

Paul, R. K. (2008). Multicollinearity : Causes , effects and remedies. https : / / api . semanticscholar . org / CorpusID : 14367289

Peppler, K., Danish, J. A., & Phelps, D. (2013). Collaborative gaming: Teaching children about complex systems and collective behavior. *Simulation & Gaming*, *44*(5), 683–705. https://doi.org/10.1177/1046878113501462

Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A., & Alfieri, L. (2013). Effects of video-game play on information processing: A meta-analytic investigation. *Psychonomic Bulletin & Review*, *20*(6), 1055–1079. https://doi.org/10.3758/s13423-013-0418-z

Rahimi, A., Roberts, S. D., Baker, J. R., & Wojtowicz, M. (2022). Attention and executive control in varsity athletes engaging in strategic and static sports. *PLoS One*, *17*(4), e0266933. https://doi.org/10.1371/journal.pone.0266933

Reynaldo, C., Christian, R., Hosea, H., & Gunawan, A. A. S. (2021). Using video games to improve capabilities in decision making and cognitive skill: A literature review [5th International Conference on Computer Science and Computational Intelligence 2020]. *Procedia Computer Science*, *179*, 211–221. https://doi.org/10.1016/j.procs.2020.12.027

Schmidt, F. L. (2002, May). The Role of General Cognitive Ability and Job Performance: Why There Cannot Be a Debate. In *Role of General Mental Ability in industrial, Work, and Organizational Psychology* (pp. 187–210). Psychology Press. https://doi.org/10.4324/9781410608802-12

Schroeder, M. A., Lander, J., & Levine-Silverman, S. (1990). Diagnosing and dealing with multicollinearity [PMID: 2321373]. *Western Journal of Nursing Research*, *12*(2), 175–187. https://doi.org/10.1177/019394599001200204

Smith, D. L., & Claytor, R. P. (2018). An acute bout of aerobic exercise reduces movement time in a fitts' task (A. P. Lavender, Ed.). *PLOS ONE*, *13*(12), e0210195. https://doi.org/10.1371/journal.pone.0210195

Stegmueller, D. (2013). How many countries for multilevel modeling? a comparison of frequentist and bayesian approaches. *American Journal of Political Science*, *57*(3), 748–761. https://doi.org/10.1111/ajps.12001

Temprado, J.-J., Sleimen-Malkoun, R., Lemaire, P., Rey-Robert, B., Retornaz, F., & Berton, E. (2013). Aging of sensorimotor processes: A systematic study in fitts' task. *Experimental Brain Research*, *228*(1), 105–116. https://doi.org/10.1007/s00221-013-3542-0

Toril, P., Reales, J. M., Mayas, J., & Ballesteros, S. (2016). Video game training enhances visuospatial working memory and episodic memory in older adults. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00206

Trick, L. M., Mutreja, R., & Hunt, K. (2012). Spatial and visuospatial working memory tests predict performance in classic multiple-object tracking in young adults, but nonspatial measures of the executive do not. *Atten. Percept. Psychophys.*, *74*(2), 300–311. https://doi.org/10.3758/s13414-011-0235-2

Tucker, P., & Warr, P. (1996). Intelligence, elementary cognitive components, and cognitive styles as predictors of complex task performance. *Personality and Individual Differences*, *21*(1), 91–102. https://doi.org/10.1016/0191-8869(96)00032-3

Valls-Serrano, C., De Francisco, C., Vélez-Coto, M., & Caracuel, A. (2022). Visuospatial working memory and attention control make the difference between experts, regulars and non-players of the videogame league of legends. *Frontiers in Human Neuroscience*, *16*. https://doi.org/10.3389/fnhum.2022.933331

Voss, M. W., Kramer, A. F., Basak, C., Prakash, R. S., & Roberts, B. (2010). Are expert athletes 'expert' in the cognitive laboratory? A meta-analytic review of cognition and sport expertise. *Appl. Cognit. Psychol.*, *24*(6), 812–826. https://doi.org/10.1002/acp.1588

Weise, J. J., Greiff, S., & Sparfeldt, J. R. (2020). The moderating effect of prior knowledge on the relationship between intelligence and complex problem solving – testing the elshout-raaheim hypothesis. *Intelligence*, *83*, 101502. https://doi.org/10.1016/j.intell.2020.101502

Yang, L., Zhang, Y., Zhang, Y., Liao, Y., Du, J., & Geng, X. (2021, September). Research on Training Method of Information Processing Ability of Military Pilots. In *Man-Machine-Environment System Engineering: Proceedings of the 21st International Conference on MMESE* (pp. 753–759). Springer. https://doi.org/10.1007/978-981-16-5963-8_102

York, J. L., & Biederman, I. (1990). Effects of age and sex on reciprocal tapping performance. *Perceptual and Motor Skills*, *71*(2), 675–684. https://doi.org/10.2466/pms.1990.71.2.675

Zorowitz, S., & Niv, Y. (2023). Improving the reliability of cognitive task measures: A narrative review. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *8*(8), 789–797. https://doi.org/10.1016/j.bpsc.2023.02.004