

Strategy variability: How too much of a good thing can hurt performance

JEFFREY T. HANSBERGER

George Mason University, Fairfax, Virginia

CHRISTIAN D. SCHUNN

University of Pittsburgh, Pittsburgh, Pennsylvania

and

ROBERT W. HOLT

George Mason University, Fairfax, Virginia

Recent research (e.g., Siegler, 1996) has discovered the important and vital role that variability plays among strategy use and development over time. However, as many researchers have pointed out (e.g., Miller, 1993), the majority of the research addressing this issue has focused on the outcomes, rather than on the potentially more informative aspects of variability, strategy development, and the process of adaptation. In this study, we examined the role of variability during strategy development, utilizing a longitudinal method. Thirteen participants were studied over 3 months as they coached a simulated football team. The results suggest that variability plays a major role in adaptive skill acquisition in a dynamic environment—in the direction, however, opposite to that predicted by previous research done with simple static tasks.

Strategies are pervasive in our everyday lives as organized ways to effectively interact with the environment and accomplish goals. A critical element in the development of these strategies is the variability that exists in human cognition and behavior. Past research in the strategy development domain, particularly in developmental psychology, has focused on the change in strategies across age groups. The typical way of describing ways of thinking or strategy use has been as a progressive sequence of increasingly proficient strategies, as knowledge or understanding increases (e.g., Ashcroft, 1987; Piaget, 1952). Over time, the result of this perspective resembles a stepwise progression of strategies. This idea of developmental stages according to age or experience with a task, however, has had inconsistent support (e.g., Bryant & Trabasso, 1971; Markman & Seibert, 1976). Recent research has proposed a contrasting perspective on strategy development and use that focuses on the influential role of strategy variability.

Individuals use multiple strategies for a given task, and the frequency of any particular strategy use varies as they

obtain experience (e.g., Reder, 1982, 1987, 1988; Siegler, 1996; Siegler & Robinson, 1982). Even with a seemingly simple task, such as single-digit addition, multiple strategies are used among college-aged adults (LeFevre, Sadesky, & Bisanz, 1996). Siegler (1996) has referred to this variability among strategies as the *wave theory*, since each strategy, over time, produces a wave-like shape, which represents the decreased use or extinction of some strategies and the successful increase or proliferation of others. Recently, evidence supporting the wave theory has been accumulating (e.g., Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; McClelland & Siegler, 2001; Siegler, 1996, 1998) in areas ranging from mathematics (e.g., Siegler & Robinson, 1982) to scientific reasoning (Kuhn & Phelps, 1982). However, most of this research has been done primarily with children interacting with relatively simple and static tasks.

Variability is thought to be instrumental in the development of strategies over time and a fundamental aspect of behavior and thought (e.g., Siegler, 1996). Past research has shown that a high degree of variability among strategies at the onset of task experience leads to greater learning and performance (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Siegler, 1995). To explain why this might be the case, Siegler (1996) presented an analogy from evolutionary biology. Change in biological systems depends on adequate variability to allow natural selection to occur effectively. This variability in biology is caused by a number of events, such as mutations at the genetic level or through the process of recombination of parents' genes. The biological variance can produce adaptive advantages to an organism. Similarly,

J.T.H. is now at the Human Research and Engineering Directorate, Army Research Laboratory. We acknowledge the feedback, advice, and involvement of Leonard Adelman and Deborah Boehm-Davis. We also thank Jamie Estock and Mary Pinney, who helped in data collection. Correspondence concerning this article should be addressed to J. T. Hansberger, Army Research Laboratory, US JFCOM J9 JCD, 115 Lakeview Parkway, Suffolk, VA 23435 (e-mail: jhansberger@arl.army.mil).

Note—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

strategy variation could produce adaptive advantages at the cognitive level.

Expanding this analogy, one would expect that variability among strategies results in the exploration of different behaviors in order to achieve an end state or goal. Among various explored strategies, some may prove more effective than existing strategies or provide useful insight to the user. This exploration provides the opportunity to learn and provides the potential for greater flexibility in the future. The exposure to more strategies, due to greater variability, equips the individual with a larger repertoire of strategies. When faced with a novel situation or set of circumstances, an individual with a larger and more diverse repertoire of strategies should be more adaptable and able to cope with the new situation than one who has experienced less variability among strategies (Schmidt & Lee, 1999).

To unpack how variability influences performance, it is important to examine the time course of strategy variability and performance. In the past, there have been only a few attempts to do this kind of examination with any level of detail. One research study by Siegler and Taraban (1986) investigated child interaction with the simple task of solving balance scale problems (e.g., an end state is visually presented and children are asked to duplicate it with the balance scale). On the basis of these results, Siegler formulated a hypothesis of strategy variability aggregated over time with experience: an inverted-U relationship between task experience and strategy variability. This relationship has been called the *moderate experience hypothesis* by Siegler (1996), which refers to the increased amount of diversity among strategies with moderate levels of experience. Given that people generally improve performance with a task over time and experience (e.g., Anderson, 1981; Ericsson & Lehmann, 1996; Fisk, Ackerman, & Schneider, 1987; Welford, 1968), *novice* individuals should display a general positive relationship between strategy variability and performance over time. However, the evidence for the moderate experience hypothesis is limited to one study (Siegler & Taraban, 1986), and so it is an open question as to whether there is an inverted-U relationship between strategy diversity and time in all or most domains.

Besides focusing more on outcomes than on processes, past research in strategy development has also typically focused on strategy change within static and unchanging situations, such as arithmetic (e.g., Siegler & Robinson, 1982), simulated elements of scientific reasoning (e.g., Schauble, 1990), and oral and written language (e.g., Ferguson, 1986; Jakobson, 1941/1968). These domains have served the research community well, since they provide problems that are well defined, simple, and unchanging. These very elements that are advantageous to experimental settings, however, do not generalize to many task settings and situations that we find ourselves interacting with on a day-to-day basis. Many everyday tasks have some, if not all, of the following naturalistic qualities: (1) ill-structured problems, (2) uncertain dynamic environments, (3) shifting or ill-defined goals and subgoals, (4) action-feedback loops, (5) time stress, and (6) medium-to-high stakes (Orasanu & Connolly, 1995).

A fundamental difference of dynamic environments is that they impose on a person the need to *adapt* his or her strategies to the unexpected changes that he or she encounters. To borrow from evolutionary biology again, adaptation occurs when a source of variability and a source of selection interact with one another. These two elements acting together provide the capability of evolving behavior to improve or maintain high task performance. Both strategy variability and the selection among that variability play distinct roles in adaptive skill acquisition over time, but it is their interaction that produces adaptive behavior.

Variability among strategies has been shown to exist and play an important positive role in performance (Alibali & Goldin-Meadow, 1993; Siegler, 1995). However, research in evolutionary biology (e.g., Lewontin, 1978) strongly suggests that mechanisms of *selection* are also important to a creature's ability to adapt. High variability among strategies provides the basis to respond to a greater range of circumstances, but strategies must be applied in an appropriate manner to maintain or improve performance. Siegler (1996) states, "Adaptive selection among the varying entities [strategies] is important because it allows organisms to perform effectively at any given time, through using effective approaches disproportionately often" (p. 17). Therefore, it is the combination or interaction of strategy variability with appropriate selection criteria that should provide optimal performance. Given a dynamic task with feedback that allows for adaptive behavior over time, it is hypothesized that strategy variability, moderated by the appropriate use of strategies, will lead to higher levels of performance or task proficiency over time.

A strategy is defined here as behavior systematically applied to a domain task/activity that serves or appears to serve an important function in the completion of a goal. Given the dynamic and complex environment we interact with on a daily basis, there are typically many different strategies available that facilitate goal accomplishment. One constant goal that we are all faced with, particularly with new and dynamic tasks, is learning.

Crowley, Shrager, and Siegler (1997) have defined two basic types of learning: metacognitive and associative learning. Metacognitive learning is explicit, effortful, and flexible. In contrast, associative learning is implicit, fast, and responsive to feedback from the environment. Crowley et al. suggested that both play a critical role in strategy development over time.

Metacognitive learning implies that a person generates new strategies by reflecting on information acquired from past experience or through external information sources. This information can be gathered, processed, and applied at two distinct conceptual levels: the strategic and the tactical levels (e.g., Pennington & Grabowski, 1990). The strategic level focuses on upper level information and knowledge that is more conducive for general planning and conceptual understanding. The tactical level, on the other hand, is focused on lower level information or the detailed knowledge required for task completion. The use of specific information-gathering strategies at both levels should aid in learning (Anderson, 1981) and was expected

to positively influence strategy variability and development in this study.

There may also be some associative learning affiliated with information gathering, but its role would be minor, in comparison with metacognitive learning. Associative learning is better applied to existing knowledge and strategies, where repeated exposure to the situation gradually refines behaviors and strategies. A direct form of experience in the form of practice would elicit associative learning more strongly than information gathering.

