MEASURING COGNITIVE LOAD MANAGEMENT IN A TRADITIONAL MARTIAL ARTS TRAINING MODEL

A Dissertation

by

HERBERT N. MAIER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2004

Major Subject: Educational Psychology

MEASURING COGNITIVE LOAD MANAGEMENT IN A TRADITIONAL MARTIAL ARTS TRAINING MODEL

A Dissertation

by

HERBERT N. MAIER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

William R. Nash
(Chair of Committee)

Joyce E. Juntune
(Member)

Robert L. Hall
(Member)

David L. Wright
(Member)

Victor L. Willson
(Head of Department)

May 2004

Major Subject: Educational Psychology

ABSTRACT

Measuring Cognitive Load Management in a Traditional Martial Arts Training Model.

(May 2004)

Herbert N. Maier, B.S., University of Houston; M.S., Texas A&M University

Chair of Advisory Committee: Dr. William R. Nash

A training method utilized in a few martial arts was found to agree strongly with current cognitive psychology theory. Further study extracted a procedural model for learning a complex set of whole-body, dyadic motor skills involving high-speed, interactive, continuous situation assessment and decision making. A broader literature survey found relevance in several fields of research, supporting the definition of four performance dimensions in the activity. Data collected from one experienced student partnering with each of ten students of various experience levels was analyzed on these four dimensions. These dimensions were found sufficient to show both individual differences and changes across an instructional intervention. Strong correlations found under linear regression were supportive of anecdotal evidence from the model's long empirical history in training. Data provided evidence of a self-organizing dynamic emerging from the interaction of a dyad participating in this activity, and of individual differences in cognitive resource management dynamically setting allocation priorities among specific aspects of a complex motor/cognitive activity. Highly individual responses demonstrate a mechanism for insight into students that are difficult to read. Numerous comparisons and contrasts show interactivity of performance dimensions. Impact is foreseen for research, training and testing in motor learning fields, as well as situation awareness, decision making and military tactical training. Further research

is recommended to replicate these findings, test hypotheses derived from them, and to extend testing of the drill-network model into other fields of learning.

ACKNOWLEDGEMENTS

I would like to thank my wife, Katherine H. Taber-Maier, Ph.D., for 2 decades of pushing me to go back to school, and then 5 years of acting as a sounding board. I would also like to thank the two people who opened the door: Dorothy Werblo, Ph.D., for pushing me to present my work at the American Creativity Association conference in 1998, and William R. Nash, Ed.D., whom I met working on the conference committee, and who invited me into the Educational Psychology Ph.D. program. Next, I would like to thank my committee members: Robert L. Hall, Ph.D., whose excellent class in cognitive psychology showed me that my ideas had validity, Joyce E. Juntune, Ph.D., whose coaching in qualitative research showed me that alternatives were available, and David L. Wright, Ph.D., who said that this research line would keep me busy for a lifetime.

I also need to thank the many outside individuals who gave early foundation or recent feedback and encouragement, including alphabetically (Oriental names are family name first):

- Donald Barthlme, AIA
- Charles Christ, Ph.D. (US Army Research Institute)
- Nancy Cooke, Ph.D. (Arizona State University)
- Mica Endsley, Ph.D. (Cognitive Technologies, Inc.)
- Robert Hoffman, Ph.D. (Institute for Human and Machine Cognition, University of W. Florida)
- Glenn James, Ph.D. (USAF retired)
- Kim Pyung Soo (GrandMaster, founder of Cha Yon Ryu martial arts)
- Stanley Roscoe, Ph.D. (retired)

- Lawrence G. Shattuck, Ph.D. (Colonel, US Military Academy)
- Wayne Shebilske, Ph.D. (Wright State University)
- Andrew Stricker, Ph.D. (TAMU, now at Vanderbilt University)
- Wang Kiu (GrandMaster, Wing Chun, under Yip Man)
- Sean Winkler, Ph.D. (Lincoln Laboratories, MIT)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

It is likely that few fields of endeavor have received as much effort, creativity and field-testing, as human combat. Records of the existence of martial arts training go back to 12th century in the Philippines (Inosanto, Johnson, & Foon, 1977), 7th century in Korea (Han, 1974; Kim, 1973), and farther back in China, India, and Greece. Some direct instructional lineages can be traced back many generations. Of equal importance to tradition is constant innovation. Innovation may be motivated by differences in circumstance, such as climate, clothing, immediate surroundings, and objects at-hand. Innovation is also required to replace knowledge lost through secrecy, death, inaccurate/incomplete transmission, and variously motivated suppression—challenges to any oral tradition. Innovation may be triggered by new insight, inspired either directly by the activity itself, or indirectly by metaphor/transfer from other experience. Thus, it is common knowledge that many “styles” of martial art exist, but not why that is so. It comes from having to deal with many different specific contexts. A dramatic example is the flying kicks that are so popular with movie and television audiences, and with young practitioners. Like the knights in medieval Europe, Oriental knights and cavalry were the equivalent of modern tanks. Flying kicks were the rocket-propelled grenade of the time—an individual hurling himself as a missile, hoping to survive, but dedicated to un-horsing the rider. There are such serious disadvantages to flying kicks that, outside of exhibitions and entertainment, they are acts of desperation (and only for the young).

This dissertation follows the style of *Human Factors*.

The traditional martial arts are a field of pure human performance. The Okinawan word “Kara-te” is translated as “empty hand”. The visible bearing of arms has a tendency to escalate violence; showing that one is unarmed is a universal indicator that a peaceful conclusion is desired. In addition, weapons tend to be confiscated and training to be suppressed in any people who have been invaded and defeated. The Han in China, the Japanese in Okinawa, the Americans in Japan, the Spanish in the Philippines, all imposed such changes. Suppression is seldom as effective as the conqueror would wish. Inosanto, Johnson, Foon (1977) described the manner in which Filipinos preserved their combat arts by concealing them in dance.

Most weapons involved in traditional martial arts training are simple farm tools or household objects. The adaptation of an object with which one is already familiar is easier than learning ab initio to work with a new object. Success with a weapon is judged on the extent to which its manipulation becomes an extension of the body. Fast, fluid, effective performance is developed through cultivation of an integrated mind and body. This is not a mystical idea; it is the same sort of integration that occurs in learning to keyboard, or to drive a car well. Fast appropriate responses require instant access to organized knowledge. The traditional martial arts have continued for many, many centuries, and will continue for many more, with absolutely no modification due to technological change. Successful methods developed around fires centuries ago are still perfectly valid under streetlights today. This fact is difficult for those committed to selling the newest technology, but resonates strongly for those with more practical interests.

The timelessness of fundamental problems is supported by twentieth century archeology and cultural anthropology, which have accumulated vast evidence that earlier generations were solving far more complex problems than recent generations generally give them credit for. James and Thorpe (1994) said “Like Needham, we feel that it is a fatal mistake to underestimate the technological and intellectual achievements of an ancient people” (p. xxiii).

Luttwak (1987) aptly described the importance of timeless fundamentals in terms of the paradoxical nature of combat. He pointed out that the more effective a technology is, the faster it is neutralized or overcome by a different one, constantly returning the deciding factor to the decisions and actions of the individual combatants. Likewise, in describing the life of a combat pilot, Watts (1996) emphasized the importance of: “…the ability of opposing aircrews to develop and sustain accurate representations of where all the participants in or near the air combat arena are, what they are doing, and where they are likely to be in the immediate future” (p. 94). As the dimension of technology goes through cycles of being neutralized, the human performance dimension keeps coming to the fore.

Luttwak’s discussion focused on the impact of higher technology, and the goal of new technology being to render older technology inferior. However, in motivated hands, lower technology can still overcome higher technology, because every technology has both strengths and weaknesses. Inosanto (ibid.) describes the U. S. Marines as acquiring the nickname of “leathernecks” from the two layers of leather worn around the neck to stop the blades of the Moros in the Philippines. It can be very much an issue of scissors/paper/stone, because strengths and weaknesses can exist on separate, even independent, dimensions. This issue of dimensionality will be extremely important in the current study, first in defining the dimensions to be measured, second on interactions between them, and third, in differentiating instructional effect on each one.

The martial arts field has not received the serious academic or research attention that some of its activities and ideas deserve. Research and academe are strongly oriented toward written/verbal material. Any field lacking written records is thus less attractive. One factor contributing to a lack of written records in the martial arts is a concern for secrecy. People have died and battles have been lost because a secret escaped. Silence and disinformation are innate to any competitive enterprise, including military, business, and political efforts, all of which rely strongly on surprise and misdirection.

Possibly a stronger reason for the lack of written records is that generating a verbal representation of a non-verbal, physical activity is not a trivial task. Much of the knowledge in the traditional martial arts is preserved in the mechanics of performance. Like dance, there is not even a successful system of notation for the superficial movements and positions, let alone a descriptive language of the functions and mechanisms sufficient to avoid slipping into metaphor and allegory. In the well-known martial arts movie, “Enter the Dragon”, Lee’s famous line is “Don’t think—feel! It is like a finger pointing the way to the moon…”

This difficulty is far from unique to the martial arts. Edwards (1979) described in the preface of her book the experience of being unable to switch modes from drawing to verbally explaining how she drew in order to teach it. She also observed a tendency of students to improve abruptly rather than gradually. She said they described, vaguely, that they were just seeing differently. This difficulty in communicating how a physical activity is actually performed certainly supports the perception in outsiders that whatever is happening is some kind of “gift” that will remain outside their reach. It does not occur to them that activities like drawing would have anything of importance in common with activities like dance, or gymnastics, or theatre, or martial arts. As Edwards showed in learning to teach drawing, the challenge of verbalizing the non-verbal is not insurmountable.

Prior to developing clear written descriptions, a physically-based activity can only be studied effectively by persons knowledgeable in that particular field. Considering the time and effort that must be invested in training a complex physical skill, and the time and effort required to train the skills of research and writing, there will be few persons investing in both. There is a long-standing cultural dissonance, in the West as well as in the Orient, that on the one hand, any sort of pugilism demeans and is appropriate only to the lowest social classes, and yet on the other hand, martial arts training potentially offers lessons about inner strength, virtue, and self-restraint valuable to those headed for nobility and leadership. There are also traditions in the martial arts emphasizing that training increases intelligence—this would concur with researchers like Torrance and Safter (1999) who cites his own and others' research supporting kinesthetic/psychomotor/rhythm/manipulation modalities in learning problem solving, basic concepts, language skills, and creativity (p. 188-189). It is possible that the perceived dissonance between education and training arises from the dual challenges of making time to advance at both activities, and of training oneself both verbally and non-verbally.

Adaptations of traditional martial arts can also diminish their attractiveness as research topics. There are now multiple significant purposes and goals of martial arts training that may obscure effects being sought. Modern elite military is still interested in actual combat application, whereas police are more focused on methods of restraining/subduing with limited force. Many civilians want to either watch or participate in sport with various levels of safety. Prior to the 20th century, most of the motivation toward martial arts training was the ability to fight effectively. Since World War II, many traditional martial arts have been adapted to non-combat uses. Japan removed the crippling and deadly techniques from Jujishu (the training of the Emperor's elite guard) and successfully promoted Judo as an Olympic sport. After the Korean war, Korea started a long but similar process toward the same goal with its synthesis of Tae Kwon Do. In the 1980's, the People's Republic of China officially declared the composite style WuShu to be the official martial art of China. Many other styles of Oriental martial arts have similarly promoted sports-oriented programs and international networks of tournaments where trophies reward competition in considerable safety. For these reasons, this researcher strongly feels that it is necessary to focus research upon proven traditional martial arts, rather than simplified modern versions, to find the most-intact structures and procedures to investigate.

Rituals, practices and procedures can encode a great deal of integrated multi-dimensional information. Like an unknown language, the trick is extracting the conceptual systems that give a practice its structure. Independent of the many "styles" being taught, there are numerous instructional methods in use, with very different results. A single factor seems to describe much of the variance. That factor, consistent with the idea of organized knowledge, is the degree of integration among the 3 common aspects of modern martial arts training: forms, self-defense and sparring. One particular instructional method appears to integrate these aspects quite well, generating knowledge that is more organized and therefore more instantly accessible. It focuses specifically on structuring and training the decision cycles that must occur at pre-verbal, and often pre-conscious speeds. This approach instills and rehearses organized integrated procedural and declarative knowledge, including distinguishing features of different situations, providing a foundation for developing recognition, judgment and very fast decision-making aimed at quick control of a situation. Its structure provides highly detailed information on which parts of the knowledge base are being mastered, and which still require focused attention. The ultimate purpose of the approach is efficient construction of expertise for quick decision making.

Appropriateness of the Study

Evidence of this study's timeliness in its cognitive, instruction/training and testing aspects, include Sweller, van Merrienboer, & Paas (1998), who wrote that "it is not common to measure cognitive load while conducting research on instruction" (p. 267). Similar comments were made in 1979 (Shea & Morgan, 1979). The timeliness of this study's movement (kinesiology) aspect is indicated by a special issue of Brain and Cognition, devoted to Movement Timing and Coordination (Krampe, Engbert, & Kliegl, 2002), and the "International Conference on Motion, Attention, and Perception", June, 2002, in France. Rumiati (2000) referenced Morris as saying that skilled movements are a primary development modality in the child, going on to say that though this ability remains important throughout life, *acquisition and structuring of this cognitive domain remain unexplained* (italics added). Chamberlain and Coelho (1993) wrote that research had modestly described decision-making, but needed more applied work on development of expert levels. They call investigation of training in decision-making "minuscule", owing to limitations in methodology including "the lack of a connection to actual game decision-making" (p.151). This study contributes to this area in its ability to closely track development of performance during learning, with strong inference of ongoing knowledge construction, and improvement of recognition and decision skills. McGarry, Anderson, Wallace, Hughes and Franks (2002) wrote that there is still little research on inter-person coordination in sport.

This researcher has often heard the advice to “go with your strengths”. In truthful assessment, a non-traditional student’s greatest strength probably resides somewhere within the years of experience brought to the graduate school endeavor. Twenty-seven years of teaching and of watching teachers in several different martial arts traditions is certainly relevant to an advanced degree in educational psychology, especially since the experience motivated returning to school. Appropriateness for this researcher performing this study is increased by the fact that there are very few people with this amount of teaching experience in the martial arts, even fewer with certification in the specialty being studied, and vanishingly few of these with an interest in learning to do research. This combination of factors makes the selection of a topic in this area more an obligation to the field, than an option to the person.

Purpose of the Study

The purpose of this study is to investigate one instructional activity drawn from traditional martial arts in order to substantiate the instructional benefits behind its long use. This investigation will include relating the activity to current research literature, and observing it in use with a small sample of students. As generations of users have found performance improvement to be directly visible, the rate of improvement can reasonably be expected to be considerable. Practices and procedures in traditional martial arts encode complex structures of information, more complex than popular modern martial arts. The lack of written records does not indicate lack of valuable knowledge. Like translating an unknown language, the trick is extracting the concepts that give a successful practice its structure.

This study focuses on an activity that will be described in detail in Chapter III. Preliminary comparison with basic cognitive theory (Maier, 1999) (Appendix A) suggested the activity to be a strong empirical demonstration of schema theory, as well as of composition and proceduralization.

Instructional purposes, goals, and challenges related to this activity can all be defined, setting a basis for clear description of the core ideas encoded in the activity. The activity's instructional purpose is to instill complex knowledge and related physical skills quickly in an organizational structure supporting instant (<0.1 s) access under high-stress conditions. Speed of access is absolutely essential—deadlines in the martial arts are unforgiving and a late response is at best useless, at worst disastrous. The activity's instructional goals are measurable components of the instructional purpose, and are defined in this study in terms of Performance Dimensions. The activity's instructional challenges include the complexity of the knowledge, the time pressure of the performance, and retaining students' focus on the goals.

There is also a "greater purpose" to this study. It is hoped that it will go some little way toward establishing the value of traditional martial arts as a research field. The hope is that this would contribute to a more diverse participation and greater general understanding of the value of the activity, and also to encourage some of the research-trained participants in various martial arts to turn a new eye toward something they already value personally.

Philosophy and Greater Purpose

According to Lipshitz, Klein, Orasanu, and Salas (2001), most [Naturalistic Decision Making] research is conducted in the field, drawing on methods from anthropology, ethnography, cognitive science, and discourse analysis. Efforts typically begin with descriptions of the phenomena, without prejudging what is or should be important to study. Descriptive approaches allow the researcher to examine phenomena in their natural contexts rather than leaping to premature attempts to narrow the focus and to test hypotheses. While field methods dominate, other methods may be used, such as simulation and laboratory techniques. (p. 343)

A motivating belief behind this study is that science is organized curiosity, that much of what is valued in scientific knowledge began as observation and description of some phenomenon, followed by a struggle to predict or even influence it. Rajan (1999) and Fischman (2003) both wrote that description is good and fundamental science, a necessary precursor to theory that produces testable hypotheses. Despite a seeming cultural bias toward leaving the earliest genesis of an idea out of the record, anecdotal evidence, or even an individual observation, frequently leads to valuable, creative research.

Though traditional theory-based hypothesis testing is the norm for the bulk of published research, it cannot be done effectively in an uncharted area. This study first explored the use of a instructional procedure traditional to certain martial arts styles, relying on expertise from a lengthy teaching career. The first phase was completed with an integrative literature review, finding in several research specialties work that is related to various aspects of the activity to be studied. Based upon performance data collected on a cohort of students, the second phase was to develop a method of interpreting the data which would concretely describe and enable measurement of one or more of the benefits of the activity to help explain its long retention in real-world training.

As there appears to be no literature directly related to the instructional procedure studied here, literature from diverse fields is referenced in discussing various aspects of the observed activity. Indicated is a need common to all these fields, as well as resources they could share. It can be supposed from this commonality that any contribution toward solution of this common need would be of interest to each of these diverse fields. An appropriate phrase heard at a recent conference was “information foraging”, a process which starts with targeted serendipity (such as is the case here, finding the first keyword which led to important areas of literature outside of mainstream education and psychology), and progresses to a more focused search, guided by the first discoveries and their effect of narrowing the definition of the topic.

