In Vivo Studies of Solo and Team Performance

Ray Perez (Co-Organizer)
Office of Naval Research (ray.perez@navy.mil)

Wayne D. Gray (Co-Organizer)
Rensselaer Polytechnic Institute (grayw@rpi.edu)

Jerad H. Moxley
Weill Cornell Medicine (jhm2006@med.cornell.edu)

David Mendonça
Rensselaer Polytechnic Institute (mendod@rpi.edu)

Jamie C. Gorman
Georgia Institute of Technology (jamie.gorman@psych.gatech.edu)

We bring together four researchers who study expertise in team or in solo (i.e., individual) performance. Team research tends to either collect a lot of questionnaire data after performance or a little data, in real-time, by human observers. Studies of solo performers are often restricted to convenience samples of task novices, who often spend less than an hour learning and performing the task. In contrast, the research of all four of our panelists is notable for using tasks which require days-to-years of practice and for the quantities of data collected. Discussions will emphasize the contributions these approaches are making to theoretical cognitive science.

Jamie C. Gorman – *Theory of Interactive Team Cognition*

By recreating environments for Drone pilots or Submariners, Jamie Gorman and colleagues collect communications and responses among team members in real-time longitudinal studies. These data allowed researchers to apply the power of nonlinear dynamical systems theory to further develop the theory of Interactive Team Cognition (ITC, Cooke, Gorman, Myers, & Duran, 2013). The approach has been extended to teams composed of humans and machines.

ITC proposes that team cognition: (1) is an activity, not a property or product; (2) must be measured and studied at the team level; and (3) is inextricably tied to context.

ITC Prop 1 maintains that team cognition is dynamic and context dependent. ITC Prop 2 leads to a systems perspective in which models and metrics are focused at the team level, with individual cognition and behavior viewed as emergent team dynamics. Team member behavior and cognition are dynamically reorganized (or rearranged) in real time (ITC Prop 1) to maintain functionality as the team adapts to changing task environments to achieve its goal. Hence, teams with high cognitive skill achieve their goal even if environmental context varies and roadblocks to team effectiveness are encountered (ITC Prop 3).

Unlike individual cognition, there are no standard tests to measure the general cognitive skill or ability of a team. One theoretical and methodological development has been to determine a generalizable way to identify and measure team cognitive skill through a team's "general adaptive response".

Our research on team cognition has shown that teams that achieve their goals have (a) a faster general adaptive response, (b) adapt their responses to the variability in obstacles they encounter, and (c) generate responses appropriate to the particular roadblocks they encounter. For examples, I will draw on research with medical teams, submarine crews, UAV teams, as well as in vitro, laboratory, team coordination tasks. This variety of teams illustrates the concept of the *general adaptive response* as an ITC-based measure of team cognitive skill. These teams also illustrate the real-time dynamical system modeling techniques that we use to track team cognition in dynamic environments.

David Mendonça – *Adaptation in Adversarial Games*

David Mendonça's prior research has focused on intensive studies of teams in high-stakes, time-constrained environments. His most recent work is an extremely retrospective analysis of "An historical perspective on community resilience: The case of the 1755 Lisbon Earthquake" (Mendonça, Amorim, & Kagohara, 2018).

In addition to being the most played game in the world (with approximately 10M active users), *League of Legends* (LoL) is an adversarial game (similar to "capture the flag") in which teams must adapt to (and even precipitate) unplanned-for contingencies. Elite players (such as those we study), have played thousands of such matches, with the average match consisting of two teams, each of 5 players, battling for 30 min.

Our work explores the relationship between (i) prematch composition of a team, (ii) decision processes within the match, and (iii) match outcomes in LoL. Respective methodological challenges include (i) characterizing team capabilities, (ii) quantifying adaptation, and (ii) validating measures of performance.

In contrast to traditional work on teams, we utilize no psychometric instruments, instead deriving measures that are validated against salient theoretical constructs and instantiated with gameplay data. And while these data are freely available, their allure is offset by some hard realities: researchers have no influence over either the data stream or the game architecture, and the formulas used to benchmark individual and team expertise are held as trade secrets. Matches are scheduled by the developers on a rolling basis and–unlike in "regular" sports–are designed so that opponents are closely matched.