Associative learning of strategies adheres to the *matching law* by matching the frequency of strategy use to the reinforcement obtained from that strategy (Herrnstein, 1961). Associative learning requires extensive experience in the targeted domain in order to obtain high levels of proficiency. It basically aids in refining the strategies that an individual already possesses. This experience is often in the form of repetitive practice for many domains and tasks (Ericsson, Krampe, & Tesch-Römer, 1993; Simon & Chase, 1973). This type of practice should gradually decrease variability among strategies as the strategies are slowly pruned down to more efficient ones according to their success rates (Lovett & Schunn, 1999; Reder, 1987).

These various approaches to learning are strategies themselves. The time spent using these learning strategies, along with their variability, should have a positive effect similar to that found in past research on learning and performance. It could also influence the type and variability of task-specific strategies used for performance. As has been mentioned before, it is also hypothesized that the information-gathering learning strategy should have a stronger positive influence at the strategic level than at the tactical information level on task-specific strategy variability. This strategy variability should, in turn, interact with the individual's selection criteria to provide more adaptive and improved performance over time.

In this study, we attempted to extend the moderate experience hypothesis beyond a static task setting to a dynamic and complex task environment. The interaction with a setting possessing more naturalistic characteristics should require individuals to show typical adaptive behaviors with their strategy selection and use over time. Most important, these changes were investigated over time in order to provide additional clues as to why and how strategy variability translates into improved learning and performance.

METHOD

Microgenetic Method

To obtain more information about what is going on during periods of rapid change, Siegler and Crowley (1991) proposed the use of the microgenetic method. This method has three primary characteristics: "a) observations span the period of rapid change in the competence of interest; b) the density of observations is high relative to the rate of change in the competence; c) observations are subjected to intensive trial-by-trial analysis, with the goal of inferring the processes that gave rise to the change" (Siegler, 1996, p. 178).

With this type of method, the focus of the analysis is at the individual level, and not on averaging across participants, which can misrepresent what the individual is doing (Estes, 1956; Sidman, 1952). Quantitative results are complemented by qualitative results

and findings to produce a better understanding of individual processes and outcomes. Therefore, in this study, we examined how strategy change occurred for 13 participants over 16 experimental sessions and a period of 2–3 months.

Task

In order to effectively study the relationship of goals to the mechanisms of variable thinking, adaptive selection of strategies, and performance, the task domain must possess several important characteristics that are evident in a natural environment. The task must possess sufficient interest and importance to elicit the types of goals and effort participants would normally exhibit in everyday life. The task should also be dynamic in order to better evaluate adaptive strategy shifts to a changing environment. Finally, the task should be complex enough that different strategies can be learned and applied to different task situations.

The task domain of simulated football coaching is interesting to many college students, dynamic across games with different opponents, and complex enough to allow several distinct strategies but still provides enough control to quantitatively analyze the results. The coaching simulation task used in this study was meaningful and engaging to the participants, which raised the stakes for them and provided more natural goals and behavior.

The football simulation allowed each participant to coach a single team with the same players over a span of 16 regular season games. Each game was conducted in the same manner and order as a real-life game. At the beginning of each play, the participants had access to a virtual playbook for offense and defense, depending on whether they had control of the ball or not. The participant had 25 sec between each play to select and begin the play. The play selected was then executed by the computer once it had begun, which relieved the participant of any direct control over the execution of the play. There were over 200 specific offensive plays and 100 individual defensive plays available.

Documentation from Electronic Arts and the Prima Strategy Guide to Madden 2001 (Cohen, 2001) suggest that the basic game mechanics of Madden 2001 are primarily guided by the player personnel. Each team is composed of different players with varying levels of proficiency at their designated positions. The more effective plays for a team are the ones that utilize the strengths of their personnel. For example, the participant's team, the Miami Dolphins, has a strong offensive line, a good fullback, and a good running back. Therefore, this team is well suited for running plays. There is no one play that is consistently dominant across all teams and situations, and success cannot be obtained by randomly selecting plays.

Effective play calling is also affected by the strengths and weaknesses of the opposing team, which changes from game to game. For example, the first opponent of the season, the New York Jets, has a good defense with a solid defensive line and a good quarterback but poor wide receivers and a weak offensive line. The solid defense will make it hard for the participant's offense to run the ball. The computer-controlled opponent will also act according to the situation (e.g., what down it is and how many yards to go) and will have its own tendencies.

It is important to note that although it is a video game, the relative importance of sensory-motor skills in this particular game is relatively low. Unlike many sports video games that require rapid perception and action, the focus on coaching and the way in which Madden Football is structured eliminates the need for fast reflexes. It is more comparable to chess than to Pac-Man. The dynamic nature of the game is in the continually changing situation, from play to play and from opponent to opponent.

Participants

Thirteen university members participated in this study: 9 undergraduate students, 2 graduate students, and 2 Ph.D. scientists (providing some diversity in education level and age). There were 11 males and 2 females; the mean age of the participants was 23 years.

All the participants had little to no experience with the football coaching simulation, but all reported that they had a basic understanding and interest in football. The participants were compensated \$80 to \$150 on the basis of their performance: \$3 for each win, \$13 for best league record, \$10 for best division record, and \$5 for winning "coach of the year."

The participants were screened to have no official football coaching experience or training, and only 1 participant had experience playing organized football beyond the high school level, which was purely recreational football. The participants had little to no experience with the specific football simulation used and little experience with football simulations in general. The participants also spent relatively little time playing sports and racing video games: an average of 1.7 h per week, as compared with respondents to a recent survey of college-aged adults, who spent an average of 8.9 h per week playing video games (Sherry, Lucas, Rechtsteiner, Brooks, & Wilson, 2001).

Materials

The football simulation was Electronic Arts' Madden 2001 and was played using the Sony PlayStation 2 game console with a Sony control pad. The participants were provided with their own copy of the playbook, which contained diagrams of all available offensive and defensive plays, and blank notebook paper for taking notes. The participants also had access to a small collection of football coaching textbooks (American Football Coaches Association, 1995, 2000; Billick, 2001; Cohen, 2001; Koehler, 1992; Long & Czarnecki, 1998) that could be studied during the pregame period of each experimental session. Microsoft Access forms were used by the experimenter to record all in-game play calling during the session.

Procedure

To reduce participant bias and demand characteristics regarding strategy variability, a cover story was provided to each participant. The cover story explained that the purpose of the study was to investigate usability issues concerning the software and hardware design of the football simulation and the Sony PlayStation 2. Questions during the postgame interview focused on usability issues and potential problems with the simulation and the console/controller. This cover story did not appear to hinder or alter typical playing behavior or attitudes toward the game or any of the sessions. At no point were comments made by any of the participants during the game on usability issues; debriefing suggested that this lack of comment on usability reflected engagement with the game, rather than doubts about the cover story.

All the participants were assigned to coach the same fixed team/players over the course of a regular 16-game season. This setup is the typical game format and provides one continuous element across the season: Team and players do not change. All survey-based measures were administered in the first session of the study (e.g., demographic information). The second session consisted of an introduction to the task, practicing navigating through the interface, and coaching a simulated game. The participants continued practice sessions until they became proficient with the controls and passed a general skills proficiency test.

Following the practice sessions, each game was coached in a 1- to 2-h experimental session (two to three sessions/week). All 16 regular season, experimental sessions consisted of the following phases: (1) pregame preparation, (2) coaching 1 of the 16 games, and (3) postgame activities included answering cover story questions verbally and making any other final comments.

During the pregame preparation phase, the participants had six different pregame strategy activities available to help them prepare for the game. Each activity mapped to one of several hypothesized variability-producing mechanisms, which we called *learning strategies* and will be described below. The participants could spend from 0 to 40 min doing any of these activities for each session. These activities had the potential to improve and diversify the strategies that

the participants had and could use during the game. The participants had no direct control over the players during the game and could only call the plays, call audibles, make roster adjustments, and call time-outs (basic duties of an actual sideline coach).

Measures

Offense versus defense. Within this football simulation, the participants were equally responsible for calling plays and making decisions for the offense and the defense. Calling plays for offense and defense required different knowledge and strategies. One of the basic differences is that in coaching a defense, one is often more *reactive* to the actions of the opponent, as opposed to being on the offense, where play calling is strongly dictated by the offensive strategy. Most individuals also favor and have more interest in the offensive side of play calling. The participants in this study showed this bias, since they spent, overall, 85% of their pregame time on offense-oriented activities. Given these differences between the offense and the defense and the fact that greater attention was devoted to the offense by the participants, offensive play calling was chosen as the focus of this study.

Learning strategies. The specific learning strategies responsible were measured by the amount of time the participants spent on various offense-related tasks available to them during the pregame preparation phase (see Table 2). Following the theoretical analysis of strategy change mechanisms, we divided learning strategies into metacognitive and associative learning strategies. The information-gathering strategy responsible for metacognitive learning was divided into two different strategies that differentiated the type of information being gathered. The first strategy is called *strategic information gathering* and was directed and focused on higher level information, such as understanding the pros and cons of using various offensive formations (e.g., the *I formation*, the *single-back formation*, etc.). The second metacognitive learning strategy is called *tactical information gathering* and was directed toward lower level information, such as the details of a task. This type of information is focused on specific instructions, details, and procedural knowledge, such as the past performance of specific players or how fast or strong position players might be.