Naturalistic approaches to psychological phenomena (e.g. Klein, Orasanu, Calderwood, & Zsambook, 1993, and Endsley & Bolstad, 1994) work to model decision making and situation awareness in real-world, as contrasted with theoretically-conceived, activities. Klein’s research, which produced the Recognition-Primed Decision (RPD) model, started from the question that since expert decision-makers are obviously better than current research says they should be, and refuse to use support tools and procedures based on current theory, what do they do naturally? Endsley and Smith (1996) started her work in situation awareness, which generated the SAGAT testing procedure, along the lines of wondering how is attention distributed in an complex, time-constrained, high-risk environment? Roscoe’s (1993) studies in situation awareness, which developed the WOMBAT testing system, began with the problem of what needs to be tested to improve prediction of pilot effectiveness before investing in their training. Schvaneveldt's (1990) Pathfinder algorithm, which has found limited but dedicated use, emerged from wondering how network models, which abound in cognitive theory, can be generated as representations of empirical data. Battig (1972), whose work was completely ignored by the education field only to be picked up and advanced by the motor-learning field, started from the observation that it makes sense that situations which required harder work in particular ways increase learning. Kandel (2001) in effect asked—what smaller system would reveal the same interactivity and emergent function as the human nervous system, but on a scale that would permit full description?'

Research Questions

A theory-based study has much of its territory already mapped out and is inquiring about one specific detail of terrain. Its final questions can be stated at the beginning. An exploratory (descriptive) study must start with very general questions, and develop more specific ones as it proceeds. The following global research questions were the starting points:

1. Why does this instructional procedure remain in use?
2. What unique or unusual effects might it be producing?
3. What unique or unusual insights into human performance might it provide?

Design of the Dissertation

This dissertation is organized into six chapters. Chapter I, Introduction, orients the reader briefly to the field of study. It also provides problem and purpose statements, research questions and definition of terms. Chapter II, Relevant Literature, draws upon and integrates several distinct lines of literature. Chapter III, Model and Observed Activity, is necessary as both may be quite alien to many readers. Chapter IV, Method Development, describes the trials, obstacles and breakthroughs found in a naturalistic inquiry. Chapter V, Demonstration: Results and Discussion, delineates analyses and appraisals of data produced in the activity. Chapter VI, Summary, Challenges and Potentials, relates the conclusions reached in Chapter V back to the literature in Chapter II, and discusses limitations encountered, challenges unsolved, and directions for further research.

CHAPTER II

REVIEW OF LITERATURE

Studying a real-world procedure is different from designing an experiment to isolate and observe a single mechanism. An effective procedure which has already endured several human generations has taken on aspects of a living thing—it competes successfully because it resembles actual application of knowledge in the field more closely than more commonly used procedures which are simpler. No single field is going to cover its significant aspects. For this reason, the approach to scientific literature in this study will ignore traditional boundaries, generate its own center, and suggest connections beyond itself.

The following literature was selected and organized to set a context to support the validity of the instructional activity observed and the means devised to observe and measure it. Incorporated are several research areas in which findings agree with expectations for the observed activity, and in which questions exist to which this work might contribute.

Few readers will be familiar with all these areas, but most will be familiar with one to three. As integration must succeed for this chapter to fulfill its purpose, each area will be introduced individually and linked to the ones before. Hundreds of works (and, no doubt, other specialties) not referenced here can be suggested by readers, but these selections were perceived as relevant at the researcher's current level of training, and should suffice to make the necessary points.

The seven areas of literature in this review build on each other to develop an engineering approach to teaching the management of cognitive load in a high-speed task-space of high cognitive and motor complexity. As the activity studied is being observed through its motor component, motor learning is the most obvious area to begin with.

Motor Learning

The motor learning field is feeling an impetus to increase study of complex or ecologically valid tasks, those that more closely approximate real-world skills. Wulf and Shea (2002) wrote that the field should include “skills that require the control of several degrees of freedom, load or overload the perceptual, cognitive and/or attention systems, have multiple emphases that need to be integrated, and/or perhaps even involve whole-body movements or real-world skills” (p. 205-6). But they do not support naturalistic methods or settings, agreeing with Banaji and Crowder that “the more complex a phenomenon, the greater the need to study it under controlled conditions, and the less it ought to be studied in its natural complexity” (p. 205).

<u>Motor Expertise and Closed/Open Skills Approach:</u> Starkes and Allard (1993) distinguish what they call “the expertise approach” from previous approaches to motor learning and control, including closed loop theory, schema theory, dynamical systems theory, or direct perception, each of which they say started with a theory, moving to predictions and laboratory tests. They contend that early theories were too cognitive, and more recent ones are not sufficiently cognitive. They would like a theory to describe expert performance cohesively with early-learning performance. They point out the general lack of data on established motor experts and the unevenness in what is called an expert across studies.

Allard and Starkes (1991) contrasted closed and open skills. Two main differences are that closed skills are performed solo in an unchanging environment, so that skill is measured in the perfection of exhibiting the motor pattern itself. Open skills have competitors or opponents (and thus a changing environment), so that skill is measured by the impact upon that environment (outcome generated) by the motor pattern. Closed skill in 11-year-old dancers showed a primacy effect similar to verbal memory, but no recency effect, conjectured as due to not needing to start a sequence in the middle. They cite a 1956 study by Laban in which professional modern dancers learned structured and random movement sequences with equal facility, clearly not needing the structure needed by the comparison group of novices.

The examples Allard and Starkes gave in 1993 of studies related to open skills were done quite differently from the closed skill studies. Instead of directly observing performance as was done in closed skill studies, these studies were all done by reconstruction of player arrangements from photographs, or card-sorts of photographs or diagrams. Though all showed significant results, it is reasonable to question how well these activities represent actual play. Recognition is not recall, nor is it effective performance. Neither does the time pressure and uncertainty of play exist. According to their own definition, these studies show no motor pattern, and no situational outcome.

Allard and Starkes point to a difference in cognitive strategy, under speed stress. Open skill activities emphasize search, and closed skill activities emphasize memory. They do not accept condition-action links as in production-system models of skill (i.e. Anderson) in motor skills, especially open skills, because of the disadvantage of predictability to an opponent. They use the examples of ‘practice players’ who excel in practice, but do not perform well in competition, and to wall practice leaving tennis players unprepared for play to set the stage for attempts to separate ‘knowing’ and ‘doing’ in an experimental situation.

They set the stage for separating ‘knowing’ from ‘doing’ in an experimental situation with the examples of ‘practice players’ who excel in practice, but do not perform well in competition, and of wall practice leaving tennis players unprepared for game play.

It could be argued first that knowledge of the next circumstance to occur is not equivalent to a change in actual knowledge about the field. Also, there may be a relationship to augmented feedback work, in that too much external feedback can actually interfere with growing reliance on internal feedback, showing that there is, indeed, considerable control. From these experiments, Allard and Starkes infer a primary importance in the linkage between knowing and doing as separate sets of knowledge.

Allard and Starkes emphasize that the human system performs much more easily in terms of producing a consistent environmental outcome than in terms of consistent production of exact motor actions. Allard (1993) describes this as an example of the direct connection between perception and action according to the ecological psychology perspective on motor control—that “variability across trials in the performance of experts is often functionally linked to accomplishing the goal of the action” (p. 21).

<u>Contextual Interference and Practice Schedule Approach:</u> The non-intuitive term ‘contextual interference’ (CI) refers to the challenge or cognitive load presented by a task or sequence of tasks. It was selected as a label for the effect observed by Battig. The term ‘practice schedules’ refers to patterns of ordering tasks during learning. In a block practice schedule, all the repetitions of one task are completed before moving on to another task. In a random practice schedule, repetitions of two or more tasks are randomized so that the next task is always unpredictable. In a sequential practice schedule, each task is different from the one before, but the sequence is repeated, and therefore predictable. Practice schedules are

used to structure the activities observed in CI research, such that random practice generates more CI than sequential practice, which generates more than block practice. Individual skills are learned most quickly in block practice, but overall performance gains, especially under random demand, and transfer are higher with random practice.

Schild and Battig established this direction of research in the specialty of verbal learning in the 1960's. Through 1979, Battig referred to the effect as the 'intratask interference principle'. Battig (1972) said that whenever learning was effortful, the participant was "likely to invoke any additional learning processes that he can" (p. 155), making the learning more resistant to forgetting. He emphasized that "individual Ss may differ substantially and often idiosyncratically as to how they try to overcome intratask interference" (p. 155). By the time he counseled Shea and Morgan (1979) in the design of their landmark experiment, Battig had switched to the term 'contextual interference' so as to include interference both within the task (intra-task) and between tasks (inter-task). It is important that Shea and Morgan claimed explicitly that results reflect cognitive and perceptual processes, not just the motor component of performance. Individualized performance optimization through engaging multiple processing strategies during acquisition is an important aspect of learning.

Research in this area appears to have advanced along two lines. A major one focusing on hypothesized mechanisms for the CI effect is not relevant to this study. The other, focusing on field research demonstrating the CI effect in various activities, is highly relevant in terms of activity selection, measurement strategy, and instrument design.

Experimental design in CI research, especially in field studies, has been a hotly debated issue, particularly in terms of amount of complexity and cognitive load to design in.

Brady (1998) reported that Hebert criticized field experimentation for having exactly what Shea and Wulf (ibid.) said was needed in future research: more complexity in the skills. Similar to Allard and Starkes' (ibid.) concern about level of cognition involved in early theories, Brady's review of research on the CI effect reported a common criticism from motor control researchers that many early experiments had too little motor demand, and were too heavily loaded cognitively. On the other hand, Brady wrote that Del Rey and colleagues believed open-skill athletics make considerable use of cognitive processing, agreeing thereby with Allard and Starkes. Brady also reported general problems of connecting field results to strong laboratory results such as Shea and Morgan's. The difficulty with obtaining field evidence has been attributed generally to insufficient sensitivity in measurement, and occasionally to inappropriate instrument selection. It is quite possible that the activities selected, such as fundamental volleyball and badminton skills, are so simple that they place too little cognitive load to induce measurable effect. Additionally, physical education classes may present convenient samples, but are notoriously casual, and seem unreliable in evoking the level of engagement needed.

Several points made by Lee, Swinnen, and Serrien (1994) will prove very useful as other literature areas are incorporated. One is the phrase "contrastive knowledge" (p.338), which concisely ties random practice to schema construction, in that each situation that differs from the previous one establishes a schema recording its differences or contrasts. Also, without mentioning reconstructive memory, they say to "approach a task by developing a strategy or plan of action" (p. 338). This allows that the next time the task is approached, a different plan may lead to the same action. This would construct alternative linkages and probably support use of the memory in novel (transfer) situations.

Practice schedules describe the activity of the system in broad terms, but not in sufficient detail to support detailed manipulation. Contextual interference is one way of viewing the effort required by the cognitive processing, but has not yet proven itself in field research.

Cognitive Psychology

Two topics in cognitive psychology bear direct relevance to this study. The first topic, Cognitive Load Theory, contributes valuable detail to what was just discussed under contextual interference, supporting an engineering approach to the discussion of managing cognitive effort. The second topic, a set of basic mechanics including schema theory and proceduralization/composition, will likewise extend the previous discussion of practice schedules toward an engineering approach.

Cognitive Load Theory: Cognitive Load Theory has been developed by Sweller, van Merrienboer, and Paas (1998) to encourage the structuring of instructional design around efficient utilization of working memory (WM). They propose a construct "Cognitive Load" (CL) to quantify the burden learning places on working memory, which they develop in a manner that strongly resembles engineering. They define 3 types of CL: intrinsic (ICL), extraneous (ECL) and germane (GCL). Being intrinsic to the nature of the material, ICL is difficult to reduce. ECL is to be avoided. As it is determined by the instructional design and therefore independent of the material itself, ECL can be significantly reduced. The basic goal is that ICL + ECL should not exceed WM. The concept of GCL came about later in the theory's development. If ECL is reduced enough that WM capacity is available, GCL, which involves or supports "construction and mindful abstraction of schemas" (p. 264) is to be incorporated through elaboration by introduction of related material. Stated algebraically, ICL + ECL + GCL $\leq$ WM.

They agree with traditional expertise research in chess that human cognition is better adapted to memory of huge numbers of situations and linked actions (even very complex ones) than to complex chains of reasoning. They consider learning and problem-solving as incompatible processes, because holding current, and goal states as well as operators, displaces schema construction based on developing recognition of problem states and associated actions. The burden is increased by sub-goals in a multiple-step problem, and by searching for useful examples in the instructional material or elsewhere. They say that element interactivity contributes substantially to CL. Sweller et al. use learning of vocabulary as an example, saying that each element can be learned in isolation, and therefore with very low ICL. However, this would seem to imply that the immense interactivity actually available between vocabulary elements in a language (which some people greatly enjoy) would be considered GCL.

Their development of CL measurement is very detailed, including the definition of 7 distinct effects. Performance measures of CL began simply with differential learning times and error counts. Interaction with individual differences can be incorporated easily into the Cognitive Load Theory approach, in that high-ability students will sometimes do additional processing and even elaboration of schema. In contrast, those of limited capacity who would benefit most from schema-formation support may ignore supplementary material until they get overloaded and into trouble. Incorporation of GCL is recommended in terms of elaboration and/or associated questions. This is a notable similarity to the benefits of observation with augmented feedback in motor learning research (e.g. Lee et al., 1994).

Sweller et al. recommend that teaching material blend gradually through examples and problems of increasing cognitive load. This is very similar to the recommendation by motor learning researchers that CI be incremented upward during practice to remain near the learner's capacity. Construction of a more efficient knowledge base is noted to be very difficult, especially as the individual student would be the deciding factor in how much material would help schema construction (GCL), or become a counter-productive overload (ECL). It requires just the sort of detailed engineering Sweller et al. describe as being more efficient for learning that supports transfer. Their advice to integrate presentations is identical to what Tufte (1983) calls good information design, with the same effect of greatly increased intelligibility and lower levels of ECL.

Unfortunately, Sweller et al. did not support their discussion by integrating the substantial CI literature already available in 1998. The fact is striking that they basically restate Battig, even use his term 'contextual interference' (p.287), and structure their ICL and ECL so similarly to his conception of intrinsic and extrinsic interference without referencing him, as is their work in vocabulary learning. Even in his 1999 book, expanding on Sweller et al., Sweller (1999) still took no advantage of the fast-growing motor learning literature, even Battig's original work. Battig struggled for years to get attention for the effect in classical education, before the motor learning field adopted the idea. Wulf and Shea (2002) and Magill and Hall (1990) both reference papers by Jelsma, who is one of Sweller's co-authors. Studies like those by Wright, Li, and Whitacre (1992), also reported in Brady (1998), and that of Wright, Snowden and Willoughby (reported in Lee et al., 1994), would have been apt support, especially in their use of verbal elaboration, and Wright et al.'s evidence that the change was in reaction time rather than in motor time.

Studies discussed by Sweller et al. provide long-expected support for CI in academic areas, in a cohesive theoretical discussion. Dimensions that become more accessible with academic tasks than with physical tasks include manner of task presentation, saliency of defining characteristics, and task context and familiarity. In one study, variability was generated by the surface story independent of the problem content (versus correlated with the content) so that one group could track the problem by the story while the other group had to see through the story to the problem. Those participants who were forced to examine more deeply categorized better. Sweller et al.'s distinction between ECL and GCL is a major contribution to the CI field in accounting for the paradoxical nature of increased load improving retention and especially transfer. As they explain, "Variability over problem situations is expected to encourage learners to develop schemas, because it increases the probability that similar features can be identified and that relevant features can be distinguished from irrelevant ones" (p.286). They also make the valuable distinction between task-based load, which encompasses all of the above, and learner-based load, which involves the capacities or resources allocated, and encompasses individual differences.

<u>Proceduralization/Composition and Propositional Modeling:</u> As described by Gagne, Yekovich and Yekovich (1993), actions can be represented as condition-action rules, or productions. They describe compilation of these small productions into larger ones as consisting of composition and proceduralization. Composition is described as aligning end conditions of one production with compatible pre-conditions of the next. The matching of conditions in the two schemas to be linked suggests illustrative metaphors from engineering and/or chemistry. Proceduralization is described as the dropping out of cues and connections to declarative knowledge, leading to automation. Allard and Starkes (1991) suggest that

automation would be a disadvantage in competition, by generating predictability. However, if the automated actions occur too fast for the opponent to effectively interrupt them, they can succeed—in fact, automation is very important for achieving this short-burst speed.

Clamann, Wright and Kaber (2002) demonstrated discrimination between types of function based on compatibility with automated support systems. In a workload analysis of air traffic controllers, they showed that computerized automated assistance was beneficial in lower-level sensory and psychomotor functions, such as information acquisition and action implementation (first and last links in the chain), but not in cognitive tasks, such as analysis and decision making (intermediate links). Cognitive automation within the person could be argued similarly, so that activity would progress in a run/poll cycle, similar to the most basic computer program, waiting for executive input to trigger the next production.

Propositional modeling, as proposed by Kintsch (Gagne et al., 1993) deals with the memory of meaning as opposed to memory of detail. It resembles the use of productions described above, in that small elements are represented, as well as larger constructions of them. A propositional network gets constructed through elaboration (Anderson, 1995) adding new propositions, which can become lost if not linked in, and difficult to access if linked in at only one or two points. This suggests that investment in more linkage makes access to a specific proposition available from more than one direction, so it is likely to be thought of under more circumstances. Spread of activation represents this ‘thinking of’ or accessing neighboring propositions by following a link-trail. The similarities between composition of productions and linking of propositions are impossible to miss.

Cognitive Load Theory provides a highly detailed method of discussing the engineering aspects of working within cognitive capacity. It is giving evidence of great value in the realm of academic learning, but has not been extended into more real-world task realms. Cognitive psychology fundamentals provide tools for describing the basic structure in a practice schedule.

Knowledge Organization

If practice schedules describe the activity in the system, networks describe the terrain. Knowledge organization provides an opportunity to put the theoretical structures of productions and propositions into application in real-world areas.

Canon-Bowers wrote: “One of the markers of expertise seems to be the organization of knowledge” (Cannon-Bowers & Bell, 1997, p. 103). Similarly, McGarry wrote: “The existence of structure in sport competition is implicated in the widespread practice of using the information gathered from a past contest to prepare for a future contest” (McGarry et al., 2002, p. 771).

As psychology has moved more from behaviorist models to cognitive and connectionist models, and as the symbiosis of psychology and computer science has flourished, with the development of hypertext, neural networks and fuzzy logic, interest in how ideas connect and relate has intensified. As described in the previous section, many efforts have been made to map concepts in a way to compare and improve structures, predict school or job performance, and describe and develop expertise.