After briefly summarizing results to date, we explore within-match performance of teams whose members have weaker or stronger histories of working together, focusing specifically on behavioral responses to the temporary loss of one or more team members. We present data on how the experience of "playing shorthanded" translates (or fails to translate) into longer-term behavioral adaptations.

The talk concludes with issues and implications for the design and/or modification of open-source, team-based games and the data associated with them. ïż£

Jerad Moxley – *Chess: The Once & Future Paradigm*

The distinction of having studied more types of game-play by solos or teams, than any other researcher on this panel may go to Jerad Moxley. His studies have spanned crossword puzzles, chess, basketball, elderly game players, videogames, as well as gender differences among SCRABBLE players.

For researchers interested in skilled performance, an important feature of chess is the reliability of the chess rating system and the fact that one experimental task (the choose the *best move task*), can measure skill and age effects about as well as tournament play, thereby making Chess ideal for studying domain-specific performance. Complimentary, another common task, the *recall task*, diverges from tournaments in ways that make it useful for studying a mixture of domain specific and domain-general abilities.

Applying the *best move task* and *the recall task* across the lifespan of chess players has increased our understanding of how domain-specific processes and domain-general abilities develop. Research on older adults and children now converges to show strong aging effects of chess tasks that tap into both specific and general abilities. In contrast, the best move tasks captures relatively small aging effects consistent with tournament performance.

As noted, performance on the best move task shows developmental trends in both youth and older adults that mirror tournament performance. Importantly, however, process tracing shows clear differences between the growth of skill in youth and the decline of skill with aging. Although skill development is broadly consistent with what we expect based on tournament performance, the age-related decline of prior skill levels shows process differences that dissociate from skill. In particularly, the age-related declines are not uniform. On easy problems, better players immediately gain an advantage over weaker players.

In contrast, on difficult problems, process tracing has shown that better players initially resemble weaker players but as problem solving continues, better players massively improve their move selection. In contrast, more time does not improve the performance of the weaker players. Methodologically, these conclusions follow from the combination of verbal protocol analysis and the traditional behavioral measures.

We view chess not as a standalone domain, divorced from the rest of human cognition but, rather, as a viable paradigm for studying the big questions in cognitive science. Indeed, the tasks and domains discussed here can easily be used by researchers who have no interest in chess itself to answer their questions of interest.

Wayne D. Gray – *Plateaus, Dips, & Leaps to Expertise*

After several years of working in applied labs, Wayne Gray became concerned that basic researchers were not working on the types of theory he needed to do his job. That concern led him to shift to academe where he has since attempted to pursue theories and research applicable to problems of interactive behavior.

Learning a new task can be hard but, apparently, learning and using a new procedure for an old task can be even harder. That is the message from work on *stable suboptimal performance* from the early 2000s. Wai-tat Fu and I demonstrated time and again that people who knew the optimal procedures would fail to apply them, falling back on older ways of doing things.

Although that battle is still being fought (e.g., Lafreniere, Gutwin, & Cockburn, 2017), the focus in my lab has shifted. After a few years of looking at learning curves for individuals, we realized that none of our curves were close to being picture-perfect power law curves. All of our curves showed plateaus, dips, and leaps. Indeed, what we had thought of as noise was, in fact, the message; namely, that learning a real-time, complex, dynamic task entails a series of explorations and discoveries, trials and errors, in search of methods or strategies that will move performance forward.

We now refer to complete mastery of a task as *asymptotic performance* and to stable suboptimal behavior as performance *plateaus*. However, the most interesting parts of the curve are those periods in which performance dips and, sometimes, leaps. The talk will provide several examples of the use of dips and leaps to identify periods of method discovery or invention.

Ray Perez – Basic Research for Complex Problems

For the last 3 decades, Ray Perez has been pursuing applied problems by finding or encouraging others to find theory-based solutions. Most recently, Ray has been the Program Officer of the Office of Naval Research's Cognitive Science of Learning program.

Ray Perez is co-organizer of this symposium as well as its moderator and discussant. In each of these three roles, Ray is focused on how complex tasks, sometime performed by a single person and other times performed by teams, are learned and executed.