The strategic information-gathering strategy was defined by the time spent examining the playbook and the reference books that were available. The tactical information-gathering strategy was defined by the time spent reviewing their own team or player statistics, their opponent's team or player statistics, and the examination of rosters and player characteristics for their own team or an opponent's team. Lastly, the practice strategy responsible for associative learning was defined by the amount of time the participants spent utilizing the practice field with their offense and the time spent playing practice games prior to their actual game. Note that metacognitive learning can occur through practice but that, relatively speaking, there is much more associative learning in practice.

Learning strategy shift. Variability between sessions among the mechanisms was measured by comparing each session with the session that preceded it, beginning with Session 2. When a specific activity (e.g., playbook examination) was used in a given session without being used in the preceding session, it was considered a strategy shift in learning strategy.

Coaching strategies. The coaching strategies focused on the participants' offensive-play-calling selection during the course of each simulated game. Each individual offensive play was categorized by its basic type (run or pass), direction for run plays, and distance for pass plays. Run plays were categorized into five categories: (1) run inside left, (2) run inside middle, (3) run inside right, (4) run outside left, and (5) run outside right. Pass plays were categorized into three categories: (1) pass short, (2) pass medium, and (3) pass long. This categorization scheme is often used by football analysts because it easily includes all the plays and allows the identification of play-calling trends. However, complex tasks, such as football coaching and play calling, typically have many layers of strategies. This cat-

egorization scheme was used to focus an already complex analysis on the salient level of offensive strategy choice.

The distribution of proportions that each type of play was called during a game summarizes the offensive strategy/strategies employed for that game. The strategy for a game consists of the profile of plays, rather than the calling of any single play type. In other words, it is the accumulation of calling various types of plays over the course of the game that defines what the strategies were for any particular game. For example, Figure 1 displays how one participant's offensive strategies (proportion of play calling) shifted from game to game. This person changed their play-calling mix from a focus on the long pass and running to the inside left in Session 6 to a more focused strategy for medium and short passes in Session 7. Due to the very dynamic nature of these strategies for each game, the analysis will examine both the variability found within a single game and the variability or shifting of play calling across games.

Between-game coaching strategy shift. Variability between games was recorded as the change or shifting of distributions of each play-calling type from Game *n* to Game *n*+1. These distribution differences between any two successive games were summed across the eight play types to provide the strategy shift score. This measure shows the degree of change occurring among the game strategies from one game to the next and is conceptually distinct from the variability *within* a game. A low score indicates stability and little change from one game to the next, whereas a higher score indicates a greater amount of shifting and variability from game to game.

Within-game coaching strategy variability. Because each game consisted of a distribution of various play types, strategy variability within a game can be measured by the distribution of play calling for each game. Strategy variability within a single game was measured using the uncertainty or entropy coefficient (Theil, 1972). The uncertainty coefficient is computed using the equation $\{-\sum P_i[\text{LN}(P_i)]\}$, where P_i is the proportion each play was called during the game. Maximum variability results when each play-calling category is called a proportionally equal amount. With eight

play-calling categories, an equal distribution for each play-calling type would be 12.5% and would provide the maximum uncertainty score of 2.08. The lowest amount of within-game strategy variability would occur if only one play type was called 100% of the time for a game, which would produce an uncertainty score of 0. In other words, the lower the uncertainty score is, the more the individual is favoring one or a few particular play types for that game.

A high uncertainty score indicates that the participants are not heavily favoring any one play type and that there is more variability in their play calling for that game. Both the learning strategies and performance are measured in the same session as within-game strategy variability. Therefore, it is hypothesized that the within-game strategy variability will be more strongly related to antecedent learning strategies and that learning strategies, in turn, will more strongly predict performance than will between-game shifts. *Strategy variability* will be used to describe the within-game strategy variability, and the between-game variability will be referred to as *strategy shift*, as described above.

Strategy (matching) selection criteria. The appropriateness or the fitness of strategy selection for each game was measured by considering the degree to which each participant optimized their play calling with respect to the strengths and weaknesses of their team and their current opponent's team. Each team, including the participant's own team, is composed of individual players who possess performance ratings provided by the software. The performance of the simulated players is reflected in these performance ratings. Therefore, if a player has a performance rating of 95 on a 100-point scale, they will be very effective at their job assignment. However, not all players on the field have a direct and consistent influence on the success of the chosen play. On the basis of input from three subject matter experts, the players most influential to the success or failure for each type of play (e.g., run inside left or pass short) on offense and defense were identified. The performance rating for these players was averaged to obtain an overall score for that play type on offense and defense. This score quantitatively helped to indicate

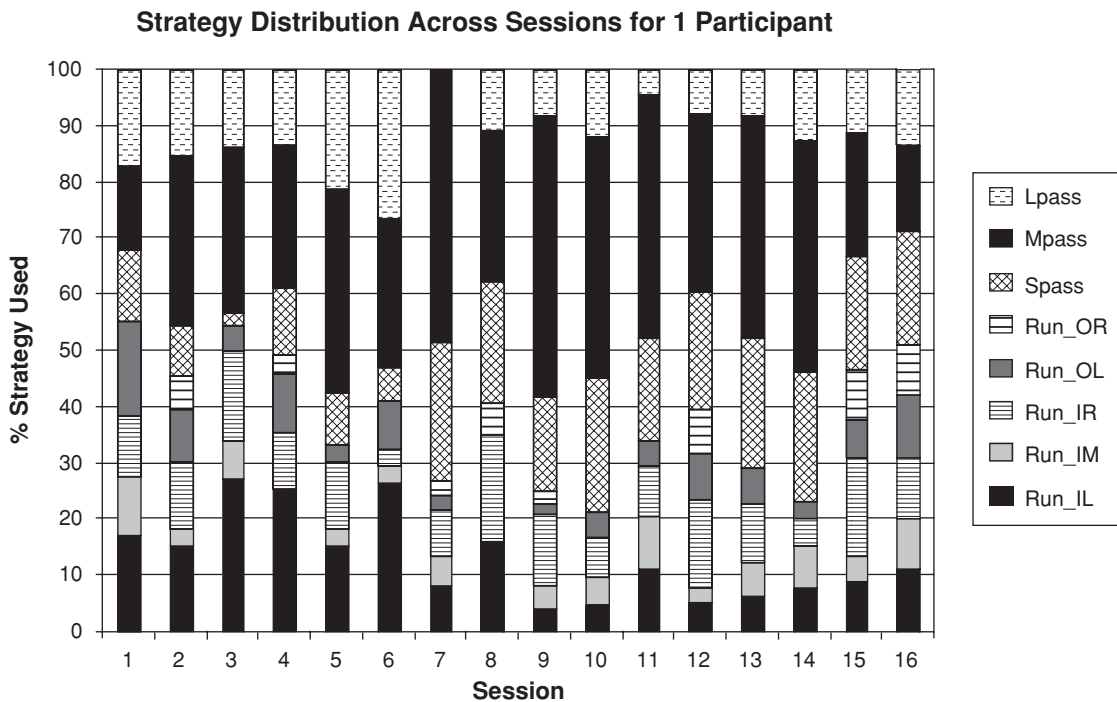


Figure 1. Distribution of the type of plays called across sessions for 1 participant. L pass, pass long; M pass, pass medium; S pass, pass short; OR outside right; OL, outside left; IR, inside right; IM, inside middle; IL, inside left.

areas of strengths and weaknesses in the participant's team and the opposing teams.

The most adaptive play selection on offense occurs when the participant's own team has a high offensive score for a play type (signifying a strength) and the opponent has a low defensive score for the same play type (signifying a weakness). For example, if the participant's team had a score of 93 out of 100 for *run inside left* plays and their current opponent had a score of 63 out of 100 for defending against *run inside left* plays, the resulting difference or *matching* score would be +30. A positive matching score reflects a favorable match-up for the participant, and a negative matching score reflects a more favorable match-up for the opponent. To the extent to which the participant selected types of plays that took advantage of their team's strengths and opponent's weaknesses, they would obtain a large positive matching score. If participants selected plays for which their team's competency was weak and the opponent's was strong, they would obtain a low to negative matching score.

To provide an estimate of how often and to what extent participants selected appropriate plays, the matching scores were weighted by the play-calling distributions for each game. The weighted scores were summed across play types to provide an overall selection criteria score for each game and each participant.

To help understand the statistical properties of the task domain in order to aid future generalization of results, several strategy-to-performance analyses were conducted. First, an analysis examining which play-calling strategies predicted performance validated the a priori matching selection criteria used. This analysis used observed choices and performance outcomes in the present data to see which strategies were, in fact, most successful across participants. The matching selection criteria identified the run inside left play as the best inside run play according to the personnel on the participants' assigned team, the run outside left play as the best outside run play and both the short and medium passes as the best pass plays. A stepwise regression of the percentage of usage of each play-calling strategy predicting performance also identified the run inside left ($\Delta R^2 = .02, p < .05$), run outside left ($\Delta R^2 = .03, p < .05$), and the short pass ($\Delta R^2 = .05, p < .01$) plays as being the most effective plays across all opponents.