Numerous systems of conceptual mapping or so-called “structural knowledge” are in use (Friendly, 1977; Jonassen, Beissner, & Yacci, 1993). One of these, network modeling, is well represented by a published tool called Pathfinder (Schvaneveldt, 1990). Pathfinder generates network representations of sets of concepts, with no particular dimensionality; the nodes representing concepts are located only in terms of proximity (representing relatedness) to each other. These proximities or connections are represented by the network's links. The person whose conceptual structure of a subject is being mapped inputs Likert-scaled values for the relatedness of each pair of concepts. The network begins fully connected. Less-related links are pruned away, based on calculations from the relatedness values input. This pruning actually reflects the neuronal pruning in the brain (Taber, 1999). The least-connected (most-pruned) settings in the software are said to create the most useful diagrams. This technique converts raw relatedness data into a form richer in information and predictivity (Johnson, Goldsmith, & Teague, 1994). Current utilizations include predicting school or job performance, evaluation and redesign of instructional materials, and measurement of expertise. Findings in the evaluation area include Britton and Tidwell's (1991, 1995) successful increase of student correlations to experts after specific changes in a text. Wyman and Randel (1998) refer to Britton and Tidwell's study in reporting that their comparisons between high and low performance job groups indicate differences in conceptual structure. Both studies also agree that though correlational differences between experts may seem small, they can be extremely important. All the studies consider selection of experts a critical question.

Wyman and Randel made several points about Pathfinder that are very consistent with Sweller et al. A well-ordered knowledge structure would contribute substantially to an instructional designer's efforts to eliminate ECL in favor of GCL. Focusing clearly enough on key concepts and critical connections, which is what Pathfinder is designed to support, may even allow a tighter definition of what true ICL is in a given field.

Centrality is put forth by Wyman and Randel as correlating highly both with performance scores, and with experience and degree of reference in comparing two experts. Centrality is measured quite simply by maximal links to a single concept. In the subjects, centrality correlated positively with performance on a test. Expert I correlated positively with this centrality effect (and had actual field experience in his profession) and expert II correlated negatively (and had only schooling, with no field experience).

Like Cognitive Load Theory, network modeling as developed in Pathfinder is extensively enough developed to offer engineering level manipulation of its area. Valuable concepts including links, proximity, centrality and pruning are contributed.

Naturalistic Decision Making (NDM)

A strong fellow-feeling developed for NDM during the course of this study, because like Battig's work in contextual interference, it starts from observation of real behavior. Also like Battig's work, it has had to find itself a home outside of the field from which it arose.

NDM is a significant new paradigm in decision making research. It began as a departure from traditional research that has long found human decision making inferior to rational decision models such as Bayesian analysis or multi-attribute utility theory. It is NDM's position that decades of laboratory experiments have been unable to explain the success of experts making fast, ill-informed, complex decisions in their field of expertise because these experiments specifically exclude factors such as context cues and past experience, which are essential to real-world decision making.

Under the older normative paradigm so many real-world factors were removed in order to design rigorous controllable laboratory-based experiments in decision making that experts appeared to perform no better than novices (Serfaty, MacMillan, Entin, & Entin, 1997; Serfaty & Michel, 1990). Lipshitz and Strauss (1996) point out that the standard RQP (Reduce, Quantify, Plug-in) heuristic is problematic in reducing the uncertainty that comes with poor information, in quantifying and in plugging into formal schemes probability values that are incalculable. These difficulties handicap the use of support systems reliant upon such quantification. The old approach ignored what military commanders actually needed and did, or wanted to do, in favor of theoretical norms, requiring them to develop three courses of action, weigh them in terms of Bayesian analysis or multi-attribute utility theory, and enact the winner. This was never done in actual practice, due to time pressure and changing unreliable information, but was “pervasive in the support systems and training products that resulted from these efforts” (Drillings & Serfaty, 1997, p. 73), creating a serious discord between the generation of training and tools, and actual practice.

In the 1980’s, various new paradigms were proposed, but few involved the expertise of the individual commander, which historically has been so crucial in the real world. Actual research at this time showed little difference between results of single- and multiple-option decisions due to ‘hedging’ behavior, and that “…experts apparently recognized prototypical situations using feature cues…” (Drillings & Serfaty, 1997, p. 74), the very cues the old research was removing from its testing. Drillings and Serfaty refer to Daniel’s 1979 work demonstrating individual expertise as being at least as important as quantity of information, in that “the best players made as good decisions with 20% of the ground truth as did the worst players with 80%…” (p. 78). Endsley had similar findings in her 1995 and 1996 work (Endsley, 1995a; Endsley & Smith, 1996).

Klein’s development of the Recognition-Primed Decision model (RPD) began during observation and description of the behavior of fire-ground commanders (FGC’s) at fires (Klein, 1997). The FGC’s were observed in more than 30 incidents averaging 5 high-risk, non-trivial decisions each, with time pressure, unreliable information, changing goals, and changing situations, in total contrast with traditional laboratory studies. These experts reported making no comparisons at all, relying rather on similarity to prior situations, and generating a workable first option, fixing or replacing the option only if mental simulation demonstrated significant flaws. Clearly, real-world conditions lead to processes not seen in, and therefore not related to, traditional laboratory findings. This collision between work based on irrelevant theory and real-world observation would be instantly recognized by Battig, who never did get educational research in the 1960’s and 1970’s to consider the idea that working harder would generate stronger learning.

The RPD model (well-described in (Klein, 1997)) has been tested in activities ranging from anesthesiology, nuclear power plant operation, software design, offshore drilling, and jury deliberations to highway design. This diversity has been somewhat masked by the interest and support quickly shown by the military services, who funded a substantial amount of early research and fully sponsored the first two NDM conferences.

The development of the NDM approach, as shown by its 6 world conferences, is noteworthy. The first conference (Orasanu & Connolly, 1993) in 1989, established as points of interest 8 factors complicating decision processes:

1. Ill-structured problems
2. Uncertain dynamic environments
3. Shifting, ill-defined, or competing goals

4. Action/feedback loops

5. Time stress

6. High stakes

7. Multiple players

8. Organizational goals and norms. (p. 7)

By the second conference (Zsambok, 1997) in 1994, the list above had become one aspect of NDM, to which three other aspects had been added: study of expert rather than naïve subjects, study of actual decisions in context-rich settings, and consideration of a broader view of the decision episode, to include situation awareness as well as option selection.

The fourth conference (Salas & Klein, 2001) in 1998, strengthened the focus on expertise, and showcased the dramatic expansion of application fields beginning to adopt NDM approaches to some of their research. This diversification is important to prevent an unrealistic assumption that only military persons and organizations are subject to the conditions listed above.

Brief consideration of the 8-factor list above will show that uncertainty is the most general problem, in that most of the other factors generate it. Thus, uncertainty, with its causes and coping strategies, is a major topic in NDM. Its confounding impact through individuals' interpretations of the problem to be solved causes its intentional removal in most laboratory research, through simplification and thorough explanation of the problem. This avoidance is understandable, given the challenge of incorporating uncertainty into a model.

Schmitt and Klein (1996) condense uncertainty into a matrix in which missing, unreliable, ambiguous or conflicting, and complex information are mapped as causes, and data, knowledge and understanding are levels (similar to Endsley's 3 levels of SA, see next section). They warn that hopes of achieving 'information dominance' over an enemy focus on the data level. Like the earlier problems with support systems requiring procedures decision makers do not use, focus on this type of dominance ignores the higher-level interpretation used by successful commanders. Their work indicates that commanders achieve most of their understanding of a situation from very few cues, often 6 or fewer, filtering out most of the data presented. One interpretation of this work is that an expert model of knowledge supporting pattern recognition is crucial to uncertainty management. They list several major events of this century (9-11 could be added) in which adequate data were received but no warning was interpreted, in part because information can reasonably be interpreted in many ways.

Endsley and Smith (1996) used a replacement task and a decision task modeled intentionally upon classic studies of decion making in chess by DeGroot (1965) and Chase and Simon (1973) and found support of an NDM model in the total lack of consideration of multiple alternatives. Endsley also found that individual subjects were significant factors on all measures (seemingly based on combat experience), especially in a tendency to plan several moves ahead of time. Also consistent is Wyman and Randel's result (ibid.) that the stronger expert instructor was the one who had combat experience.

Canon-Bowers and Bell (1997) condensed a conference panel discussion that cited a lack of empirical investigations of effectiveness of NDM-generated training. A question they posed was "how do we train people to be flexible, or to engage in decision making that does not follow a set of predetermined steps?" (p. 100). They recommended three steps to create this training: "use NDM theory to establish a set of characteristics that describe expert decision makers", examine the processes, knowledge, and skills that NDM theories suggest…" (p. 101), and use this information to design methods, strategies, and content. This desire for un-predetermined decision processes is similar to what Battig called invoking any processes possible.

A major point contributed by NDM is the argument that psychological research cannot approach real-world problems without incorporating options, uncertainty, real stakes, and change. It also strongly argues that starting from theory without first observing what naturally occurs, and trying to describe it, is fruitless.

Situation Awareness (SA)

Like NDM, SA arose from dissatisfaction with traditional research and theory in a field, for its failures in describing or predicting real-world processes, due to a preponderance of bottom-up development. New testing was needed which accessed and assessed real human function with its high adaptability, flexibility, and complexity.

This research area is evolving interactively with that of NDM, as shown by its inclusion in NDM's 1994 conference (Venturino, 1997), and those since. What is now called situation awareness, pilots used to call "'staying ahead of the aircraft', 'good judgment', and 'airmanship'" (Roscoe, 1993) (p. 48), and researchers used to call 'residual attention' (Damos in O'Hare, 1997).The importance of SA is vivid in terms of the consequences of pilot error: "…while in close combat, many pilots report that they are interested only in where their opponent is. Too frequently, however, though they are successful in avoiding enemy missiles, they end up flying into the ground…" (Endsley, 1995b, p. 38).

Roscoe and WOMBAT: Roscoe's (1997) definition of SA derived from his work during WWII, since. Roscoe defined SA as the overarching ability to:

- attend to multiple information sources,
- evaluate alternatives,
- establish priorities,
- estimate probable outcomes for different courses of action,
- work on whatever has the highest momentary urgency, without losing sight of the routine,
- reorder priorities as situations deteriorate or improve, act decisively in the face of indecision by others (p. 11).

A failure was seen in traditional test batteries in accounting for only around 25% of the variance in training success in aviation, and no correlation seen with post-training performance (Roscoe, 1997). As the human and financial costs of errors among key personnel continued to grow, and as system automation increased the challenges of tracking and deciding to intervene, a more effective screening method was needed. For this reason, Roscoe and Corl developed the WOMBAT Situational Awareness and Stress Tolerance Test from Roscoe's definition of SA.

In their view, the notion that performance depends on a collection of simple abilities had to be replaced. The complexity of what was being predicted had to be reflected in a measuring instrument of comparable complexity. At the same time, confounds from experience with specific equipment or job background had to be eliminated. An adaptive scenario was developed which presents both high-pressure and monotony as well as multiple information sources. Multiple response alternatives must be tracked, evaluated and integrated. WOMBAT was developed to run several different kinds of tasks at the same time, with constantly changing priorities on the tasks adding executive-level demand, the focus of the test. WOMBAT's individual tasks are: target tracking with two joysticks, figure rotation, quadrant location, and digit cancelling. Each task can be learned quickly. The critical element in the test is that they all run concurrently, with constantly changing point-values, so that attention management is what gets rewarded.

O'Hare (1997) compared elite pilots, average pilots and non-pilots, concluding that "the WOMBAT test measures some ability over and above those measured by the battery of component ability measures" (p. 549) against which it was compared. Among four sections selected from the Walter Reed Performance Battery, the only one to correlate significantly (r(24)=.59) with WOMBAT performance was Pattern Recognition. General results were that elite pilots obtained clearly higher WOMBAT scores than did average pilots, who obtained higher scores than matched non-pilot controls. O'Hare distinguished between different circumstances of testing, in pointing out that "…the domains in which expertise has traditionally been studied—chess, music, physics, and so forth—may emphasize *focused rather than divided* attention. In real-world domains requiring high levels of SA, attention management skills may be pivotal to successful performance" (p. 551, emphasis added). O'Hare cited a 1972 study by Damos, in which measures of residual attention increasingly correlated with performance in flight training over the first 30 hours. It also matched his own results in that students in the professional program outscored students in the basic program.

Like NDM, applicability of SA extends well beyond either military or flight personnel. Roscoe (Klein, 1997) pointed out the frequency with which paramedic trainees may learn all the skills and knowledge of the field, and still be unable to apply them under the stress of a real disaster.

Endsley and SAGAT: Whereas Roscoe's work came from a background in psychology, Endsley's work came from human-factor-based information system design work. Her definition is quite different from Roscoe's, and produced a very different testing instrument, in that it specifically eliminates decision and action. Her original SA definition was: "perception of the elements of the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1995a, p. 65). She (1995b) later structured this definition much more fully as:

- Level 1: Perception of the Elements in the Environment,
- Level 2: Comprehension of the situation …based on a synthesis of disjointed Level 1 elements …in light of pertinent operator goals,
- Level 3: Project the future actions of the elements…at least in the very near term. (p.36-37)

As Roscoe used his SA definition to develop WOMBAT, Endsley used her definition, which specifically excludes decision-making as a function separate from awareness, to develop SAGAT. This is a data-collection and interviewing technique (later computerized) utilizing stop-action simulations. The initial testing of SAGAT (Endsley, 1995a) indicated that situational information in memory far outlasts traditional limits of short-term memory, being available for 5-6 minutes. This supports the importance to the subject (salience) as a source of the contrast, and the possibility of long-term storage prior to short-term "highlighting".

Endsley and Bolstad (1994) compared experienced combat pilots' scores on SAGAT with their scores on a battery of tests of basic skills hypothesized as important to SA. Consistent individual differences appeared, including a point of potential importance to this study. A high correlation with SAGAT scores was found, not with the tracking task itself, but with the level of difficulty reached in the tracking task. They comment, "…it is not immediately obvious why a psychomotor task should be related to SA…" (p. 257).

Jones and Endsley (cited in Endsley & Robertson, 2000) analyzed a report of SA errors according to Endsley's 3 levels, resulting in 76.3% at Level-1, 20.3% at Level-2, and the remaining 3.4% at Level-3. Their error analysis appears to pose a contrast to Schmitt and Klein's warning against a reliance on "information dominance" in that ¾ of the failures were interpreted as being generated at the perception level. Resolution may be found in the sub-headings of the analysis indicating that by perception they mean perception of information already filtered for relevance. They note that inexperienced pilots are less able to assess timing, risk, capabilities, consequences and severity, and tend to focus on Level 1 rather than Level 2; this agrees with Schmitt and Klein's scaling of experience in command roles. Their list of higher-order cognitive skills constituent to SA is very similar to that given for macrocognition (Klein, Klein, & Klein, 2000; Woods, 2002).

O'Hare (1997) discussed Space Fortress, a computer activity designed to challenge a number of abilities. Eight different secondary tasks combined with the game indicated that major aspects of performance were response accuracy and timing, and monitoring multiple information sources. Over 25 sessions, correlation of scores on Space Fortress with other games decreased, while correlation with IQ increased. O'Hare suggests that both involve mastering a complex system of rules, attend selectively, manage working memory, and make quick accurate predictions.

Observing individual differences is to be expected with the direct study of expertise. NDM research opened a whole range of human decision performance to measurement, not considered by the older normative viewpoint discussed earlier. This range has no visible correlation with rank or seniority (Pascual & Henderson, 1997), but significant correlation with factors such as: familiarity of situation, type of experience, and decision making / information-processing strategies used (Endsley & Smith, 1996; Pascual & Henderson, 1997) (p.242). Lipshitz and ben Shaul (1997) found that experts 'read' the situation more accurately, engage in more efficient information search, and (seemingly as a result) make fewer bad decisions. They also noted more frequent and elaborate communication by experts than nonexperts, showing that "experts are more likely to consider other players' perspectives in their decision making" (p. 296). They argue that differences in expert and novice construction of mental models, as well as the function of schemata in the feedback loops of decision making, are crucial to understanding the processes. Serfaty, MacMillan, Entin and Entin (1997) agree that "…the mental model allows the expert to act under uncertainty" (p. 243), and suggest that "an expert commander has a mental model of the tactical situation that differs in measurable ways from that of a novice" (p. 235). Barber (1999), an Army Ranger, says quite simply, "My M-16 is zeroed a little differently than yours" (p. 1).

Both Endsley's and Roscoe's definitions came about in researching the problems of combat pilots. Watts (1996), a retired Air Force fighter pilot himself, provided a clear, but colloquial, definition of SA: "…the ability of opposing aircrews to develop and sustain accurate representations of where all the participants in or near the air combat arena are, what they are doing, and where they are likely to be in the immediate future" (p. 94).

With its strong components of attention management, filtering and abstraction of complex input, and executive action, SA certainly advances the demand for clear definition of constructs, tests, and instructional procedures at a complex level.

Adult Learning

Mature adult cognition is contextualized, fitting the abstract to the concrete limitations of everyday life. Tennant and Pogson (1995) contrasted typical academic test problems with workplace, family and community problems. Referencing Sternberg, Neisser and Gardner, they say that selection of a problem (often ill-structured or unstated) out of a problem-space is a defining aspect of adulthood on which formal testing and traditional education fall short. This problem-selection or problem-definition corresponds to Endsley's Level-2 SA in its synthesizing of pattern from perceived elements and Level-3 SA in its projection of future actions, with the added impact of one's own goals, purposes and actions. The number of 'correct' answers in a real-world situation, resulting in part from this issue of problem selection, is also an important factor in both adult development and NDM literatures.

Whereas Schmitt and Klein (1996) wrote of tolerance of uncertainty, Tennant and Pogson (1995) wrote of the development of tolerance for contradiction and ambiguity. Both fields write of the importance of incomplete information and/or feedback. Tennant and Pogson could be writing a chapter in an NDM book when they say, "Ambiguity, poor feedback, unclear problem boundaries, the vagaries of the relationships we have with others, and many other factors all combine to constitute the very loose framework of our adult experience of intelligent action in the everyday world" (p. 33).

Likewise, O'Hare's (1997) interest in focused attention measured by traditional methods versus divided or shared attention, required in most real-life venues is deeply involved with selection of a problem from an ever-shifting multi-dimensional space. In the background of this attention must be a constant prioritization function, which is a major aspect of what WOMBAT was designed to measure.

It can be seen that many of the major points of the NDM and SA perspectives are also found in the literature of adult development and learning.

Neuroscience

Modeling of Complex Systems: In biology, a smaller system is used as a surrogate model to study processes that also occur in larger systems. Kandel (2001) selected the sea slug Aplysia as a model for study because it had few enough neurons and synapses to view and track completely, but enough to retain the emergent properties seen in much larger mammalian nervous systems. The roach's nervous system has so few that it is only suitable to model hard-wired functions like simple robotics (Beer, 1990).

Salience, Pattern Recognition, and Embodied Cognition: Klin, Jones, Shultz and Volkmar (2003) used eye tracking (also much used in ergonomic fields) to study inappropriate cueing and search-for-meaning disability in autism, in explicit versus naturalistic situations. It was clear that autistics could learn to cue correctly in explicitly directed laboratory situations, but still fail to independently use that learning in real settings. The attachment and constant adjustments of relative salience to situation elements is similar to Endsley's Level-2 SA, and hyperlexia is a disconnect between Level-1 and Level-2 SA. Schmitt and Klein (ibid.) would probably agree with what they refer to as the false hopes of "information dominance" without higher-level interpretation as a sort of "organizational autism". The conditionality of shifting priorities, based on the context-dependence of complex environments is an everyday variety of expert field-knowledge, foundational to NDM. All of these points are also harmonious with those made in the section on Adult Learning, tying that area together with prioritization as a constant supervisory function, as shown in WOMBAT.

Taber (personal communication) pointed out that relative salience and prioritization is related to pattern recognition. O'Hare (1997) found that pattern recognition was the only significant correlant between SA scores on WOMBAT and the Walter Reed Inventory, and Schmitt and Klein (ibid.) say pattern recognition is crucial to uncertainty management. 'Salience tracking' would be an appropriate term for what NDM calls 'attention management', incorporating changing salience from changes of situation and/or goals. Klin, et al.'s term "topology of salience" is quite evocative. Research efforts into team cognition, including shared situation awareness and shared/distributed decision making might benefit from a neurological view on the developmental aspects of meaning and shared meaning, as discussed by Klin, et al.. Another point of overlap is the extended memory of important (salient) points in SAGAT, reported by Endsley (1995a).

The indispensability of the motor aspect in this study's observed activity is supported by Klin, Jones, Shultz and Volkmar's (2003) embodied action approach, drawn from a neuroscience framework called 'embodied cognitive science' and its literature on the inseparability of perception and action in cognition as lived. They say: "These experiments illustrate the point that meaningful cognition of objects… cannot be formed by means of visual extraction alone; rather, there is a need for perceptual processes to be actively linked with action in order to guide further action upon these objects" (p. 352), and describe developmental skills as 'perception-for-action' systems. This would also support the high correlation of a psychomotor task with SA, found by Endsley and Bolstad (1994), described previously in the section on SA.

Klin et al. (2003) reference neuroimaging studies which have shown overlapping brain circuitry subserving action observation and action generation (p.352). This would relate to observational learning as described by Lee, Swinnen, and Serrien (1994), under motor learning.

It is impossible to escape consideration of the brain in relation to any cognitive function. But only recently available has been technology to enable direct concrete study of brain activity itself. It can be expected that grounded theory based on actual observation of behavior in 'field' or real-world conditions will interact favorably with research into the brain.

Summary

The most important single point to be drawn from this literature review is that several fields concerned with human performance in complex fields, both psychomotor and academic, would find value in an experimental and/or instructional paradigm which incorporates problem selection from a problem-space, coping with pressure in time, uncertainty, and shifting priorities, and in which situational control is taken and lost frequently.

CHAPTER III

DESCRIPTION OF OBSERVED ACTIVITY

In this dyadic activity, partners face each other, trading blocks and punches in a steady, continuous rhythm. Two terms are used here to distinguish different levels of specificity: a hand set is a movement sequence which can be performed with either hand, a drill is the left- or right-handed version of a hand set (i.e.: Hand Set-1 is expressed as Drill-1L and Drill-1R, for left and right). Partners switch unpredictably between 3 unique hand-sets. Each hand-set consists of a specific block and a specific punch, and can continue indefinitely. Thus the studied activity has a total of 6 drills.

This activity is learned incrementally. Students begin each new hand-set by learning basic skills for its two drills, working alone. Partnering begins as soon as manageable, because interaction with a partner is necessary to achieve the purpose of the activity. Figure 1a shows the fundamental partner interaction of each drill. Addition of a partner directs attention to new factors including distancing, timing, and balance. Partners change frequently, to incorporate all the individual differences of working with different people. It is important to interpretation of this study’s results to see the dynamic that emerges from the coordination of the two partners’ actions. The partners merge as a system that relies on both partners maintaining a dynamic stability in speed, distance, and timing. At this stage of instruction, partners practice single drills, as shown in Figure 1b. This level of integration is all that is provided in many instructional programs in the martial arts.

Transitions between drills are learned soon after block performance of two hand-sets has become stable. These transitions allow seamless switching among the drills. The

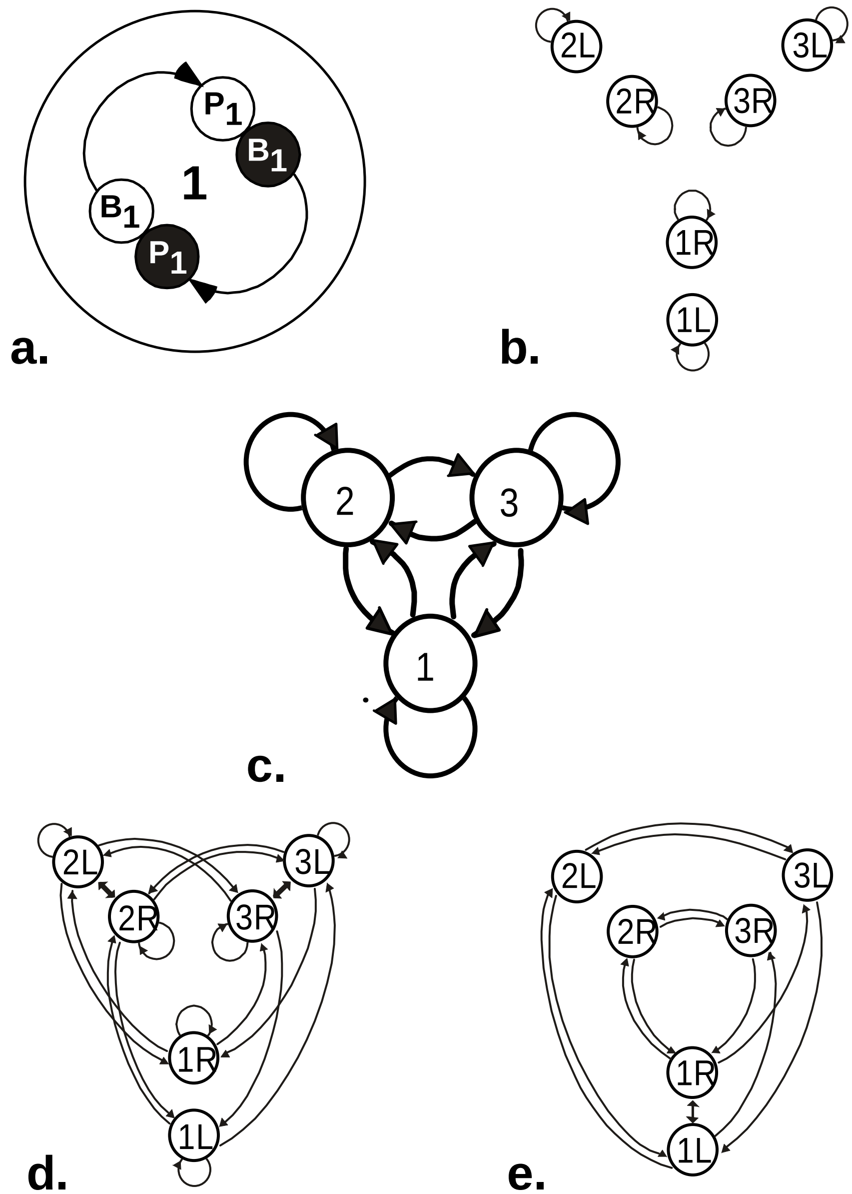


Figure 1. Structue of drill network. a: Drill-1 is shown with partners (coded as black and white) cycling smoothly and continuously. Drills 2 and 3 are similar. Each responds to the other's punch (P) with the block (B) associated with the drill, then gives the associated punch as a stimulus in return. b: Each drill is learned separately, in block practice format. Arrows are standard Markov Process notation e.g. (Pfeiffer, 1978) to show repetition. c: Transitions are learned which smoothly link the drills (nodes) into continuous action. This is the classical generic structure that is instantiated or contextualized to suit specific activities. Variations of this diagram are found in several traditions. d: As instantiated or contextualized by the specific drills used in this study, the network is pruned to remove links that are either physically unavailable, or logistically or tactically undesirable. e: Pruned out links are shown for completeness. Portions here are adapted from (Maier, 1999) (Appendix B) and (Maier & Taber, 2003) (Appendix C).

technical knowledge contained in these transitions is unique and different from knowledge in the drills. The only place this *link knowledge* appears is under the circumstances of switching drills. As in every other field of knowledge, it is the integrative factors that make the basic knowledge accessible and usable. Neither the methods of switching drills, nor the selection of which drill to switch to, is reliably intuitive. Concrete presentation and discussion of the technical and judgment aspects of this *link knowledge* is a necessary step in the construction of expertise in a student. It not only provides technicalities for an individual situation, it also teaches a method and motivates a habit for critical thinking in both student and teacher. Figure 1c shows the generic model of the fully integrated activity, in which all drills are connected by transitional actions that keep the activity smoothly continuous while switching between drills. The relevance of such ideas as procedural networks in cognitive psychology is now evident, with each drill represented as a node, and each transition represented as a link. Also evident is the interest in Markov Processes, as this researcher discussed in (Maier, 1999) (see Appendix B). Figure 1d shows the pruned network of links that results when the generic structure in Figure 1c is instantiated (mathematics term) or contextualized (cognitive term) with the specific hand-sets used in this study. Each node has 3 links to choose from, one that continues the same drill, and two that switch to the other drills. As an aid to understanding, Figure 1e shows the links that are discouraged in instruction, due to mechanical, logistic or tactical concerns.

Actually specifying these integrative factors has a profound editing effect on the basic skills selected for a curriculum, and also the timing and order of instruction. In providing a foundation for detailed rational curriculum structure, a coherent evaluation tool is also generated. Control of the activity passes back and forth between training partners. This incorporates selective perception (Gagne et al., 1993) into every cycle, to assess the situation for changes made by the partner. The stimulus factors selected-for in this perception training are a learned set, found in the three block/punch motor schemas making up the drills.

Either partner can transition to a different drill on any cycle, so that each must respond to unpredictable input, which is perceived as random. Each participant progresses in training by challenging the partner with an increasing level of randomness as training progresses, while simultaneously coping with the partner's increasingly random patterns of transitions. The rate of transitions produced by each partner is limited by the cognitive demands that partner is experiencing at that point in time. If the load is too great, the activity disintegrates and must be restarted. This second level of dynamic stability becomes the foreground to the background dynamic described above. Students are encouraged to increase the frequency and diversity of transitions at the highest rate that maintains the integrity of the activity. Thus, the activity is self-regulating.

The breakdown of stages of procedure in teaching this activity is the result of empirically discovering a level of cognitive load at each stage that requires concentration, but is within capacity for the vast majority of students. For example, consider the case of two relatively new students who are not yet making any transitions. They are comfortably performing Hand-set 1. One, who is a little more motivated, or more prepared by other experience, takes the unchanging, ongoing drill (i.e., 1L) as a static background to conceive and execute a transition (to 2R). The other partner must now devote himself to regaining a stable background for his own thoughts, which now must construct a completely different plan (switching from 2R instead of from 1L) than he was working on previously. The new background is familiar from earlier block practice on it, but the total cognitive load is still quite high.

Though the block/punch sequence repeats seamlessly, the order (block, then punch) is important, and makes a completed thought that is important to the cognitive aspect of the activity. An observer would believe that the purpose of the block is to keep from being hit—in some martial arts programs understanding is often left at that level. In the particular program this activity is drawn from, that purpose exists, but is of secondary importance. The primary purpose of the block is to set up favorable conditions for the punch following it, thus tightly linking the two. Considered in this way, the sequence resembles a set-up and spike in volleyball. In the drills used here, each block is associated with only one punch; thus when the partner is seen executing a particular block, the coming punch can be anticipated. In this way, the stimulus is presented in the same form as marching soldiers receive from their commander—command content followed by an execute signal. The block is the only point-of-choice available in the drills. This gives each partner the longest possible time for processing of recognition, decision, execution of his response. Each block (response) is considered an “event” which gets recorded as described in Chapter IV. The time between one participant’s block and the same participant’s next block is considered a “cycle”, containing two events, one from each partner, as seen in Figure 1a. Thus, the processing time available to each partner is one-half cycle.

As described by Maier and Taber (2003) (Appendix C), the activity generates 1) complexity, 2) uncertainty, and 3) time pressure.

Complexity is is generated by each partner having three options at each cycle: continue in the same drill by using the same block, or transition to one of the other two available drills by using a different block. This is shown in Figure 1d by three arrows departing each node, including the one directly returning. Both partners are encouraged to transition between drills as often as manageable, forcing the other partner to respond correctly to the transition while also planning his own transitions. Complexity is also generated by the additional knowledge involved with each transition—its method, and conditions.

Uncertainty is generated by neither partner knowing what he will have to respond to next. The drill-network structure is thus a random-practice model compatible with the definitions of such in Chapter II, and can be expected to produce greater retention and transferability than block practice. This study agrees with practice schedule theory in that these effects are due to high cognitive load.

Time pressure is generated by each partner's punch as a stimulus for the other partner's block. In this study, the time available before each block must intercept the partner's punch ranged from 0.85 s down to 0.44 s.

The result is a 3-alternative, forced-choice task. It resembles a scaled-down chess game in its strictness of structure, and an expansion of scissors-paper-stone in that all moves can "win" over all moves. Barber's (1999) phrase: "…two tigers chasing each other's tail and running to keep the other from grabbing it" (p.58) captures the perceived experience, including the constant iteration or feedback going on between the partners in the dyad. This experience is very different from that of either a solitary student, or a dyad limited to block practice, and helps eliminate any tendency toward laziness in practice (Lee et al., 1994).

In summary, this activity has several aspects that make it useful for research. The drill network is a clear, simple example of a practice schedule that adjusts seamlessly from pure block practice on any single skill to a level of randomness exceeding that conventionally provided in motor learning research described in Chapter II.

This activity distinguishes itself from any found to-date in contextual interference literature:

I. The activity is both physically and cognitively much more complex than activities like shooting baskets or knocking down blocks.

II. The locus of control in the presentation of changes in the activity is removed from the researcher, and placed in the activity itself. One effect of this shift is that the distribution of repetitions of each drill is very unlikely to be either uniform or predictable. Evenly distributed counts are imposed in conventional studies to make the data within-set compatible with the needs of ANOVA.

III. The activity is truly continuous. Block/punch hand-sets are not finished units of activity which start over at each cycle. Each partner is not simply waiting for the next problem to be presented. Part of dealing with each cycle is generating the problem to present in return with the knowledge that a heavily loaded partner presents less challenge on the next cycle. This integrating of function eliminates any possibility of an inter-task interval.

IV. The results of all decisions are observable in the actions taken (block/punch sequences selected and used).

Probably the most important single characteristic of this activity is that it is not scripted, but extemporaneous, within certain guidelines. This quality allows sufficient complexity of action to resemble natural function—control of all specifics as to order, distribution, intensity, and pace are in the hands of the participants, making the dynamics of their emergent system observable. This training activity was developed historically, not for research, but for efficient instruction. The goal is not to measure a single mechanism while restricting others, but to maximize the total output of the intact system, encouraging all mechanisms to interact maximally and, in Battig's word, idiosyncratically. In requiring different sorts of observation, this activity also permits measurement of performance factors which traditional practice schedule (PS) designs do not.

CHAPTER IV

METHOD DEVELOPMENT

Participants

The participants were 12 males, undergraduate college students, age 18-24, members of a college martial arts club. The all-male sample is representative of the preponderance of males in the martial arts. Their seriousness, engagement, and focus were in all probability impacted by the fact that the activity is part of a club, without the grading motivation of a credit class. Participants (Table 1) had 2-6 semesters of moderately supervised practice with a curriculum including the drill-set being observed. Their background consisted of 2-4 hours of instruction in the drill-set each semester from the researcher, and irregular practice under the faculty advisor of the club, who is also a student of the researcher. These far-from-ideal training conditions would be expected to reduce effect-size of training, providing a challenging test for the proposed measurements.

TABLE 1: Experience levels of participants

Participant	P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
# semesters	6	2	4	1	1	2	2	4	4	6	6

Procedures

<u>Data Collection:</u> Each participant partnered with 10 other members of the sample for 2-minute periods, video recorded for later data-extraction by slow-motion viewing on a personal computer (the 11th set of dyads was missed due to a clerical error). The 2-minute data-collection period at each time-point provided sufficient data-points to assure that the full variety of patterns each dyad was likely to exhibit would be seen, even in the least productive dyad. A pre-test was followed by two approximately 2-hour periods of instruction, with an

hour-long break in between. A post-test, with the same instructions as the pre-test, followed. Test instructions were quite simple: to perform all the drills that could be recalled, and to make as many transitions as possible. Speed was intentionally not referred to in the instructions.

For this study, all dyads including participant P_0 were chosen for analysis, based on P_0's relatively high level of experience and the higher quality of data available due to his position relative to the camera. Thus, P_0 was the central participant in the 10 dyads analyzed in this study. In comparing performance of P_0 across the 10 partners at each time-point, it is assumed that he is not changing significantly over a 20-minute period, and therefore the variance seen is generated by a combination of the interaction with partner, and whatever fatigue/tedium was perceived by participants during that testing period, and over the entire four hours. Based on this assumption, P_0 is responding to 10 versions of a single test. Similarly, participants 1-10 (P_{1-10} , referred to as P_n) are assumed to be experiencing the same test. They contribute to the variance through their individual interactions and relationships with P_0. So P_0 is viewed as taking 10 parallel versions of the same test with different types and degrees of difficulty, while P_n are seen as experiencing a highly sensitive adaptive test.