An analysis by individual opponent shows that the more effective play-calling strategies do vary across the different teams, depending on their strengths and weaknesses and how they match up with the participant's team. The correlations (one-tailed) between the play-calling strategies and performance by team were examined to identify what strategies were most effective by opponent. Each of the run inside middle and right strategies was significant against one team, the run outside left strategy was significant against three different teams, the short pass was significant against two teams, the medium pass was significant against five teams, and the long pass was significant against three teams. This demonstrates that there is no one particular play-calling strategy that can be used all the time. The participants must consider the strengths of their team and of their opponent and must consider the situation they currently face.

Performance. Performance was measured by the number of offensive points scored for each game by the participant's team. Points scored on offense was used over defensive points allowed or relative points (offensive points scored – defensive points allowed) in order to preserve the focus on offense and avoid potential confounds concerning defensive performance.

General football strategy knowledge was evaluated at the beginning of the study (preknowledge) and upon completion of the study (postknowledge) for each participant. Two primary knowledge areas were evaluated: (1) offensive and (2) defensive strategy knowledge. Twenty-four questions in a multiple-choice format (12 items for each offense and defense) were taken from published football coaching texts to evaluate offensive and defensive strategy knowledge.

Each of the three knowledge sections was completed by four subject matter experts. Three subject matter experts were Division I-A college assistant coaches, and the fourth was a former high school

football coach. The average subject matter expert score for both offensive and defensive questions was 84%, as compared with the preknowledge scores for the novice participants of 59%. Considering the expertise continuum for the coaching profession from novice to long-time coaches of professional football teams, the subject matter experts used here would fall into the intermediate area in terms of football coaching expertise. The preknowledge scores of the participants, in comparison with the intermediate football experts, and the lack of any professional training or football experience strongly suggest that the participants were in the novice range in terms of football coaching knowledge and expertise.

RESULTS

Overall Performance

The overall performance mean among all the participants was 26.9 points per game and ranged from 16.4 to 33.5 points per game across the players. There was a general trend of increasing improvement in performance over time; the mean improvement across the 16 games was 0.4 points per game, ranging across participants from –1.2 to 3.3. The main focus of this article is on examining what role strategy variability played in this improvement and variability in offensive point performance.

Learning Strategy and Performance

To investigate the relationship between the learning strategies and performance, two repeated measures regression analyses were performed (Colquitt, Hollenbeck, & Gully, 1998; Pedhazur, 1997). Between-subjects effects were controlled for by subtracting the participant's mean across games for each independent and dependent variable. The effective sample size was 208 (13 participants \times 16 games) for each analysis. The presence of autocorrelations was also tested through the use of Durbin–Watson scores, which showed no significant autocorrelations.

The first regression analysis tested the relationship of learning strategy variability between games (strategy shift) and performance in terms of offensive points scored. The first half of the moderate experience hypothesis predicts a positive relationship with strategy variability among novice users as they obtain experience. The relationship was, surprisingly, in the opposite direction and was borderline significant [$\beta = -0.12; t(193), p = .08$]. This weak relationship suggests that a low degree of learning strategy variability might be related to a high degree of performance.

The second regression analysis tested the relationship between time spent using the learning strategies and performance. A similar but stronger relationship was found between the time spent on the various learning strategies and performance. All of the strategic [$\beta = -0.18; t(204), p < .01$] and tactical [$\beta = -0.14; t(204), p < .05$] information-gathering and practice [$\beta = -0.16; t(204), p < .05$] learning strategies significantly predicted performance with negative relationships. These significant negative relationships support the weaker finding among strategy variability and suggest that the less time individuals spend on the learning strategies, the higher their performance will be. This finding appears counterintuitive.

tive and contradicts the original hypothesis of a positive relationship.

Coaching Strategy and Performance

Like the hypothesis above, our second hypothesis instantiated the first half of the *moderate experience hypothesis* as the relationship between *coaching* strategy variability and strategy shift to performance, moderated by appropriate selection of the strategies. To test the main and moderation effects, two hierarchical repeated measures regression analyses were conducted with two steps each (see Table 1) for both strategy variability and shift. The first regression analysis tested the relationship between coaching strategy variability and performance. The first step entered was strategy variability and matching selection. The second step tested the interaction of strategy variability and matching selection. As was expected, more appropriate matching selection positively predicted performance. That is, more appropriate play calling significantly increased offensive performance. However, strategy variability again had a surprisingly significant *negative* relationship with performance, which indicates that low strategy variability leads to high performance. This result is the opposite of the effect that would be expected from the first half of the *moderate experience hypothesis* but is consistent with the learning strategy results.

The moderated interaction of strategy variability (within game) and the matching selection criteria was also significant; the combination of the matching selection criteria and strategy variability predicted performance above and beyond what each individually predicted. In other words, as was originally predicted, both strategy variability and adaptive or appropriate selection criteria are important to performance.

Further investigation revealed the nature of the interaction relationship. A median split of points scored by

strategy variability and the match-up scores showed that individuals with high match-up scores and low strategy variability had the most offensive points scored (30.65; see Figure 2). The graph shows that strategy variability hurts both low- and high-selection participants but has a larger effect on those participants using better selection criteria.

This moderating interaction effect reinforces the effects seen in the first step of the moderated regression, where strategy variability negatively influences performance, whereas the use of more appropriate selection criteria improves performance. The combination of low strategy variability and the use of appropriate selection criteria led to maximal performance, whereas the use of poor selection criteria (i.e., the lack of play calling based on one's own team and the opponent's strengths and weaknesses) and a high degree of variability among strategies proved to be the worst.

The second regression analysis tested the same type of moderated relationship with strategy shift (between games) and performance moderated by the matching criteria. The same hierarchical two-step procedure was used, and a similar pattern of results was found (see Table 1). Strategy shifts had a negative relationship with performance, whereas the matching selection criteria positively predicted performance. As for strategy variability, these results suggest that reduced strategy shift between games leads to higher levels of performance and that greater use of the appropriate selection criteria aids performance. The moderated interaction of strategy shift and the selection criteria was also significant and showed the same pattern of results as that for coaching strategy variability.

The correlation between strategy variability within games and strategy shifting between games was very low ($r = -.04, p > .05$). This lack of a correlation indicates that these two facets of strategy variability are independent

Table 1
Moderated Repeated Measures Regression Results With Strategy Variability and Strategy Shifts: Predicting Performance

Step Predictor	β at Step	β Final	R^2	ΔR^2	95% Confidence Interval for β	
					Lower Band	Upper Band
DV: Points scored						
1. Strategy variability (SV)	-.15*	-.18**			-0.28	-0.01
Matching selection (MS)	.22**	.19**	.06**	.06**	0.08	0.35
2. Interaction (SV \times MS)	-.18**	-.18**	.09**	.03**	-0.32	-0.04
DV: Points scored						
1. Strategy shift (SS)	-.16*	-.16*			-0.30	-0.02
Matching selection (MS)	.18**	.20**	.06**	.06**	0.04	0.32
2. Interaction (SS \times MS)	-.14*	-.14*	.08**	.02*	-0.27	0.00
DV: Points scored						
1. Strategy variability (SV)	-.15*	-.17*			-0.29	-0.01
Strategy shift (SS)	-.16*	-.15*			-0.30	-0.02
Matching selection (MS)	.21**	.20**	.08**	.08**	0.07	0.34
2. Interaction 1 (SV \times MS)	-.17*	-.17*			-0.31	-0.03
Interaction 2 (SS \times MS)	-.13‡	-.13‡	.13**	.05**	-0.27	0.01

Note—DV, dependent variable. 95% confidence intervals for β apply to the first step of each variable entered. ‡ $p \leq .06$. * $p \leq .05$. ** $p \leq .01$.

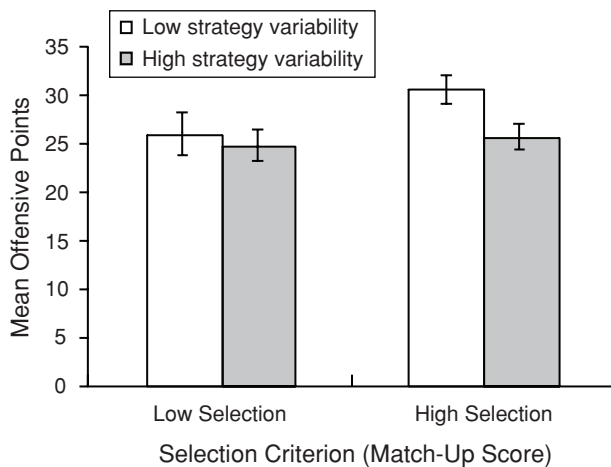


Figure 2. Mean offensive points scored by median splits for the interaction terms of strategy variability and matching selection criteria.

of one another. A third regression analysis was conducted to examine the potential independent effects of strategy variability and shift, along with the moderated interaction effects. The third regression in Table 1 shows that both strategy variability and shift independently and uniquely predict performance. The second step of the regression also displays unique significant effects for the interaction terms for strategy variability and the selection criteria ($p < .05$) and a nearly significant interaction effect for strategy shift and the selection criteria ($p = .06$).