<u>Data Extraction:</u> Data was extracted from the video recordings of performance by means of stop-action viewing through video editing software on a personal computer. The frame-rate of approximately 30 frames per second allowed very detailed appraisal of all actions. As shown in Table 2, an event was tallied each time a participant blocked a punch. Entries into columns for P_0, and for his partner in each dyad, recorded the drill performed at each cycle. A category labeled "Undefined" was added to the list of 6 drills to be tallied, to

account for events that differed from those defined in the drills so substantially that they could not be considered performance variations.

TABLE 2: Example of data extraction method

P_n		P_0	
Event #	Drill #	Event #	Drill #
5	1L	6	1L
7	1L	8	3R
9	3R	10	3R

Data Analysis: Four instructional goals were measured by analyzing the raw data in terms of four Performance Dimensions, which are described fully in Chapter V, Results. Speed was a simple count of number of events recorded during the 120-second observation period. Undefined was also a simple count of such entries, with a zero-count being the most desirable. Diversity was determined by a standard deviation of the counts in each of the 6 drills, such that a lower standard deviation indicated a more even distribution of observation of the 6 drills. Fluency was calculated by counting the consecutive pairs of drill-numbers that had non-identical entries, showing a change of drill (transition). This dimension was the only one for which individual as well as dyad scoring was possible.

Analysis Goals: The lack of prior literature on this activity made identification of critical main effects and interactions a major goal. For the same reason, power analyses of tests were unavailable, so sample size was based on an available group with sufficient range of experience. All statistical treatments used in this study are post hoc tests. Spatz (1997) said, “Any and all differences for which a story can be told are fair game for post hoc tests” (p. 232).

Raters and Ratings

Inter-rater reliability was not considered important for the purposes of this method-development study. The level of field knowledge needed by a rater for this particular activity depends upon the degree to which participants remain within the prescribed activity. This cohort strayed frequently. Some of the actions observed were recognized as variants of, or substitutions for, expected actions. At some points, especially in particular pairings, arousal overcame structure and control had to be regained.

CHAPTER V

RESULTS AND DISCUSSION

Differences between individual students were seen, as well as changes within-student across an instructional intervention. Evidence of members of a dyad performing and changing together in coordinated ways supported the expectation that a system with its own stability and dynamics emerges from the iteration of response-to-response, intention-to-intention, and capacity-to-capacity.

To crystallize the more advanced research questions for this study required considerable interaction between the model in Chapter III and the various literatures in Chapter II. In much the same way that research questions and hypotheses are different formulations of the same thought, the evolved research questions in Chapter I and the performance dimensions presented below are the same thoughts. The global questions dealing with perceived value, nature of training effects, and insights on performance expressed themselves as the Performance Dimensions of Speed, Diversity, Undefined, and Fluency (Table 3).

Speed is the easiest Performance Dimension to measure, and offers a gross estimate of cognitive load of the activity in that a steady pattern of small (or a few major) processing delays add up to a reduced production of events. Speed is determined by counting an event each time a partner blocks, and includes the block and the hit following it; a cycle, therefore, includes two events, spanning the time between two blocks from the same partner.

Two Performance Dimensions, Diversity and Undefined, are concerned with the breadth and boundaries of the ‘option space’ allowed by the activity. As each of the drills and transitions in the activity is practiced, its cognitive load reduces. There is a tendency for those options learned first to become comfortable and those less well-learned to become more uncomfortable by contrast. This relative comfort is interpreted as an inverse of effort required, and therefore, of relative cognitive load. Thus, Diversity measures distribution of activity that stays within the bounds of the assigned activity. Undefined counts the times a participant breaks the boundaries, departing the activity's option space. It indicates that a participant has gotten lost or confused, and is interpreted as showing a moment of cognitive load exceeding the participant's capacity.

These first three Performance Dimensions focus on the drills, represented as nodes in the network (see Figure 1 in Chapter III). As the two participants interact in the larger system of the dyad, their individual performance is seen only in terms of the transitions they perform. These transitions literally "slip through the cracks" between the drill counts on which the first three dimensions are based. Shifting focus from drills to transitions, from nodes to links, is the step necessary to view Fluency.

Fluency is the first dimension that makes visible the individual participant's contribution to the dyad's activity. Just as in language, fluency measures the freedom with which a participant moves through the knowledge base, enabling him to express his thoughts. In this activity, Fluency shows the nimbleness with which a participant copes with, and provides the challenge of unpredictability. The first measure of Fluency discussed is a simple count of transitions between drills. Two goals of the intervention were 1) to increase performance of less-familiar hand-sets (Diversity) and 2) to increase switching between individual drills (Fluency). Speed was neither encouraged, nor discouraged. Structure of the model indicates that Diversity and Fluency are not strictly independent as the addition of a drill implies a minimum addition of one transition to enter it. However, the dependency is quite loose as two drills are sufficient to generate a transition score of 100%, by transitioning between the two drills at every event.

TABLE 3: Performance dimensions, associated measures and performance factors indicated

	Performance Dimension	Associated measure	Indexes:
1	Speed	Event count	Processing speed, as gross indicator of cognitive load
2	Diversity	Drill distribution	Perceived cognitive load of individual drills
3	Undefined	Departures from activity	Lack of recall
4	Fluency	Transition count	Cognitive load tolerance

One participant (P_0) was in all 10 dyads. Other participants ($P_{1\text{-}10}$, refered to as P_n) were numbered to match the order in which they partnered with P_0. Each dyad is referred to by the number of P_0's partner (i.e.: Dyad 1, Dyad 2). This sample group is sufficiently diverse to show a wide range of performance on each dimension. It will be seen on each performance dimension that one or two participants fall significantly above or below an otherwise strong pattern. Such outliers generate challenges in teaching situations. As P_0 was in all 10 dyads, this broad range of outcomes indicates the interactions between P_0 and 10 different partners.

Performance Dimension 1: Speed

Speed of performance is derived from the total events (blocks) observed in each dyad for each 2-minute session (Table 4). Event-time is defined as: the time elapsed between one event (blocking response by one partner) and the next (blocking response by the other partner). This is the critical time unit in the activity, as it is the time available for processing and responding to the triggering stimulus of the hit.

Productivity during the pre-intervention test ranged from 141-291 events per 2-minute session. Productivity decreased after intervention, as was predicted due to the increased cognitive load of unfamiliar material from the intervention. Event counts shifted to a range of 115-233 events per 2-minute session. The high event score of 291 indicates that Speed significantly higher than the pre-intervention mean of 222.7 events per 2-minute session is quite possible.

Two dyads (8, 9) produced dramatically reduced event counts on the post-test. The video record showed the reason for this was that they went off-task before the session was completed. This fact actually makes their posted scores inaccurate, because scores are generated based on event counts for a full 120 seconds. Thus, data from these two dyads required adjustment for valid comparisons to be made (Table 4 and Figure 2). The level of production during the on-task period of the post-test was used to estimate the number of events that would have been produced had these dyads maintained the same production-rate for the entire 2-minute session. Estimated values were generated by dividing actual production by number of seconds on-task and then multiplying by 120 seconds (2 minutes). After adjustment, event production by Dyad 8 was near the mean for the group, and Dyad 9 had the largest post-intervention event count. The group mean of post-intervention event counts was shifted by this adjustment, without substantially changing the SD. The adjustment to Dyad 8 was relatively minor, as they actually performed for approximately 90% of the period. However, the adjustment to Dyad 9 was more questionable, as they performed for less than half the period. The adjustment does not make them the only dyad to produce more events post-intervention than pre- intervention however, as Dyad 3 also increased.

Increasingly strong relationships (pre-: r=0.684, α=0.042; post-: r=0.742, α=0.022) are also seen between event counts and experience (Figures 2b and 2c), if Dyad 10 is excepted. Though Dyad 10 performs with the group in terms of pre-/post- change, they perform on this dimension well below a level consistent with 6 semesters of experience.

TABLE 4: Total event production of 10 dyads (Speed)

Dyad	1		2		3		4		5		6	
Condition	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Events	252	232	209	204	169	186	141	120	232	203	226	198

Dyad	7		8		9		10		Group Means		Group SD	
Condition	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Events	261	223	291	187	261	115	185	173	222.7	184.1	46.81	39.12
Seconds Ontask				108.6		51.2			Adj.		Adj.	
Adj. Events				206.6		269.5				201.5		39.10

Note. Rows are added to show adjustments to two dyads.

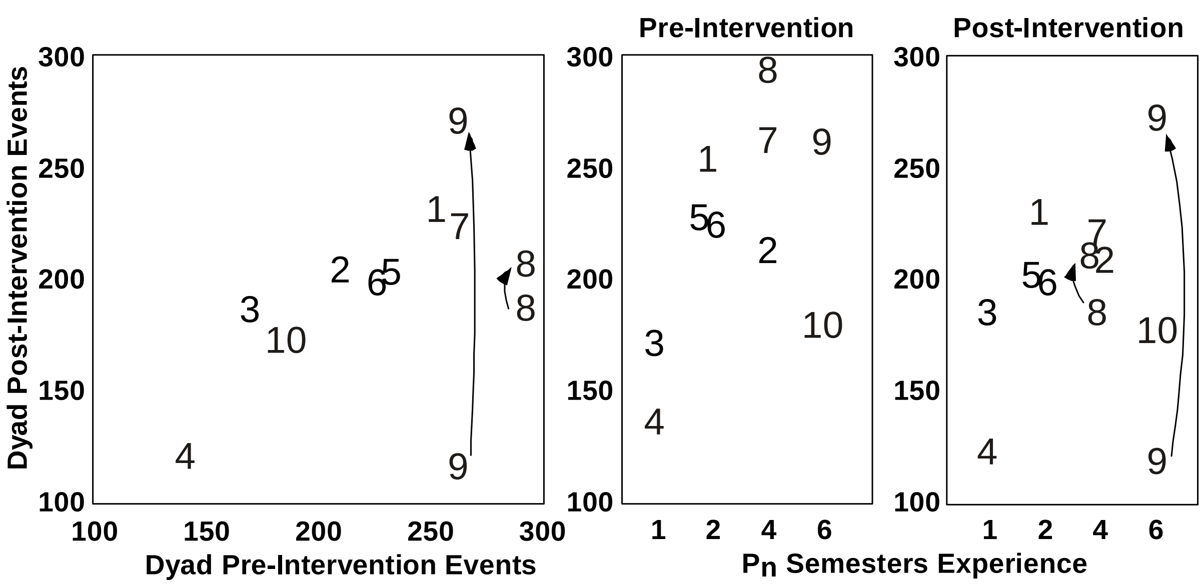


Figure 2. Event production scores (Speed).

Regression analysis with pre-intervention event counts as the independent variable, and post-intervention event counts as the dependent variable yielded r=0.361, α=0.306 (actual counts); and r=0.802, α=0.005 (adjusted counts). The validity of this adjustment was tested in two ways: 1) the group is otherwise consistent, such that the previous regression produces r=0.915, α=0.001, with Dyads 8 and 9 removed, 2) regressions in Table 5 show high stability of speed within each dyad by means of regressions of their production at 51.2 s and 108.6 s time-points against their total productions on pre- and post-tests. It also shows a high stability in the productivity decrease after intervention. These results support the idea that performance that slows on the post-test is related to the intervention and may be one coping mechanism for dealing with increased cognitive load.

TABLE 5: Regression results of event production rates (Speed)

Time Point	51.2 s		108.6 s	
	r	α	r	α
Actual Pre-	0.945	<.001	0.998	<.001
Adjusted Post-	0.933	<.001	0.998	<.001

P_0 was clearly capable of performing with each partner within that individual's range. His adjustment to each partner further indicates that the club views the activity as cooperative learning, not as competition. This perspective on the part of the participants will be extremely significant on later measures. It will also be important later that it was 2 of the 4 fastest dyads who failed to complete the post-test.

Interlocutory: Though obviously important as a performance measure, large differences in total event-counts effect the meaning of simple counts on other measures. To filter out this variance, all further results will be standardized as percentages of that dyad's, or that individual's, total actual events under each condition (pre- and post-intervention). It will be seen that the performance of the two non-finishers on further measures is quite consistent with the rest of the dyads throughout their shorter performance, as it was on event production. Performance quality, including intensity, is shown stable throughout performance duration. As described in Chapter IV, the 2-minute time period was semi-arbitrary, based on generating a large number of data-points.

Performance Dimensions 2 & 3

Distribution of activity among available drills is shown in Table 6. When this study was performed, the participants had already been exposed to all 3 Hand Sets. Hand Set-1 (Drills 1L and 1R) was the most familiar, and Hand Set-3 (Drills 3L and 3R) was the least familiar to this group. One goal of the intervention was to increase utilization of the less familiar drills.

There are complexities in the interpretation of this data. As was shown in Chapter III (Figure 1), the fully-linked generic network was heavily pruned by technical considerations to the specific Hand Sets used here. In the resulting sparse network, the system (dyad) is not free to move from one state (drill) to all five of the others. At each event, the current state (drill) limits the available next states (drills). The least challenging approach to the activity is to cycle through a limited subset of the possible sequences of drills, resulting in heavy weightings toward a few drills. Thus, greater Diversity shown by a more even distribution within-sample is considered an indicator of greater mastery on this dimension.

TABLE 6: Percentage of cycles performing each drill (Diversity)

Dyad	1		2		3		4		5		6	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Drill-1L	27.8	28.9	28.2	34.3	21.3	26.9	48.2	21.7	29.3	32.0	31.0	28.8
Drill-1R	21.0	7.3	24.9	11.3	14.8	8.6	41.1	29.2	6.9	3.0	14.2	6.6
Drill-2L	16.3	18.5	22.0	14.7	25.4	29.0	1.4	12.5	17.7	21.2	25.2	25.3
Drill-2R	27.8	28.4	13.4	27.0	30.8	28.0	0.0	11.7	28.9	36.0	13.7	25.3
Drill-3L	4.8	11.6	1.0	8.3	2.4	5.9	2.8	3.3	12.1	4.9	13.3	6.1
Drill-3R	2.4	1.7	10.5	4.4	2.4	1.6	0.7	21.7	1.7	3.0	2.7	8.1
Diversity (SD)	11.1	11.2	10.2	11.6	11.9	12.6	22.6	9.2	11.4	15.1	10.0	10.8
Undefined	0.0	3.4	0.0	0.0	3.0	0.0	5.7	0.0	3.4	0.0	0.0	0.0

Dyad	7		8		9		10		Means		SD	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Drill-1L	32.6	23.8	32.0	23.0	36.4	34.8	34.1	19.1	32.1	27.3	7.0	5.4
Drill-1R	12.6	1.8	19.6	5.3	23.0	7.0	27.6	11.6	20.6	9.2	9.6	7.7
Drill-2L	29.9	37.2	16.2	16.6	17.2	11.3	10.3	20.2	18.2	20.7	8.2	8.0
Drill-2R	8.0	28.3	10.0	27.3	1.9	15.7	4.9	15.6	13.9	24.3	11.4	7.5
Drill-3L	7.3	8.1	8.2	14.4	8.0	15.7	5.9	19.7	6.6	9.8	4.0	5.3
Drill-3R	8.8	0.9	11.0	7.5	11.5	5.2	15.7	13.9	6.7	6.8	5.3	6.5
Diversity (SD)	11.6	15.2	8.8	8.5	12.2	10.6	12.0	3.5	12.2	10.8	3.8	3.4
Undefined	0.8	0.0	3.1	5.9	1.9	10.4	1.6	0.0	1.9	2.0	1.9	3.6

Performance Dimension 2: Diversity

The intention of the intervention on this dimension is to increase familiarity with the less-practiced drills, decreasing their cognitive load past a threshold where practice of those drills is comfortable enough to be included spontaneously. The SD of each dyad's 6 drill scores under each condition is here interpreted as a measure of the diversity of the distribution, such that the smaller the SD, the more uniform is the distribution. The Undefined score is intentionally left out of this measure, as its goal is to reach zero, and would obscure the Diversity of the other scores. This Diversity measure shows a very modest effect of the intervention on the group overall, with considerable variability in the individual dyads. As seen in Figure 3, Dyad 4 improved the most, and Dyad 10 achieved the best

Diversity overall. The Diversity measure is most useful in cases like Dyad 2 or Dyad 5, where the complexity of the redistribution makes simple scrutiny of Table 6 unreliable. On this measure, 4 dyads (4, 8, 9 and 10) improved their scores, one (Dyad 1) stayed the same, and the other 5 got worse. The magnitude of the 4 improvements was large enough to make the mean Diversity score improve. Of particular interest here is the fact that Dyads 4 and 10, each of whom produced the lowest speed in their experience level, turned out to be vastly the most successful in increasing their Diversity scores. This indicates that their focus was on at least one of the two purposes declared for the intervention. Dyads 5 and 7 got significantly worse on Diversity, breaking away from the group. This indicates that their focus was elsewhere. As P_3 and P_4 were the least experienced participants (1 semester), Figure 3 indicates that the intervention was sufficient on this dimension for the newest students, but perhaps not for the more senior ones.

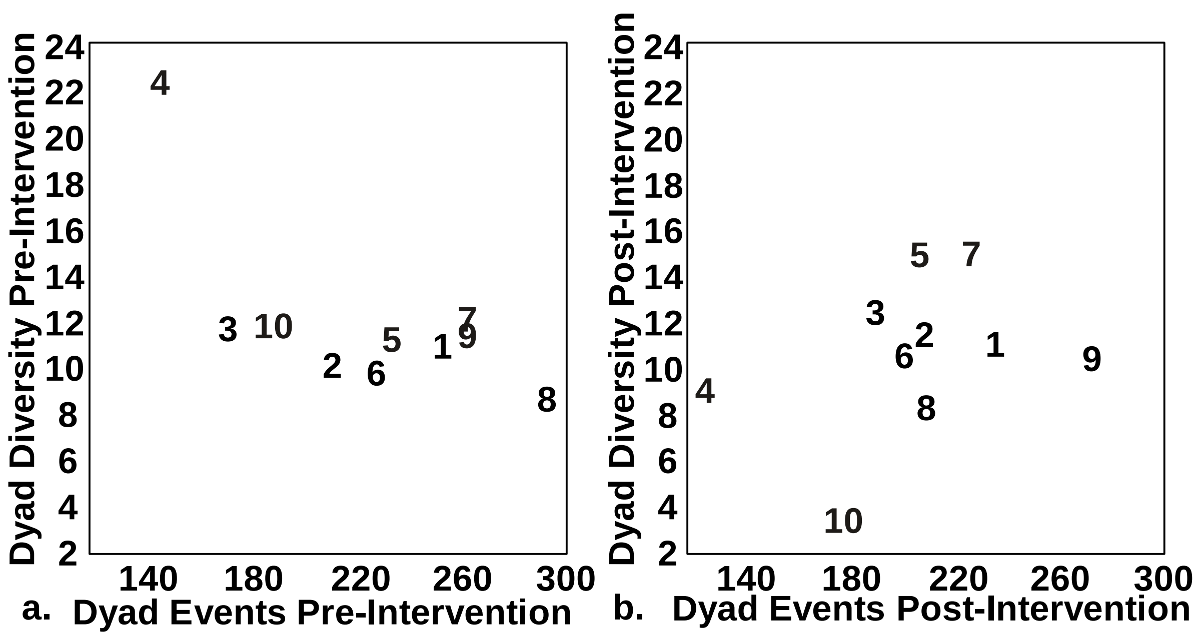


Figure 3. Diversity scores.

Performance Dimension 3: Undefined

An Undefined event is one which departs the instructed set of drills, interpreted unfavorably as a lack of either understanding or focus. As shown in Figure 4, most dyads generated 0.0 Undefined on the post-test, with two of them also generating 0.0 on the pre-test. This was a desired outcome and may indicate increased focus after the intervention. Three dyads (1, 8 and 9) increased their production of Undefined events. All 3 were among the four fastest dyads on Performance Dimension 1 (Speed). Two of them failed to finish the post-test. There is a direct relationship here, as Dyad 1, who finished, and Dyad 8, who nearly finished, show similar increases, while Dyad 9, who performed less than half the time, also generated the highest (poorest) Undefined score.

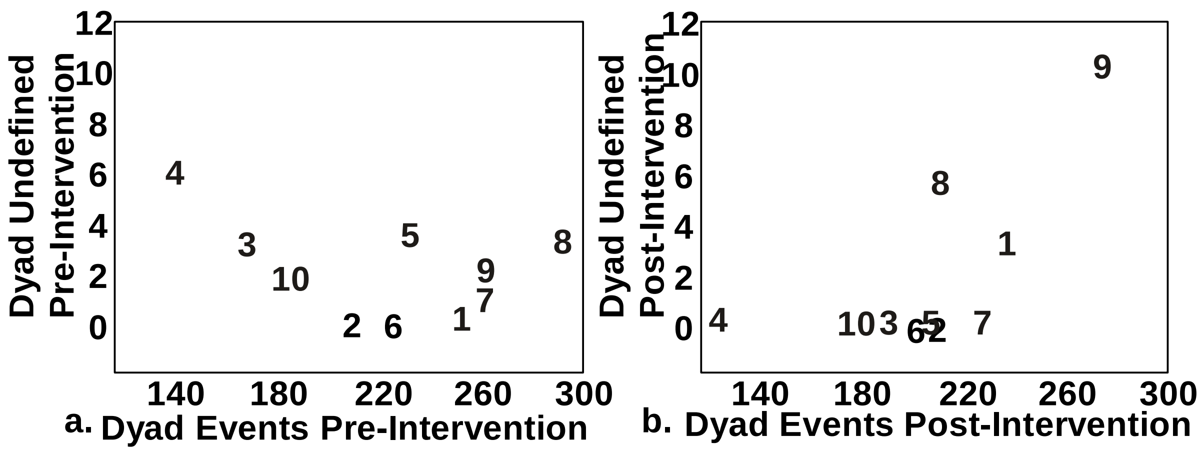


Figure 4. Undefined scores.

Interlocutory: In addition to its direct contribution to understanding performance, the data that is shown as sums in Table 6 made possible the next level of analysis, which focused on changes in consecutive pairs of events. A consideration of sequence is necessary for both based on data, and ecological reasons. Reasoning based on data relies on the fact that numerous samples, unique in sequence, would generate the same sets of totals seen above.

Ecological reasoning is based on the fact that the activity is produced and perceived as a sequence, one event at a time, and that memory of history is significant in selection of next action. The moment of transitioning is also expected to be the highest in cognitive load for both partners in the dyad: the producer in terms of generating it, and the responder in terms of coping with it. Participants were specifically instructed to transition as frequently as they were able to, so scoring this dimension is critical.

Performance Dimension 4: Fluency

Fluency of performance is indicated by the frequency, and therefore the freedom, with which participants switch (transition) from one drill to another. The second goal of the intervention was to increase transitioning activity in the practice. Two contributions this dimension makes to understanding are its focus on changes, which is clearly where the greater cognitive challenges lie, and the opportunity to view individual performance for the first time.

A transition is indicated in a Markov transition matrix by a two-event sequence in which the two events are different. In focusing on transitions between drills, Markov Processes count total traffic in each link rather than in each node of a network, quantifying what is shown in principle in Figure 1c in Chapter III. A shift in activity (in Markov terminology, probability mass) outwards from the cells on the diagonal of a Markov matrix into the rest of the cells was pointed out by Maier (2001) (see Appendix B) as an indication of a change in behavior from repeating a drill to transitioning to a different drill. Tables 7 and 8 provide this view of the data, in terms of dyads, and in terms of individuals.

<u>Dyads:</u> Table 7 shows that, post-intervention, 3 dyads increased transitioning by 11-12%, while 6 dyads increased by 4.9-8.5%, and 1 decreased by 6.3%. So, 9 out of 10 dyads increased transitions after the intervention, indicating a positive effect from the intervention on this dimension. Note that the 2 dyads who went off-task during the post-test (Dyads 8 and 9) exhibited the highest and 4th-highest increases in transitions while they were on-task, and also the two highest percentages of transitions on the post-test. Regression of post-test on pre-test transition scores for the 10 dyads in Table 7 yielded a modest r=0.581, α=0.078.

TABLE 7: Dyad transition percentages (Fluency)

Dyad	1		2		3		4		5		6	
Condition	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Transitions	20.2	31.9	30.1	41.2	33.1	39.2	24.8	33.3	31.5	36.9	31.0	36.9
+ Change		11.7		11.0		6.1		8.5		5.0		5.9

Dyad	7		8		9		10		Means		SD	
Condition	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Transitions	36.8	30.5	38.5	50.3	34.1	41.7	36.8	41.6	31.7	38.3	5.6	5.8
+ Change		-6.3		11.8		7.6		4.9		6.6		5.3

In correlating experience with Speed, Dyad 10 was interpreted as an low outlier. This low score was later argued as being a trade-off for being a high outlier on Diversity. A similar rationale suggests seeking a single outlier here on Fluency. Figure 5 shows this outlier to be Dyad 7. Dyad 7 was among the 4 fastest in Speed, and also one of the least successful in Diversity, as seen in Figure 3a and 3b. These factors support an evaluation of Dyad 7 as focusing on Speed rather than on increasing either Diversity, Undefined or Fluency, thus not being responsive to the instructions in the program. Excluding Dyad 7 on this basis strengthens the relationship to r=0.874, α=0.002.

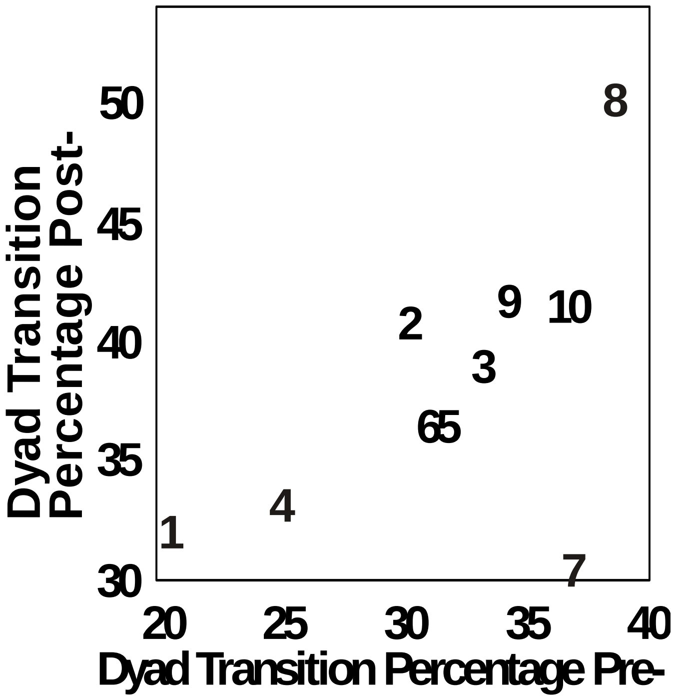


Figure 5. Dyad transition percentages (Fluency).

Regression analyses of experience on pre- and post-test scores showed a declining relationship (pre-: r=0.588, α=0.074, post-: r=0.452, α=0.189). Once again, these summary statistics do not tell enough of the story. Figure 6 shows a very consistent response to the intervention in 9 out of 10 dyads, with the higher-performing dyads in each experience group remaining high, but the lower-performing dyads narrowing the gap. The exception again is Dyad 7, as discussed above. The intervention clearly benefited Dyads 1 and 4, who lagged the group pre-intervention, but joined it afterwards.

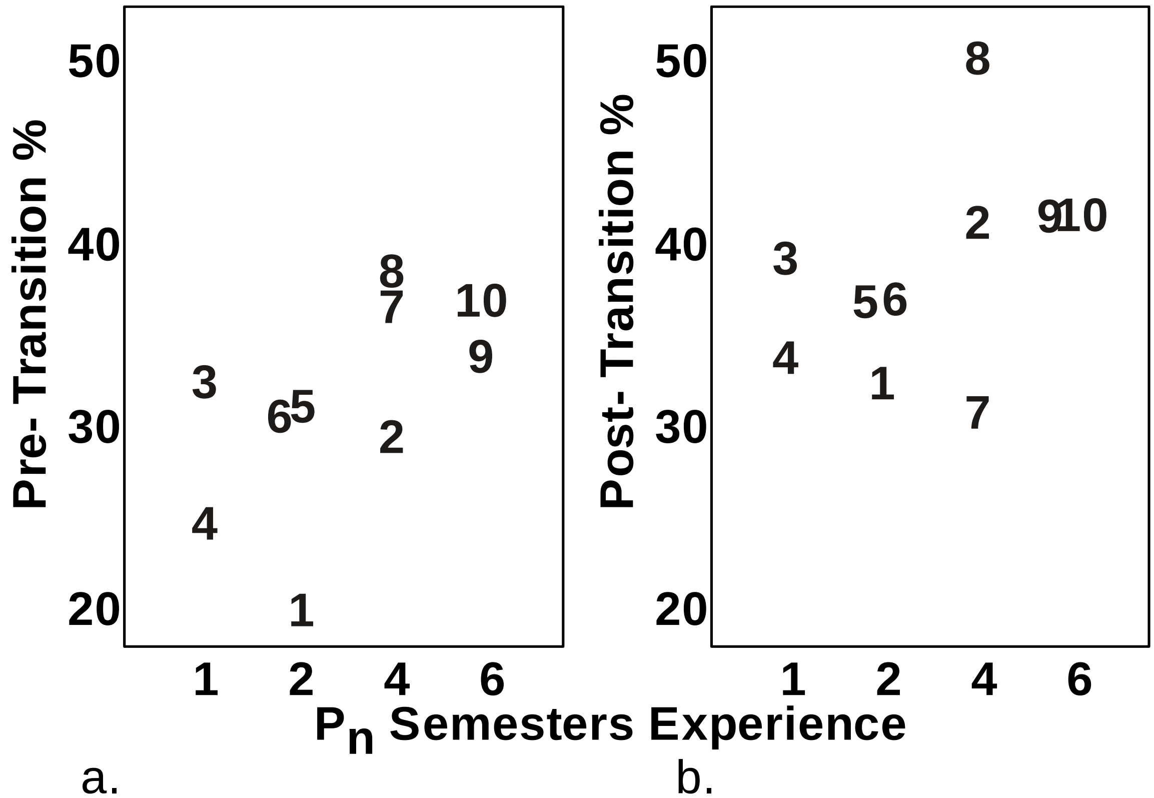


Figure 6. Dyad transition percentages versus experience (Fluency).

<u>Individuals:</u> Table 8 is the first data representation that provides a view of individual participants within their dyad by further partitioning the variance shown in Table 7. Several worthwhile comparisons are possible here. P_0 reduced transitions from his own pre-intervention performance with two partners (P_3 and P_7), and tripled them with another (P_2), while making widely varying increases (1.0%-19.2%) with the rest. P_0's margin of improvement on the post-test, in terms of the mean of the additive changes (7.4), is similar to that of the rest of the group (6.0%). The difference in mean transition scores between P_0 and his partners (P_n) decreased post-intervention, indicating that they are catching up to him. This interpretation is further supported by the fact that P_0's pre-test score in Dyad 2 is nearly 2 SD's below the mean. Removing Dyad 2 would shift P_0's mean pre-test score to 49.7,

reducing his mean additive change to 4.1, and would shift his partners' mean pre-test score to 13.9, increasing their mean additive change to 9.0. So removing the exception of P_2, P_n's increase would be more than twice that of P_0.

P_0 dramatically outperformed almost every partner on both pre- and post-tests (the sole exception being the pre-test with P_2). Dyad 2 gives the only opportunity in this small sample to view P_0's interaction with a strong competitor. The uniqueness of Dyad 2's interaction is seen clearly in Figure 7. P_2's pre-intervention transition score (Table 8) is 2-5 times higher than any of the other partners. P_0's pre-test score against P_2 is his lowest in the entire study, apparently a response to the only real challenge among the 10 partners. On the post-test, P_2 drops back to the upper end of the range generated by P_n and P_0 performs equivalently to his own scores with other partners. This single example is an indicator that the kind of interaction this model system is proposed to explore is, in fact, observable. It supports the idea that P_0's limited cognitive resources are allocated first to coping with the external load, and when that load diminishes, re-allocated to the internal load of generating transitions.P_2's and P_0's pre-intervention scores (42.9, 17.3) are approximately the inverse of P_n's and P_0's mean scores (16.8, 46.4) on the pre- test. The shift in dominance from P_2 to P_0 suggests that the partners other than P_2 do not have the cognitive resources available to initiate transitions while coping with the cognitive load of P_0's frequent transitions. Excluding Dyad 2 as the special case in this dimension, P_0 pre/post did not quite reach conventional significance ($r=0.640$, $\alpha=0.064$), but P_n pre/post did ($r=0.704$, $\alpha=0.034$).

TABLE 8: Individual transition production (Fluency)

Dyad	1		2		3		4		5		6	
Participant	P_1	P_0	P_2	P_0	P_3	P_0	P_4	P_0	P_5	P_0	P_6	P_0
Pre %	7.9	32.5	42.9	17.3	17.9	48.2	14.3	35.2	11.1	52.2	19.5	42.5
Post %	12.1	51.7	29.4	52.9	33.3	45.2	20.0	45.0	16.7	56.4	30.3	43.4
Add. Change	4.1	19.2	-13.4	35.6	15.5	-3.1	5.7	9.8	5.6	4.3	10.8	1.0

Dyad	7		8		9		10		P_n		P_0	
Participant	P_7	P_0	P_8	P_0	P_9	P_0	P_{10}	P_0	Mean	SD	Mean	SD
Pre %	13.8	58.8	11.6	65.5	10.8	57.3	18.5	54.8	16.8	9.9	46.4	14.6
Post %	14.3	46.8	31.9	69.9	13.8	70.2	26.7	56.3	22.9	8.3	53.8	9.7
Add. Change	0.4	-11.9	20.3	4.4	3.0	12.9	8.3	1.5	6.0	9.1	7.4	13.1

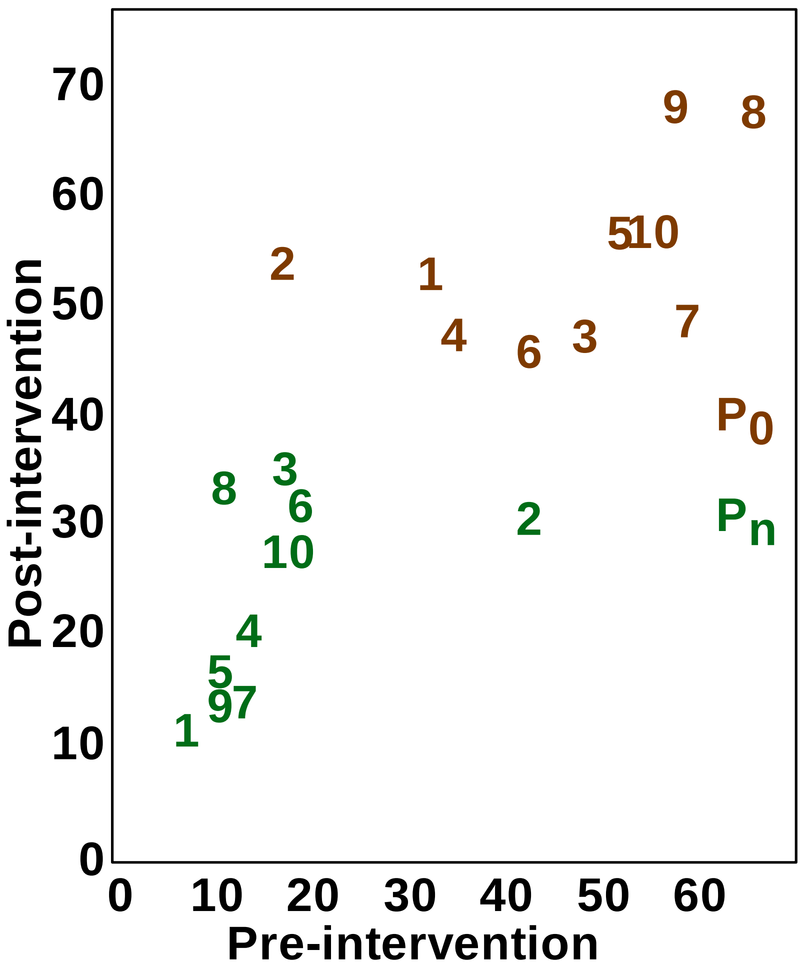


Figure 7. Individual transition production (Fluency).

Figure 8 could easily be interpreted to say that there is little difference shown by P_n across levels of experience. But this would be forgetting the interaction with P_0 that produced each transition count, as shown in Figure 9. P_0's transition production is the external cognitive load that each partner must handle prior to investing in the internal cognitive load of his or her own production. The fact that 4- and 6-semester students showed similar production under a 60% external load that 1-semester students showed under a 40% external load resembles a weight-lifter producing the same number of repetitions under heavier weights.