Learning and Coaching Strategies

The use of various learning strategies should predict coaching strategy variability and strategy shifting in play calling. Two repeated measures regression analyses tested the relationship (within subjects) between time spent on the various learning strategies and the amount of offensive coaching strategy variability in each game, along with the amount of strategy shift between games. Between-subjects effects were removed by subtracting

participant means across games for independent and dependent variables.

The first regression analysis tested how well the three learning strategies, along with their variability before the games, could predict coaching strategy variability during the games (see Table 2). It was hypothesized that the learning strategies would be more strongly related to coaching strategy variability than to coaching strategy shift, because the learning strategies occurred immediately before the game was coached. Only the strategic information-gathering strategy significantly predicted variability in offensive strategies (Table 2). These results indicate that greater use of information gathering at a strategic or more conceptual basis leads to more variability in play calling.

The second regression analysis tested the relationships of the three learning strategies to the amount of strategy shifts that occurred between games (see Table 2). Among the learning strategies, only the practice strategy showed a weak and borderline significant relation. The lack of significant and strong relationships between the learning strategies and strategy shifts suggests that other variables besides the ones measured here influence the shifting that occurs with strategies *between* games.

Summary of Results

Evidence for several of the original hypotheses was found, as well as a key unexpected result. Figure 3 summarizes the *significant* ($p < .05$) relationships found among the general variables in this study. The repeated unexpected result was that strategy variability and strategy shift for coaching strategies and time spent on the learning strategies were negatively related to performance. The use of the matching selection positively predicted performance, as was expected, and moderated the significant relationship between strategy variability and shift to points scored. Information gathering at the strategic level was the only learning strategy to positively predict coaching strategy variability. The focus of the next section will be to explain the moderated relationship between strategy variability and performance further, through the use of quantitative and qualitative data.

Table 2
Repeated Measures Regression Results:
Predicting Strategy Variability and Strategy Shift

Predictors	β	Semi-partial r	95% Confidence Interval for β	
			Lower Bound	Upper Bound
DV: Offensive strategy variability				
Strategic information gathering	.17*	.16*	0.04	0.30
Tactical information gathering	.07	.07	-0.08	0.22
Practice	.10	.09	-0.003	0.20
Learning strategy shift	.00	.00	-0.14	0.15
DV: Offensive strategy shifts				
Strategic information gathering	.11	.10	-0.03	0.24
Tactical information gathering	-.02	-.02	-0.18	0.13
Practice	.13‡	.13‡	0.03	0.23
Learning strategy shift	-.02	-.02	-0.16	0.12

Note—DV, dependent variable. ‡ $p \leq .08$. * $p \leq .05$. ** $p \leq .01$.

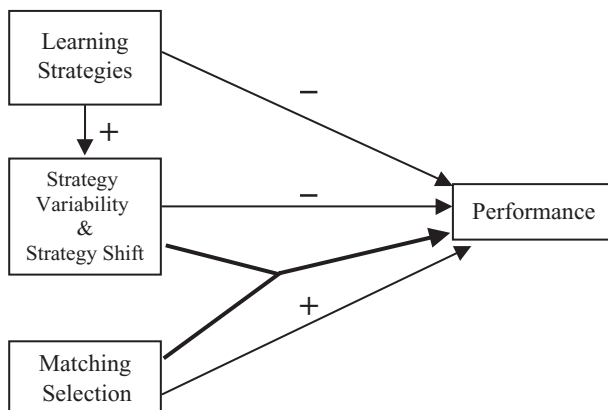


Figure 3. Overview of significant relationships, including the moderated interaction between strategy variability and shift with matching selection.

Exploratory Analyses on Strategy Variability and Development

Learning. Football strategy knowledge was evaluated on a between-subjects basis at the beginning of the study (preknowledge) and upon completion of the study (postknowledge) for each participant. The amount learned (postknowledge – preknowledge) was strongly correlated with time spent on strategic information gathering ($r = .64, p < .01, N = 13$). In other words, individuals who more often sought out information at the strategic or conceptual level learned more between the beginning of the study and the end. This result confirms the utility of the strategic information sources available to participants for improving declarative knowledge (although not for improving performance directly).

Trends over time. Due to the more proximal nature of strategy variability and performance occurring in the same game, strategy variability will primarily be used, instead of the strategy shift measure, for the following analyses. However, the general trends and relationships are similar for both strategy variability and strategy shift.

The effect of strategy variability and performance over time was examined. Figure 4A displays the average strategy variability scores for each game across all participants, which show a significant negative linear trend [$\beta = -.70, p < .01, 95\% \text{ CI } (-0.95, -0.45)$]. Congruent with the negative relationship between strategy variability and performance discussed above, performance displays the opposite linear trend, with a strong positive linear relationship over time, as shown in Figure 4B [$\beta = .69, p < .01, 95\% \text{ CI } (0.43, 0.94)$]. In other words, as the participants progressed over time and gained experience with the coaching task, they gradually reduced the amount of variability among their play-calling strategies, while increasing their performance. These results suggest that the individuals did associative learning and adaptively selected or deselected strategies on the basis of past success rates (Lovett & Schunn, 1999; Reder, 1987) through some type of strategy-pruning process across games.

Case studies. The effect of strategy variability on performance will be illustrated with data from 2 participants. These 2 participants are presented because they illustrate the basic trends many of the other participants displayed. The first participant, referred to as “Tom,” exhibited a high degree of strategy variability in the early games, which gradually decreased over time [$\beta = -.73, p < .01, 95\% \text{ CI } (-0.97, -0.49)$]. Figure 5A illustrates the same type of negative linear trend for Tom that was evident for the group as a whole (Figure 4A). Tom also showed a very strong positive linear relationship for performance over time [$\beta = .92, p < .01, 95\% \text{ CI } (0.79, 1.06)$], shown in Figure 5B. The relationship between strategy variability and performance for Tom was a robust $r = -.78, p < .01$. This pattern was typical for many of the participants and resulted in the overall trend presented in Figure 4B.

The second case, who will be referred to as “Jerry,” produced very different patterns of behavior across games. Jerry displayed a consistently high level of strategy variability across time and did not show the negative linear trend evident from the group data or Tom (Figure 5A). There was no significant linear trend for strategy variability [$\beta = -.24, p > .05, 95\% \text{ CI } (-0.58, 0.09)$]. Performance over time for Jerry was low (lowest total points scored among all participants) and does not display the same positive linear trend as that evident in the results for the group or for Tom. The trend line is flat and nonsignificant, with $\beta = -.01, p > .05, 95\% \text{ CI } (-0.36, 0.34)$ (see Figure 5B). Despite the lack of clear-cut trends shown by many of the other participants, the correlation examining the trial-by-trial relationship between lower strategy variability and higher performance for Jerry was still marginally significant ($r = -.38, p = .07$). In other words, even within Jerry’s profile, more variability reduced performance.

The results from these specific participants and the general results from the original hypotheses suggest that the gradual reduction of strategy variability over time may have a positive effect on performance. Tom displayed this pattern of behavior and showed unequivocal improvement in performance over time. Jerry, however, displayed a consistent level of strategy variability and did not display any significant performance improvements over time. The evidence examined so far suggests that a process of *pruning*, or adaptively selecting the more successful and efficient strategies, may have occurred over time for some of the participants (Lovett & Schunn, 1999). This gradual elimination of the inappropriate or less successful strategies may explain the negative linear trend evident for strategy variability over time. Stated another way, by coming to focus on a smaller set of strategies and continuing to use them more consistently over games, participants will have developed better expertise with how and when to apply those particular strategies.