Figure 8 indicates that P_n's experience is a greater factor on P_0's scores than on those of P_n. Regression analysis of the contribution of the partner's experience in predicting pre-test transition scores generated $r=0.369$, $\alpha=0.294$ with Dyad 2 included, and $r=0.671$, $\alpha=0.048$ with Dyad 2 excluded, re-confirming its belonging to a different pattern on this measure than the other dyads. Regression of experience on post-test scores generated $r=0.667$, $\alpha=0.035$.

Likewise, by forgetting the nature of the interaction, one would expect P_0's production to be greater with the less-challenging 1-semester partners, producing a negative slope in Figure 8. The nature of the activity leads strongly toward cooperative learning rather than competition, as a competitive approach would cause frequent breakdowns and re-starts, little learning and high frustration. In order to 'keep the volley going' in tennis terms, the pattern of evidence fits P_0 actively loading each individual partner about the same, in other words, the emergence of a self-regulatory function.

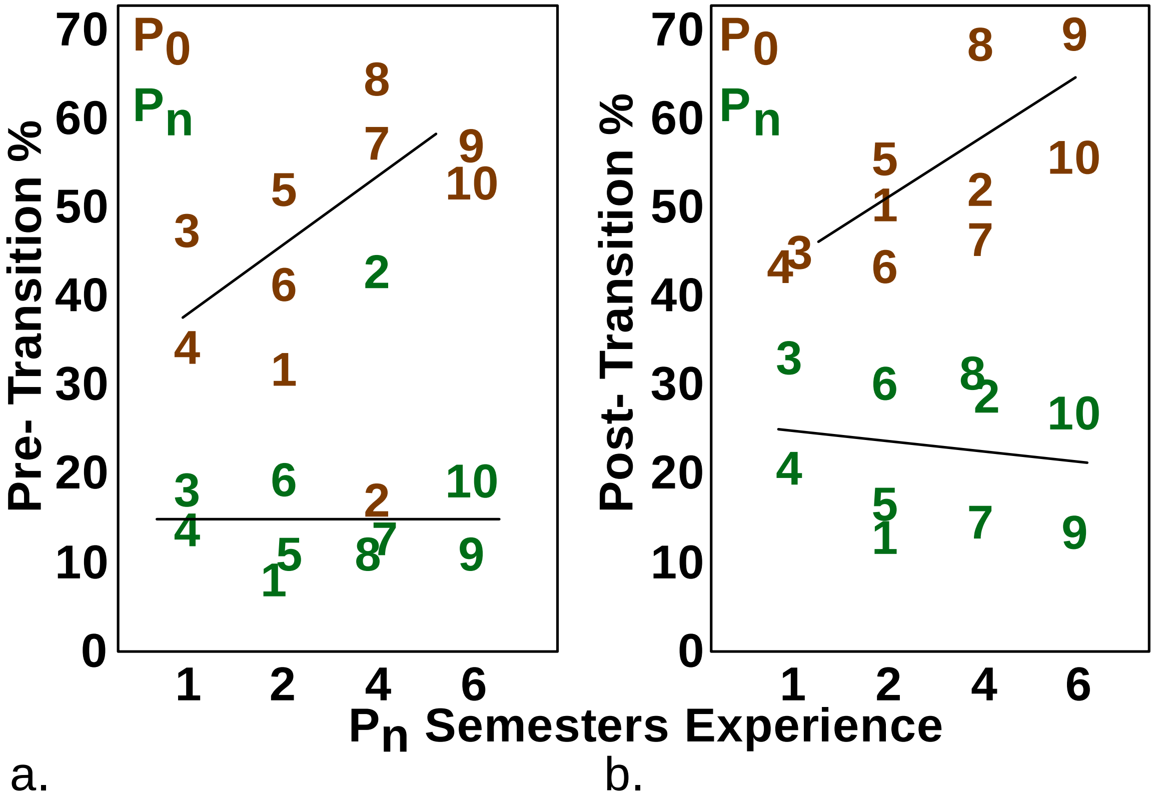


Figure 8. Individual transition production versus experience (Fluency).

In a system as complex as this one, the puzzle is to find the perspectives on the data that reveal its strongest relationships. Figure 9 shows the interactive aspect of additive changes in transition percentages of individuals within each dyad, pre- and post-intervention. Emphasizing this interactive aspect instead of independent individual scores as above, inclusion of Dyad 2 significantly strengthens the correlation (r=0.619, α=0.056), instead of weakening it. It is noteworthy that an aspect of performance that seems so divergent in terms of individual participants and isolated scores turns out to be so consistent in terms of the interacting dyad. Along with 2, 3 supports the pattern that the dynamic extends into the regions where one partner actually looses ground—and that this relation may be linear. The exception to this pattern is Dyad 7, in which P7 did not increase transitions in response to decreased transitions from P0. Thus Dyad 7 gives further evidence of ignoring the stated goals of the intervention. Without Dyad 7, r=0.875, α=0.002. This relationship of one partner increasing more than the other strongly supports the interpretation of a dyadic dynamic distributing the cognitive load of the activity.

Figure 9b displays the same data as Figure 9a, but reveals a different aspect of the relationship. A strong pattern is again immediately visible, this one indicating a threshold value. P_0 improved more against a partner whose improvement was less than 5%, and improved less against a partner whose improvement was more than 5%.

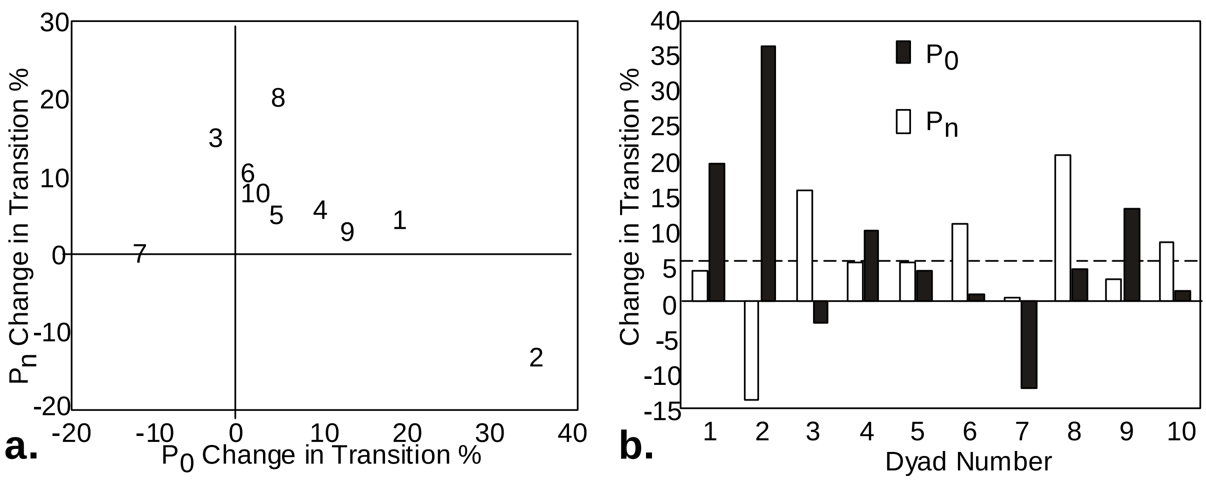


Figure 9. Changes in individual transition production (Fluency).

Interlocutory: A critical shift of focus onto change instead of pre- and post- scores focuses on an interaction that would more closely represent the empirically perceptible dynamic. It further develops the idea of a self-organized self-regulatory interaction between the improvements made by the partners in a dyad.

As in Figure 9, the interaction with P_7 presents a dramatic exception to this pattern. As noted in Table 7 and Figure 1, Dyad 7 was one of the fastest (along with 1, 8 and 9). Dyad 7 succeeded where Dyads 8 and 9 failed in finishing the post-test. Dyad 7 may have accomplished this by reducing performance demands on a different dimension, where they diverged from the group in a rather subtle way. In post-intervention Diversity (Table 6), three other dyads (3, 5 and 6) show a weighting similar to that of Dyad 7, with very heavy emphasis on Drills 1L, 2L, and 2R. But the change each dyad made to arrive at those similar scores may be highly important. Scores on Hand Set 2 (summing Drill-2L and -2R) show these changes: Dyad 3: +0.8%, Dyad 5: +10.6%, Dyad 6: +1.7%, and Dyad 7: +27.6%. This is a huge consolidation, resulting in 2/3 of Dyad 7's total activity collecting in the two versions of one Hand Set.

According to the theoretical structure of the model, Dyad 7 may have maintained stability at this high speed by reducing the load on both Drill Distribution and Fluency Performance Dimensions. Additionally, consistent with the dynamic observed in Dyad 2, and with the cooperative nature of loading suggested above, P_0 reduced transitioning enough that P_7 had the resources available to manage a barely-measurable increase in transitions. All this supports a conclusion that, like in Dyads 8 & 9, engagement with the assigned task was low, such that focus was on speed rather than content.

Summary of Results

This study has described and demonstrated a new model for studying dyadic behavior. The alternating pattern of choosing to continue an ongoing activity or switch to one of 5 other options in a limited option space, at a rate of around once per second forces participants to develop and utilize management strategies. These strategies include prioritizing which dimensions of a complex task will be allocated the limited attentional (cognitive) resources available. They also include reliance on structured knowledge which is complete, organized and easily available. Behavior is measured on 4 dimensions of perfomance: Speed, Diversity, Undefined, and Fluency. These dimensions have been shown to have strong interactions, reflecting the aforementioned management strategies. Learning has been tracked by showing effects of a 4-hour instructional intervention with data collections immediately pre- and post-.

Discussion

Careful study of this traditional activity permitted definition of four Performance Dimensions: Speed (measuring processing time), Diversity (measuring exploration of the option space), Undefined (staying within the activity's boundaries), and Fluency (measuring frequency of switching defined routines by dyads and by individuals) for evaluating group, dyad and individual performance. These were sufficient to show individualized approaches to prioritizing and trading-off among the Performance Dimensions as a mechanism to dynamically maintain the stability of the activity. The individualized approaches are interpreted as showing selective cognitive resource allocation in a multi-dimensional activity.

Despite a small sample of only ten dyads, very strong correlations ($r > .800$, $\alpha < .01$) were found on three of the four dimensions, supporting the basic thesis that learning stimulated by the activity would be clearly visible in a classroom or application setting. For the majority of dyads, performance before and after the intervention were strongly correlated on the Speed and Fluency dimensions. On the Diversity dimension, change in performance strongly correlated with semesters of experience.

Moderate correlations ($.800 > r > .650$, $\alpha < .05$) were found between semesters of experience and pre- and post- measures of Speed, between experience and pre- and post-measures of individual Fluency shown by P_0, and between pre- and post- Fluency shown by P_n. Borderline correlations ($.650 > r > .450$, $.05 < \alpha < .075$) suggest further coherence (at least in the company of the stronger ones) which a larger sample, stronger intervention, or more controlled conditions might show more convincingly. These correlations offered a clear background against which to view individual dyads presenting divergent performances on particular dimensions. These divergences point toward interactions between the dimensions and indicate important individual differences in approaches and priorities in a complex activity. These dimensional interactions create complexities in discussing each performance dimension separately, but support the thesis that the activity is irreducible because a dyad emerges as a system rather than appearing as independent individuals. The fact that one participant (P_0) was in all 10 dyads indicates that the individual P_n (and his interaction with P_0) in each dyad was driving the differences recorded.

Several correlations of pre- and post- measures with experience (Speed, Undefined, Dyad Diversity, and P_0 Diversity) strengthened after the intervention, suggesting a difference between the effectiveness of four hours of instruction from an expert instructor versus extended instruction from an assistant instructor. The intervention focused on Diversity and Fluency, but appeared to have greater effect on Fluency than on Diversity.

Two major conclusions arise from this study. The first is that it is feasible to track and measure the management of cognitive resource allocation, thereby identifying the likely focus of attention within a complex activity. The second is that a self-organizing dynamic emerges from the structure of a dyad engaging in the activity.

<u>Cognitive Resource Management & Attentional Focus:</u> Cognitive resource management is a skill that must develop to meet the challenges of a given activity. Allocation

of cognitive resources is closely related to attention distribution and therefore, directing the focus of attention. A person's attention might be focused from within by his purpose or goal, or from without by the stimulation of novelty. Successful attention management scans for external stimulation, but filters by means of purpose. Communication, and therefore teaching, is most successful if the student's focus is aligned with the teacher's. It can be critical to successful teaching to know where the distractions in a topic are, and how to recognize the signs of a student who might be energetic, but mis-directed.

The following contrasts support a conclusion that distribution of cognitive resources is accessible and measurable in the framework of the investigated training activity.

Dyad 7 showed high Speed as well as poor Diversity that declined further on the post- test, and also the only decline in dyad Fluency in the sample. P_0 decreased individual Fluency in Dyad 7 almost as precipitously as did P_2 in Dyad 2, described below, under Conclusion 2. However in this case, the partner, P_7, did not make up the difference. In terms of cognitive resource management, P_7 clearly withdrew resources from Fluency and Diversity to invest them in Speed. This set of priorities was in direct contrast with the instructions given, and strongly indicate that P_7 was not focused on the assigned goals.

Dyads 7 and 1 were the highest in Speed who also managed to finish the post- test. Dyad 1 was unusual in declining on Undefined and showed no improvement in Diversity. In contrast to Dyad 7, Dyad 1 produced the second greatest increase in Fluency, but this change came mostly from P_0. This contrast supports the idea that the activity makes visible different ways of budgeting cognitive resources to the different dimensions of performance.

Dyads 8 and 9 were also very high in Speed, but in contrast to Dyads 7 and 1, disintegrated in the post-test. They showed by far the greatest declines in the sample on

Undefined, but held stable just above the mean on Diversity. Dyad 8 produced the highest increase in dyad Fluency, while Dyad 9 was above the mean. This suggests that the attempt to invest in both Speed and Fluency was too much for Dyads 8 and 9 to maintain, while Dyad 7 emphasized only Speed and managed to finish.

Like Dyad 1, the majority of Dyad 8's Fluency increase came from P_0, while the majority of Dyad 9's increase came from P_9. Dyad 8 posted the highest pre- test score on Speed, but reduced it dramatically, to score near the mean on the post- test. Evidently, this reduction was almost enough to balance out the added load of P_0's increased individual Fluency, as Dyad 8 finished 90% of the post- test. Dyad 9 started high on Speed in the pre- test, and pushed far higher than the rest of the sample on the post- test. This may have combined with P_9 being the major contributor to Fluency increase to cause Dyad 9 to finish only 42% of the post- test.

Dyad 7 and Dyad 3 were the only ones in which P_0 decreased individual Fluency. Unlike P_7, P_3 made up the difference, to generate an increase in dyad Fluency to just below the mean. In fact, P_3 was second only to P_8 on increase of individual Fluency. Dyad 3 showed a marginal decline on Diversity, and Dyad 7 showed a greater decline. Whereas Dyad 7 was among the highest on Speed, Dyad 3 was one of the lowest, supporting a difference of focus between P_7 and P_3. It thus strongly supports alternative choices made in cognitive resource allocation.

Dyads 3 and 4 were the least experienced in the sample. Quite reasonably, they were at and near the bottom on Speed. On the post- test, Dyad 3 was the only one in the sample to increase Speed and also finish, whereas Dyad 4 decreased Speed. Dyad 4 showed the greatest increase in the sample on Diversity, starting from being dramatically the lowest to finishing above the mean, while Dyad 3 remained near the mean. Dyad 3 achieved higher scores than Dyad 4 on Fluency, both pre- and post-intervention. Clearly, Dyad 3 invested more in Speed and Fluency, while Dyad 4 invested very strongly in Diversity, and modestly in Speed and Fluency. This further supports alternative choices made in cognitive resource allocation, and differences of focus.

Dyads 9 and 10 were the most experienced in the sample. They behaved very differently. As discussed above, Dyad 9 was very high in Speed, and disintegrated quite early on the post- test. By contrast, Dyad 10 was among the lowest in Speed in the sample. On Diversity, Dyad 9 was at the mean both pre- and post-, while Dyad 10 was near the mean on the pre- test, and dramatically improved to achieve the best score in the sample on the post-test. On Undefined, Dyad 9 went from the mean to dramatically the worst score in the sample, while Dyad 10 went from above the mean to maximum possible score. On individual Fluency, P_9 produced most of Dyad 9's increase, while P_0 produced most of Dyad 10's increase. As noted above, P_9 shifted resources from Undefined to Speed and individual Fluency, while maintaining on Diversity. P_{10} shifted resources from Speed to becoming the best on Diversity, and to coping with P_0's Fluency.

Comparing Dyads 3 and 4 as the least experienced with Dyads 9 and 10 as the most experienced, Dyads 4 and 10, as the lowest Speed in each experience level, achieved the greatest improvements in the sample on Diversity. This suggests that increasing Diversity, that is, accessing and utilizing the newest knowledge one possesses may pose a higher cognitive load than does increasing Fluency, which can be produced by simply accessing familiar knowledge more frequently.

Self-Organizing Dynamic: This dynamic is interpreted to be maintained by pressures of cognitive load management for two reasons:

1. Where Diversity, Undefined and Fluency improved post-intervention, Speed decreased, suggesting a need for more processing time. Management failures were seen in (a) sacrificing Diversity, Undefined and/or Fluency in favor of Speed, and (b) disintegration of the activity and failure to finish in the post-test when greatly overloaded.
2. Stability in dyad Fluency (totals of transition generation) was maintained by strongly patterned redistribution of transition generation between the two members of each Dyad (Figure 9).

On the global level of P_0's pattern of interacting with 10 different partners, P_0 appears to have adapted his performance to the Fluency conditions generated by each partner (Figure 2). P_0's transition production correlates well with each partner's experience level, while the 10 partners transition production is virtually flat across experience. Thus P_0 clearly regulated his production of transitions to provide each partner with a level of load appropriate to their capabilities. Viewed conversely, P_0 showed his own capacity in the 5% threshold in the partner's transition production, below which he increased and above which he decreased (Figure 9).

On the local level of P_0 interacting with the only partner who truly challenged him (P_2), Dyad 2 was near the mean on Speed and Diversity, and maintained the highest possible score on Undefined. On transition production (Fluency), Dyad 2 started near the mean, and made the third highest increase. As individuals, P_2, who was vastly the highest individual producer of transitions pre-intervention, decreased precipitously, while P_0 more than tripled. Viewed as scores, Dyad 2 appears to be a dramatic outlier. However, when viewed as relative change in scores, Dyad 2's pre/post redistribution correlates strongly with the group (see Figure 9a). Dyad 3 provided the opposite extreme point of this pattern, with P_3 increasing and P_0 decreasing. The fact that Dyads 2 and 3 show the pattern extending into regions in which one partner actually reduces transition production (Figure 9a) is interpreted as additional support for the self-organizing dynamic. The strength of the regressions on these patterns is remarkable, and strongly supports the idea of opposing forces organizing themselves into a dynamically stable system.