Game-level patterns of strategy change. The same process of reducing strategy variability over time was also evident at a microlevel, or the individual game level. Each game was separated into two separate halves of the game,

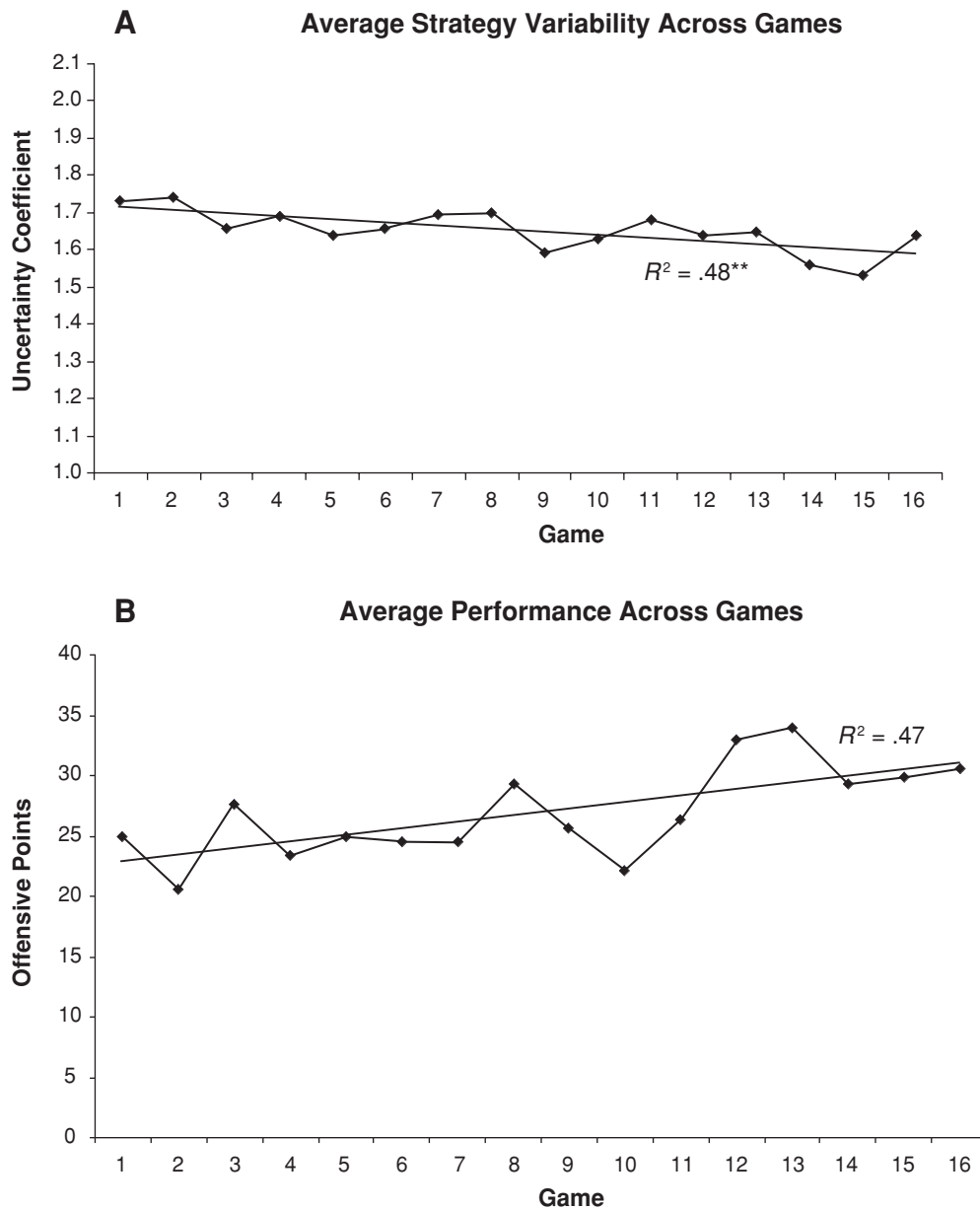


Figure 4. (A) Average strategy variability measured by the uncertainty coefficient across games for all participants. (B) Average performance measured by total offensive points scored across games for all participants.

the first and the second. Each half of the game consisted of two equally timed quarters (6 min of play-clock time each). A repeated measures ANOVA testing the differences between halves showed a significant difference in strategy variability [$F(1,207) = 8.76, p < .01$]. The first-half strategy variability was higher than that for the second half (1.58 vs. 1.52). Post hoc t tests for each participant showed a significant difference for 3 out of the 13 participants (P1, P2, and P3; see Figure 6). Each of these participants had higher variability in the first half and reduced strategy variability in the second half.

To assess whether the change in variability at the end of games was the result of conservative game playing in games in which players were far ahead or behind, strategy variability was compared between close scoring games (<10 points difference) and games that were not close in score (>10 points difference). Strategy variability scores were almost identical (1.66 for close games and 1.64 for nonclose games), suggesting that the drops in strategy variability were not the result of endgame play-calling conservativeness. Strategy variability between the first and third quarters of each game was also examined to see

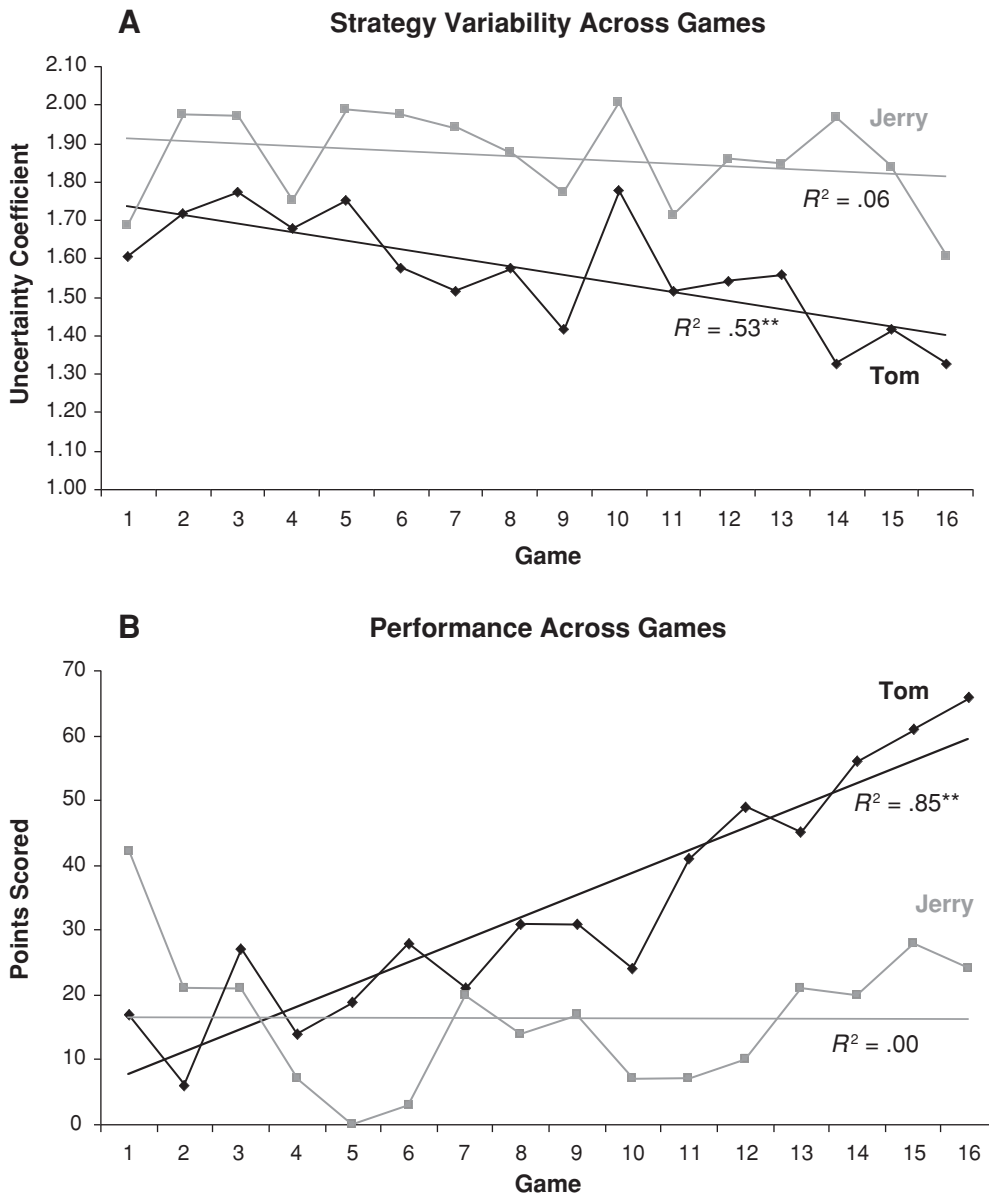


Figure 5. (A) Strategy variability measured by the uncertainty coefficient across games for Tom and Jerry. (B) Performance measured by the total offensive points scored across games for Tom and Jerry.

whether the drop in variability across halves was the result of endgame conservatism. Indeed, there was a drop from first quarter (1.61) to the third quarter (1.54), and it was, if anything, larger than the drop from the first to the second halves [$F(1,207) = 9.85, p < .01$].

Figure 6 displays the strategy variability, by half, for each participant. Even though only 3 participants showed a significant difference in halves, 10 of 13 participants displayed the general trend of higher variability in the first half, as compared with the second half. Therefore, this evidence suggests that a plurality of the participants conducted a dynamic microlevel pruning process in which

they used their experience from the first half to adaptively select strategies for the second half of the game.

DISCUSSION

The results of this study support the importance of considering behavior in a complex and dynamic task over time and the role of variability in strategy development. Interestingly, the important role of variability was upheld, but in a way opposite to that suggested by earlier work with simpler, static tasks. The present study also unpacked various learning strategies and the way in which those

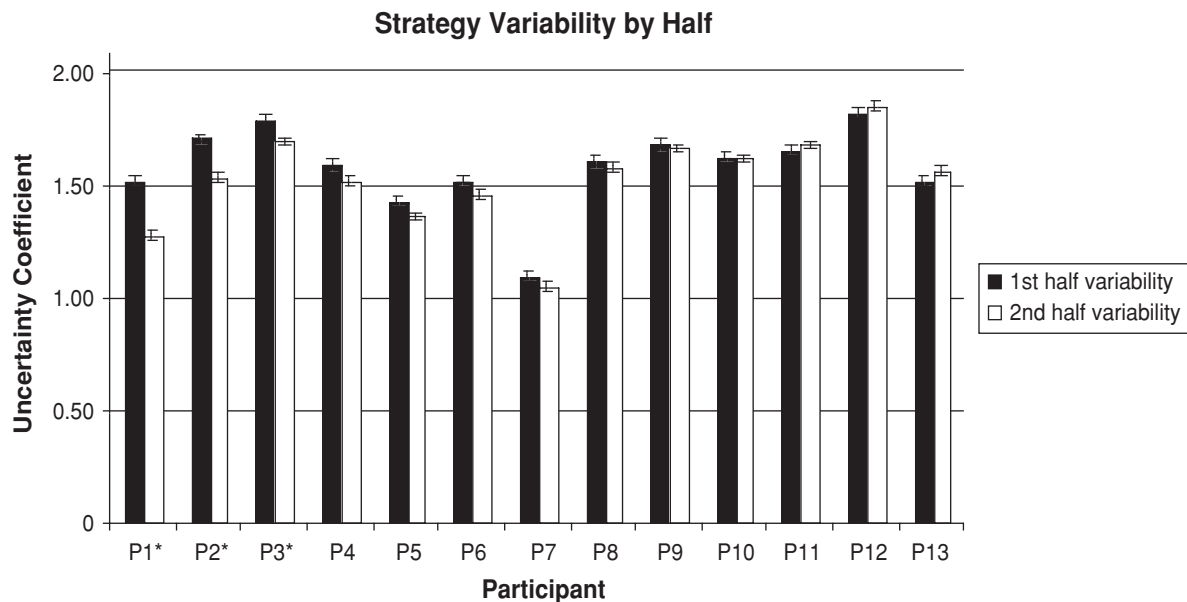


Figure 6. Comparison of the average strategy variability for each participant by halves, first half versus second half. Error bars represent standard error. * $p \leq .05$.

learning strategies contributed to variability and adaptive performance strategy use.

Learning Strategies

This study focused on instruction as one type of strategy that influences both the development of task-specific strategies and performance. Instruction was further defined and specified in order to examine observable behaviors that were self-motivated. These more specific learning strategies included information gathering at a strategic and tactical information level and a basic type of practice that emphasized repetition.

As was hypothesized, time spent using strategic information gathering positively predicted strategy variability. This positive influence suggests that when individuals gather information at a conceptual level in order to learn more about the task or domain, they are, in some sense, exploring the problem space conceptually. This exploration produced greater variability during the games among their play calling and could have led to the discovery or realization of new knowledge about the target task (Anderson, 1981).

On a between-subjects basis, use of strategic information gathering showed a strong relation with the amount of knowledge learned from the beginning of the 3-month study to the end of the study. In other words, information gathering at the strategic level does facilitate learning at a metacognitive level, which was displayed in the amount of knowledge learned, as well as in increased strategy variability.

Tactical information-gathering learning strategies showed no significant relationships to coaching strategy variability or shift. This level of information deals with

specific data and may be too detailed for novices such as the ones used in this study to use and apply effectively. For instance, a participant trying to evaluate roster players at the running back position needs to understand the role and responsibilities of that position, what physical skills and characteristics are important for a running back, and how those player attributes interact with certain plays and other teammates. If the participant does not have a basic understanding of how to correctly interpret and apply such knowledge, it will be largely ineffective. Additional research is needed to follow up on various possible relationships of tactical information-gathering strategies to performance and task-specific strategy development.

The practice strategy also showed no significant relationship to coaching strategy variability or shift. Ericsson and Lehmann (1996) have provided evidence not only that an important aspect of practice is the quantity or time spent during practice, but also that the quality of practice affects skill acquisition and performance. They differentiated practice that consists primarily of simple repetition and another type that requires more effort, or *deliberate practice*. The simple repetition type of practice is easier and would seem to foster primarily associative learning based on the consequences or outcomes. Repetitive practice should gradually decrease variability among strategies as the strategies are slowly pruned down to the more efficient ones according to their success rates.

Deliberate practice, however, is more structured than repetitive practice, and there is a conscious effort made to extend, explore, and elaborate skills and domain-related knowledge (Ericsson & Lehmann, 1996). Therefore, deliberate practice should cause both metacognitive and associative learning and should involve the information-

gathering and practice strategies responsible for each. This type of practice is congruent with Crowley et al.'s (1997) competitive negotiation theory between metacognitive and associative learning. Crowley et al. provided a model of learning and development that involves both metacognitive and associative learning interacting in a competitive negotiation according to the situation and circumstances. More information on the learning strategies related to deliberate practice might help shed light on skill acquisition. A clear distinction between the types of practice was not done in this study and could be a reason why there were no significant relationships found with practice strategy and coaching strategy variability. This would be an interesting area to investigate in future research.

Coaching Strategy Variability and Shift

It was initially hypothesized that greater variability among strategies would be positively related to performance, on the basis of results with first half or novice users described in the moderate experience hypothesis (Siegler, 1996; Siegler & Taraban, 1986) and other past findings (Alibali & Goldin-Meadow, 1993; Siegler, 1995). The participants in this study were novices, as was indicated by the lack of football coaching experience, little to no experience with the football simulation used, and their lack of general football strategy knowledge, in comparison with the subject matter experts. The surprising negative relationship, consistently found, of learning strategies, coaching strategy variability, and shift with performance suggests that greater variability is evident *earlier* in experience with this task. This finding is incongruent with the first half of the moderate experience hypothesis (Siegler, 1996), which predicts that novice individuals will possess low amounts of variability among their strategies and that variability will increase with experience until some level of expertise is obtained.

A closer inspection of the participants, divided by their preknowledge scores, showed the same general negative trend for strategy variability across time among *both* high- and low-preknowledge participants (Hansberger, 2004). The participants were also divided and examined by how much they learned from the beginning of the study to the end. Again, the same basic negative trend for strategy variability over time was evident for both high and low learners. Strategy variability began at a high level and gradually decreased over time, regardless of the participants' incoming knowledge levels of football coaching or how much they learned during the course of the study. As the case studies presented here suggest, there are differences among individuals. However, there was only 1 individual out of 13 who showed a significant positive trend of strategy variability over time. That individual showed no improvement over time, like Jerry.

The reason our results are inconsistent with the first half of the moderate experience hypothesis may be attributable to the nature of a dynamic versus a static task. The primary evidence for the moderate experience hypothesis comes from a study in which a moderately simple and static task environment was used with a balance scale

(Siegler & Taraban, 1986). One primary difference between the coaching football task and the balance scale task is the changing nature of the football task over games/opponents and the complex problem space. The complex problem space allows for a greater variety of strategies to be selected and used, and the dynamic nature of the task does not present the participants with a readily transparent structure or easily identifiable optimal solution for the task. There is also no assurance of a single correct or winning strategy across all games. The uncertainty created by these characteristics requires a different approach by the participants. On the basis of the results from this study, the best approach may be to immediately begin with a high degree of strategy variability and start adaptively pruning out the unsuccessful strategies. If this is correct, the basic human approach to complex dynamic situations may be to begin with a large variety of different strategies in an attempt to quickly prune out the less successful ones and focus on the successful ones.

Along with the difference in task complexity, there may also be a difference in the participant knowledge level between this study and Siegler and Taraban's (1986). There may be a category prior to the novice on the coaching continuum that would include "naive" people who have no or almost no knowledge of the domain. There is little question that the participants in this study are novices on a coaching continuum, since they do have some knowledge of the domain (as evidenced by their moderate preknowledge scores), but lack the experience, knowledge, and training of many others in the domain.

It is important to note, however, that in the real world, people tend *not* to enter into practice activities for complex dynamic tasks (e.g., driving, sports, teaching, and air traffic control) without some level of "book knowledge." That is the reason why the coaching continuum starts at the novice level and does not include the naive level. The novice level of the participants in this study is typical of novices that usually begin real practice and training for real tasks.

Past research has shown that pretest measures of high strategy variability lead to greater learning in such domains as math (Alibali & Goldin-Meadow, 1993; Siegler, 1995). These findings suggest that high variability among strategies should be positively related to learning and performance. The results of this study are consistent with the view that high strategy variability early in a participant's experience with a task aids final performance levels. However, additional information is added to the relationship between strategy variability and performance. Decreasing amounts of variability among strategies was associated with increasing levels of performance over time. This discovery may have been undetectable in past research, due to the simple tasks used, the limited number of possible strategies available, or the inability to track and record the strategies in the same manner as in this study over several months.

Further study is needed to verify these negative strategy variability trends among novices in dynamic settings. Future investigations should focus on the development of

strategy variability over longer periods of time from that of novice, and even naive, users to that of intermediate users or should compare the strategy variability of novices with that of intermediates in dynamic environments.

Strategy Pruning Process

Past research and much of the probability-matching research has shown that individuals are sensitive to the success rate of past behaviors and strategies (Herrnstein, 1961; Lovett & Schunn, 1999; Reder, 1987). On the basis of that prior information, individuals will adapt future behaviors and strategy use for simple and relatively static tasks. In light of these research findings and the general trends of strategy variability and performance over games in this study, there appears to be mounting evidence for the application of the matching law to strategies.

The application of the matching law to strategies suggests that a basic process of pruning exists where successful strategies get passed on for future use and unsuccessful ones die out. There is also mounting evidence that this process exists at several levels. Schunn and Reder (2001) found that individuals could adapt their strategy use at both macro- (an hour) and microlevels (minutes). The present research extends the macrolevel to longer time periods of over 2 months. It also supports the hypothesis that strategy adaptation occurs at a lower level (i.e., within an hour).

Siegler (1996) stated that adaptation was the output of a combined source of variation and selection, which is precisely what was measured through the moderated interaction of strategy variability and the matching selection criteria. The interaction proved to be significant for strategy variability within and between games. These interactions provided evidence that combining appropriate selection criteria with lower coaching strategy variability and shift provides the highest levels of performance and adaptation to a complex and dynamic task. This interaction further supports the effect of appropriate strategy pruning on performance. The evidence for the pruning process and these significant interaction effects help define what it means to successfully adapt to a dynamic task over time.

Summary and Future Directions

Past research has shown that high strategy variability leads to greater learning (Alibali & Goldin-Meadow, 1993; Siegler, 1995). Individuals are also sensitive to the success rate of past strategies and will adapt future strategy use on the basis of prior information for simple and relatively static tasks (Lovett & Schunn, 1999; Reder, 1987). This study integrated these findings with a highly dynamic task over a sustained period of time. The results from this study suggest that variability plays a major role in adaptive skill acquisition, from the developmental process of strategies to task performance and learning. These findings suggest that the strategies involved in skill acquisition of this football-coaching task were pruned over time to improve performance. Individuals began their dynamic task interactions with a wide range of strategies and began to prune or eliminate the unsuccessful strategies while uti-

lizing appropriate selection criteria over time, ending up with a limited repertoire of successful strategies.

In order to better understand the role of variability in more dynamic and realistic situations, future research should continue to use dynamic task settings in the study of strategy development and deployment. Continued use of dynamic tasks is also needed for the generalizability of strategy development research findings to real-world training, skill acquisition, and performance scenarios. It is also important for future research to examine other levels of strategic choice in similar complex tasks in order to investigate interrelationships between strategy variability across different layers of complex tasks.

REFERENCES

- ALIBALI, M. W., & GOLDIN-MEADOW, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, *25*, 468-523.
- AMERICAN FOOTBALL COACHES ASSOCIATION (1995). *Football coaching strategies*. Champaign, IL: Human Kinetics.
- AMERICAN FOOTBALL COACHES ASSOCIATION (2000). *Defensive football strategies*. Champaign, IL: Human Kinetics.
- ANDERSON, J. R. (Ed.) (1981). *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.
- ASHCROFT, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology: Progress in cognitive development research* (pp. 302-338). New York: Springer.
- BILLICK, B. (2001). *Developing an offensive game plan*. Monterey, CA: Coaches Choice.
- BRYANT, P. E., & TRABASSO, T. (1971). Transitive inferences and memory in young children. *Nature*, *232*, 456-458.
- CHURCH, R. B., & GOLDIN-MEADOW, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*, 43-71.
- COHEN, M. (2001). *Prima's official strategy guide: Madden 2001*. Roseville, CA: Prima Publishing.
- COLQUITT, J. A., HOLLENBECK, J. R., & GULLY, S. M. (1998, April). *Repeated measures regression: Decomposing variance in cross-level research*. Paper presented at the Thirteenth Annual Conference of the Society for Industrial and Organizational Psychology, Dallas.
- CROWLEY, K., SHRAGER, J., & SIEGLER, R. S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review*, *17*, 462-489.
- ERICSSON, K. A., KRAMPE, R. T., & TESCH-RÖMER, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.
- ERICSSON, K. A., & LEHMANN, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273-305.
- ESTES, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, *53*, 134-140.
- FERGUSON, C. A. (1986). Discovering sound units and constructing sound systems: It's child's play. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 36-57). Hillsdale, NJ: Erlbaum.
- FISK, A. D., ACKERMAN, P. L., & SCHNEIDER, W. (1987). Automatic and controlled processing theory and its applications to human factors problems. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 159-197). Amsterdam: North-Holland.
- HANSBERGER, J. T. (2004). The effects of prior knowledge on strategy development & performance. In *Proceedings of the 48th Annual Human Factors and Ergonomics Conference* (pp. 453-457). Santa Monica, CA: Human Factors and Ergonomics Society.
- HERRNSTEIN, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267-272.

- JAKOBSON, R. (1968). *Child language, aphasia and phonological universals* (A. R. Keiler, Trans.). The Hague: Mouton. (Original work published 1941)
- KOEHLER, M. D. (1992). *Football coach's survival guide*. West Nyack, NY: Parker.
- KUHN, D., GARCIA-MILA, M., ZOHAR, A., & ANDERSEN, C. (1995). *Strategies of knowledge acquisition* (Monographs of the Society for Research in Child Development, Vol. 60, No. 4). Chicago: Society for Research in Child Development.
- KUHN, D., & PHELPS, E. (1982). The development of problem-solving strategies. In H. Reese (Ed.), *Advances in child development and behavior* (Vol. 17, pp. 1-44). New York: Academic Press.
- LEFEVRE, J.-A., SADESKY, G. S., & BISANZ, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 216-230.
- LEWONTIN, R. C. (1978). Adaptation. *Scientific American*, **239**, 212-230.
- LONG, H., & CZARNECKI, J. (1998). *Football for dummies*. Foster City, CA: IDG Books.
- LOVETT, M. C., & SCHUNN, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General*, **128**, 107-130.
- MARKMAN, E. M., & SEIBERT, J. (1976). Classes and collections: Internal organization and resulting holistic properties. *Cognitive Psychology*, **8**, 561-577.
- MCCLELLAND, J. L., & SIEGLER, R. S. (Eds.) (2001). *Mechanisms of cognitive development: Behavioral and neural perspectives*. Mahwah, NJ: Erlbaum.
- MILLER, P. H. (1993). *Theories of developmental psychology* (3rd ed.). New York: Freeman.
- ORASANU, J., & CONNOLLY, T. (1995). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. Zsombok (Eds.), *Decision making in action: Models and methods* (pp. 3-20). Norwood, NJ: Ablex.
- PEDHAZUR, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York: Harcourt Brace.
- PENNINGTON, N., & GRABOWSKI, B. (1990). The tasks of programming. In J.-M. Hoc, T. R. G. Green, R. Samurcay, & D. J. Gilmore (Eds.), *Psychology of programming* (pp. 45-62). San Diego: Academic Press.
- PIAGET, J. (1952). *The child's conception of number* (C. Gattegno & F. M. Hodgson, Trans.). London: Routledge & Paul.
- REDER, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, **89**, 250-280.
- REDER, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, **19**, 90-138.
- REDER, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 227-259). San Diego: Academic Press.
- SCHAUBLE, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology*, **49**, 31-57.
- SCHMIDT, R. A., & LEE, T. D. (1999). *Motor control and learning: A behavioral emphasis* (3rd ed.). Champaign, IL: Human Kinetics.
- SCHUNN, C. D., & REDER, L. M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal of Experimental Psychology: General*, **130**, 59-76.
- SHERRY, J., LUCAS, K., RECHTSTEINER, S., BROOKS, C., & WILSON, B. (2001, May). *Video game uses and gratifications as predictors of use and game preference*. Paper presented at the International Communication Association Conference, Washington, DC.
- SIDMAN, M. (1952). A note on functional relations obtained from group data. *Psychological Bulletin*, **49**, 263-269.
- SIEGLER, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, **28**, 225-273.
- SIEGLER, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- SIEGLER, R. S. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General*, **127**, 377-397.
- SIEGLER, R. S., & CROWLEY, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, **46**, 606-620.
- SIEGLER, R. S., & ROBINSON, M. (1982). The development of numerical understandings. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 16, pp. 241-312). New York: Academic Press.
- SIEGLER, R. S., & TARABAN, R. (1986). Conditions of applicability of a strategy choice model. *Cognitive Development*, **1**, 31-51.
- SIMON, H. A., & CHASE, W. G. (1973). Skill in chess. *American Scientist*, **61**, 394-403.
- THEIL, H. (1972). *Statistical decomposition analysis: With applications in the social and administrative sciences*. Amsterdam: North-Holland.
- WELFORD, A. T. (1968). *Fundamentals of skill*. London: Methuen.

(Manuscript received December 20, 2004;
revision accepted for publication September 9, 2005.)