Consistent with the anecdotal evidence of the activity's long history of usage, it is important to note that the activity's success as a training exercise depends on the participants viewing the activity itself as cooperative, although its ultimate purpose is skill-development for competition. A crucial meta-lesson learned in the activity is maintaining stability under more and more challenging conditions. It is well-known in engineering fields that a cargo aircraft is designed to be so stable that a stall or a spin is actually difficult to induce, whereas a fighter aircraft exists on the edge of instability to permit the abrupt maneuvers necessary to attack or evade an opponent. The contrast between competition and cooperation, and the need for cooperation in skill development before the loads of actual competition are added, must be strongly grasped. Since the core of competition is to generate instability leading to collapse in the opponent while remaining stable oneself, the fact that this activity would collapse every few seconds if performed competitively would defeat its training purpose.

The primal urge to speed is valid for actual application of the skills being developed in activities such as this. Considering that, in application, few exchanges would go more than

3 cycles, and many would be heavily prejudiced if not decided in the first cycle, speed is important. However, overcoming this valid primal urge is absolutely necessary in a learning situation, because if the time needed for cognitive processing is denied, solid learning is replaced with suboptimal maladaptions, which can be disastrous if habituated, because they are almost always ill considered in consequences and/or mechanics.

The drill-network activity studied here has been shown highly applicable to the needs of several fields. It does measure cognitive load during instruction, which Sweller et al. (1998) and Shea and Morgan (1979) said was uncommon. It meets the criteria stated by Wulf and Shea (2002) in providing both motor and cognitive complexity, and the potential to study performance in overload. It also satisfies their concern about maintaining controlled conditions for the study. It satisfies Starkes and Allard (1993) in integrating study of expert and early-learning performance together, as well as balancing motor and cognitive demands. It is also a true open skill, in that within a framework, the environment is totally generated by the interaction between the partners.

The high dimensionality of the activity shows considerable individuality in how the challenges are selected and met, going beyond what is likely to be anticipated in a conventional design. These were concerns which Battig (1972) referred to as ‘idiosyncracies’ and additional learning processes, Klein (1997) referred to as ‘processes not related to traditional laboratory findings’, and Canon-Bowers and Bell (1997) referred to as ‘decision making that does not follow a set of predetermined steps’. Like WOMBAT, the drill-network provides multiple response alternatives in requiring a response selection at each cycle, enforces distributed attention and adjustment of priorities between the tasks of tracking the partner’s actions and planning one’s own.

The significance of the drill-network in connecting performance studies to neuroimaging studies may be considerable. Attention management in terms of tracking changing salience values in the environment connect to adult learning and expert knowledge. Pattern recognition is important in several approaches to attention management, in maintaining/adjusting purpose, assigning salience, and retention of memory traces. Studies using the drill-network may offer some answers to Endsley and Bolstad's (1994) question as to why a psychomotor task should correlate to SA, through Klin et al.'s (2003) embodied action approach.

Since dyads in this study gave evidence of overtaxing their resources to the point of failure, there is no doubt that the drill-network generates sufficient cognitive load to induce measurable effect, which has been a concern in CI field studies (Brady, 1998). The self-organizing aspect of it provides interaction with individual differences, and adjusts to changing capacities without teacher intervention, both of which were concerns of Sweller er al..

In providing a meaningful unpredictability, the drill-network extends random practice into the real world, because individual cycles are unpredictable to a participant, but integral parts of a developing context. This adds load sources in terms of *n*-back memory in recognition of patterns that will aid prediction through context.

As a curriculum guide, the drill-network records and measures the reproduction of expert knowledge (past experience) of a field (Cannon-Bowers & Bell, 1997; McGarry et al., 2002). It is a direct application of cognitive science's best models of both procedural and declarative knowledge networks, and a condensation of key concepts and critical connections

(Wyman & Randel, 1998). Also, block, sequence and random practice schedules, as used, all collapse under any number of skills approaching real-world value.

Though the first use of the drill-network's structure is unrecorded, it is difficult to participate in the field for long without seeing directly that it is a condensed description of naturally-occurring behavior (Klein, 1997). Likewise, this study's results show that the conditions created by the drill-network lead to processes not observed in traditional practice schedule studies.

In direct response to the Research Questions in Chapter I, this instructional procedure remains in use primarily because it produces a learning experience that is rich in technical content, but also induces a state of high attentiveness, including the natural tension between dealing with unpredictable input while devising one's own actions. This distribution of attention, under the general goal of making the activity more predictable by taking control away from the partner, while maintaining enough structure to the activity to work on correct technique is an unusual effect. Far more common in martial arts practice is one extreme of solo or dyad practice in a block or sequence format, which loses all engagement as soon as the actions are memorized, or totally free practice, in which not-getting-hit reduces technical advancement to mere polishing of the simplest skills, or habituation of a spontaneous invention which is rarely well-conceived. The drill-network provides an insight into human performance which is needed by every student—that learning comes about through patient focused practice and elaborative thinking and discussion which reveal cohesive reasoning within a well-conceived field of study.

One mechanism which may support the anecdotal reports that martial arts training may in some cases improve concentration and self-control is a gradual realization that overload from excessive speed and insufficient attention to method generates unsatisfying performance. This steady redirection to germane issues of cognitive resource management and attentional focus could be very valuable.

CHAPTER VI

SUMMARY AND RECOMMENDATIONS

The purpose of this study was to investigate one instructional activity with a long history of use in traditional martial arts, substantiating its instructional values.

Summary

A literature investigation showed that the drill-network activity was not published in its own right, but was solidly supported by theory and research in seven fields. Careful consideration of the activity itself, in the light of the literature, produced four performance dimensions (Speed, Diversity, Undefined, and Fluency) that were capable of distinguishing individual performances as well as changes in these performances across an instructional intervention.

An experimental trial of the activity was performed by collecting performance data from a small sample of students both before and after an instructional intervention.

Effects of strategies of cognitive load management were visible both within individuals as they prioritized which aspects (Performance Dimensions) of the activity to concentrate on, and between individuals as the balance of higher-level function (measured as Fluency) shifted from one to the other while the total for each dyad remained remarkably stable.

This study’s results strongly support the traditional use of the training activity to teach complex skills along with the highly integrated knowledge base of conditional procedures requisite to those skills. The results also strongly support the adoption of the structure of the knowledge base and the activity as a stable research model of the highly complex interaction of two persons in a dyadic system. While very simple in structure, the activity generates strong, highly individualized learning experiences.

Recommendations for Future Research

There is no more fundamental level of learning than manipulatives—especially one's own body. Indeed, everything a person ever accomplishes is expressed in some manner, through his/her body, even if it is only spoken or written. There is no more fundamental lesson of human relations than learning to disagree productively, and to keep competition brisk and nimble without toppling into breaches of trust which unravel the social fabric which affords us everything we value. The possibility of doing both these things at once, while also practicing timeliness and calm orderly thinking under pressure deserves promotion. The possibility of measuring numerous dimensions of performance, and of performance improvement, while doing these things, should be irresistible.

Of course, replications of this study are recommended. Sample size was sufficient to produce strong statistical significance, as well as effect sizes. So long as a rater reviewing slow-motion video must do data extraction, limiting the sample size would be a boon. Access to a computerized extraction process, such as image-recognition software or motion tracking equipment, would facilitate larger samples. Data collection over a 60-second time period instead of a 120-second time period would probably be sufficient for the measures used here, as speed was demonstrated to be quite uniform, and performance did not change qualitatively to any great extent. And, extension of this study would include finding established tests with which to compare results.

It should be quite clear from its compatibility with concepts in cognitive psychology, organizational design, and the myriad other fields that benefit from the use of network diagrams, that the drill network deserves attention in designing learning systems in other realms, including language (reading, writing), math skills (from arithmetic to advanced statistics and other specialties). In rationalizing, testing and verifying complex patterns of interconnected facts, issues, and/or procedures the multi-dimensionality afforded by a network is probably without equal for, as Schvanveldt (1990) said, a network will encompass hierarchies without be limited to them.

Personal communications with mathematicians have supported the idea that dynamical systems, synergetics, non-linear systems approaches to extracting a mathematical model from this operational model should be practical and rewarding. Collaborations in this direction will be sought.

Personal communications with staff officers in the military, and with professionals directly involved in systems design and performance testing for military personnel have pointed out very direct value to multiple venues. Opportunities to bring these potentials to fruition are being cultivated.

REFERENCES

Allard, F. (1993). Cognition, expertise, and motor performance. In J. L. Starkes & F. Allard (Eds.), *Cognitive Issues in Motor Expertise* (pp. 17-34). Amsterdam: Elsevier Science Publishers.

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance, and other domains. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise: Prospects and Limits* (pp. 126-152). Cambridge: Cambridge University Press.

Anderson, J. R. (1995). *Learning and Memory: An Integrative Approach*. New York: John Wiley & Sons, Inc.

Barber, B. E. (1999). *Ranger School: No Excuse Leadership*. Rockwell, TX: Patrol Leader Press.

Battig, W. F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R. F. Thompson & J. F. Voss (Eds.), *Topics in Learning and Performance*. New York, NY: Academic Press.

Beer, R. D. (1990). *Intelligence as Adaptive Behavior: An Experiment in Computational Neuroethology* (Vol. 6). Boston: Academic Press.

Brady, F. (1998). A theoretical and empirical review of the contextual interference effect and the learning of motor skills. *Quest, 50*, 266-293.

Britton, B. K., & Tidwell, P. (1991). *Shifting novices' mental representations of texts toward experts': diagnosis and repair of novices' mis- and missing conceptions.* Paper presented at the Thirteenth Annual Conference of the Cognitive Science Society, Chicago.

Britton, B. K., & Tidwell, P. (1995). Cognitive structure testing: a computer system for diagnosis of expert-novice differences. In P. D. Nichols & S. Chipman & R. L. Brennan (Eds.), *Cognitively Diagnostic Assessment*. Hillsdale, NJ: Erlbaum.

Cannon-Bowers, J. A., & Bell, H. H. (1997). Training decision makers for complex environments: implications of the naturalistic decision making perspective. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 99-110). Mahwah, NJ: Lawrence Erlbaum Associates.

Chamberlain, C. J., & Coelho, A. J. (1993). The perceptual side of action: decision-making in sport. In J. L. Starkes & F. Allard (Eds.), *Cognitive Issues in Motor Expertise* (pp. 135-157). Amsterdam: Elsevier Science Publishers.

Clamann, M. P., Wright, M. C., & Kaber, D. B. (2002). *Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing*. Paper presented at the 46th Annual Meeting of the Human Factors and Ergonomics Society, Baltimore.

Drillings, M., & Serfaty, D. (1997). Naturalistic decision making in command and control. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 71-80). Mahwah, NJ: Lawrence Erlbaum Associates.

Edwards, B. (1979). *Drawing on the Right Side of the Brain*. Los Angeles: J. P. Tarcher, Inc.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *International Journal of Aviation Psychology, 4*(3), 241-264.

Endsley, M. R., & Robertson, M. M. (2000). Training for situation awareness in individuals and teams. In M. R. Endsley & D. J. Garland (Eds.), *Situation Awareness Analysis and Measurement* (pp. 349-365). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Endsley, M. R., & Smith, R. P. (1996). Attention distribution and decision making in tactical air combat. *Human Factors, 38*(2), 232-249.

Fischman, D. A. (2003). The "descriptive" curse. *The Scientist, 17*(9), 17.

Friendly, M. L. (1977). In search of the M-gram: the structure of organization in free recall. *Cognitive Psychology, 9*, 188-249.

Gagne, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The Cognitive Psychology of School Learning* (2nd ed.). New York: Addison Wesley Longman, Inc.

Han, B. S. (1974). *Hapkido: Korean Art of Self-Defense*. Los Angeles, CA: O'Hara Publications, Inc.

Inosanto, D., Johnson, G. L., & Foon, G. (1977). *The Filipino Martial Arts as Taught by Dan Inosanto*. Los Angeles, CA: Know Now Publishing Co.

James, P., & Thorpe, N. (1994). *Ancient Inventions*. New York: Ballantine Books.

Johnson, P. J., Goldsmith, T. E., & Teague, K., W. (1994). Locus of predictive advantage in pathfinder-based representations of classroom knowledge. *Journal of Educational Psychology, 86*(4), 617-626.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural Knowledge: Techniques for Representing, Conveying, and Acquiring Structural Knowledge*. Hillsdale, NJ: Erlbaum.

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science, 294*, 1030-1038.

Kim, P. S. (1973). *Palgue 1, 2, 3 of Tae Kwon Do Hyung*. Los Angeles, CA: Ohara Publications, Inc.

Klein, D. E., Klein, H. A., & Klein, G. (2000). *Macrocognition: linking cognitive psychology and cognitive ergonomics.* Paper presented at the 5th International Conference on Human Interactions with Complex Systems, Urbana-Champaign.

Klein, G. (1997). *Making Decisions in Natural Environments* (Special Report 31): Research and Advanced Concepts Office, U.S. Army Research Institute for the Behavioral and Social Sciences.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambook, C. E. (Eds.). (1993). *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex Publishing Corporation.

Klin, A., Jones, W., Shultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences, 358*, 345-360.

Krampe, R. T., Engbert, R., & Kliegl, R. (2002). Human movement, timing, and coordination special issue. *Brain and Cognition, 48*(1).

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest, 46*, 328-344.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Focus article: taking stock of naturalistic decison making. *Journal of Behavioral Decision Making, 14,* 331-352.

Lipshitz, R., & Shaul, O. B. (1997). Schemata and mental models in recognition-primed decision making. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 293-303). Mahwah, NJ: Lawrence Erlbaum Associates.

Lipshitz, R., & Strauss, O. (1996). *How decision-makers cope with uncertainty.* Paper presented at the Human Factors and Ergonomics Society 40th Annual Meeting, Philadelphia.

Luttwak, E. N. (1987). *Strategy: The Logic of War and Peace*. Cambridge, MA: Belknap Press of Harvard University Press.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9*, 241-289.

Maier, H. N. (1999). CogPsych foundations for a wing chun curriculum [PowerPoint presentation]: unpublished.

Maier, H. N. (2001, March 21). *Structuring random practice through Markov processes.* Paper presented at the Educational Research Exchange, Texas A&M University.

Maier, H. N., & Taber, K. H. (2003, May 15-17, 2003). *Measurement of initiative in high-speed tactical decision making.* Poster presented at the Naturalistic Decision Making 6, Pensacola Beach, FL.

McGarry, T., Anderson, D. I., Wallace, S. A., Hughes, M. D., & Franks, I. M. (2002). Sport competition as a dynamical self-organizing system. *Journal of Sports Sciences, 20*, 771-781.

O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors, 39*(4), 540-552.

Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein & J. Orasanu & R. Calderwood & C. E. Zsambook (Eds.), *Decision Making in Action: Models and Methods* (pp. 3-20). Norwood, NJ: Ablex Publishing Corporation.

Pascual, R., & Henderson, S. (1997). Evidence of naturalistic decision making in military command and control. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 217-226). Mahwah, NJ: Lawrence Erlbaum Associates.

Pfeiffer, P. E. (1978). *Concepts of Probability Theory* (2nd Revised ed.). New York, NY: Dover.

Rajan, T. V. (1999). Commentary: would Harvey, Sulston, and Darwin get funded today? *The Scientist, 13*(9).

Roscoe, S. N. (1993). An aid in the selection process—WOMBAT. *Civil Aviation Training, 4*(2), 48-51.

Roscoe, S. N. (1997). *Predicting Human Performance*. Quebec: Helio Press, Inc.

Rumiati, R. (2000). Cognitive neuroscience of actions: a whole is greater than the sum of its parts. *Brain and Cognition, 44*, 117-123.

Salas, E., & Klein, G. (Eds.). (2001). *Linking Expertise and Naturalistic Decision Making*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schmitt, J. F. M., USMCR, & Klein, G. A. (1996). Fighting in the fog: dealing with battlefield uncertainty. *Marine Corps Gazette*, 62-69.

Schvaneveldt, R. W. (Ed.). (1990). *Pathfinder Associative Networks: Studies in Knowledge Organization*. Norwood, NJ: Ablex.

Serfaty, D., MacMillan, J., Entin, E. E., & Entin, E. B. (1997). The decision making expertise of battle commanders. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 23-246). Mahwah, NJ: Lawrence Erlbaum Associates.

Serfaty, D., & Michel, R. R. (1990, June 12-14, 1990). *Toward a Theory of Tactical Decisionmaking Expertise.* Paper presented at the 4th Symposium on Command and Control Research, Naval Postgraduate School, Monterey CA.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology*, *5*(2), 179-187.

Spatz, C. (1997). *Basic Statistics: Tales of Distribution* (6th ed.). Pacific Grove, CA: Brooks/Cole.

Starkes, J. L., & Allard, F. (Eds.). (1993). *Cognitive Issues in Motor Expertise*. Amsterdam: Elsevier Science Publishers.

Sweller, J. (1999). *Instructional Design in Technical Areas*. Camberwell, Victoria, Aus.: The Australian Council for Educational Research Ltd.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology review, 10*(3), 251-296.

Taber, K. H. (1999). Asst. Prof., Baylor College of Medicine, Frensley Center for Imaging Research. personal communication.

Tennant, M., & Pogson, P. (1995). *Learning and Change in the Adult Years: A Developmental Perspective*. San Franscisco: Jossey-Bass.

Torrance, E. P., & Safter, H. T. (1999). *Making the Creative Leap Beyond*. Buffalo. NY: Creative Education Foundation Press.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Venturino, M. (1997). Interference and information organization in keeping track of continually changing information. *Human Factors, 39*(4), 532-539.

Watts, B. D. (1996). *Clausewitzian Friction and Future War* (Vol. 52). Washington, DC: National Defense University Press.

Woods, D. O. (2002). *GUTs or no GUTs (Grand Unified Theories): does/can/should cognitive engineering have G.U.T.s?* Paper presented at the 46th Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD.

Wright, D. L., Li, Y., & Whitacre, C. (1992). The contribution of elaborative processing to the contextual interference effect. *Research Quarterly for Exercise and Sport, 63*(1), 30-37.

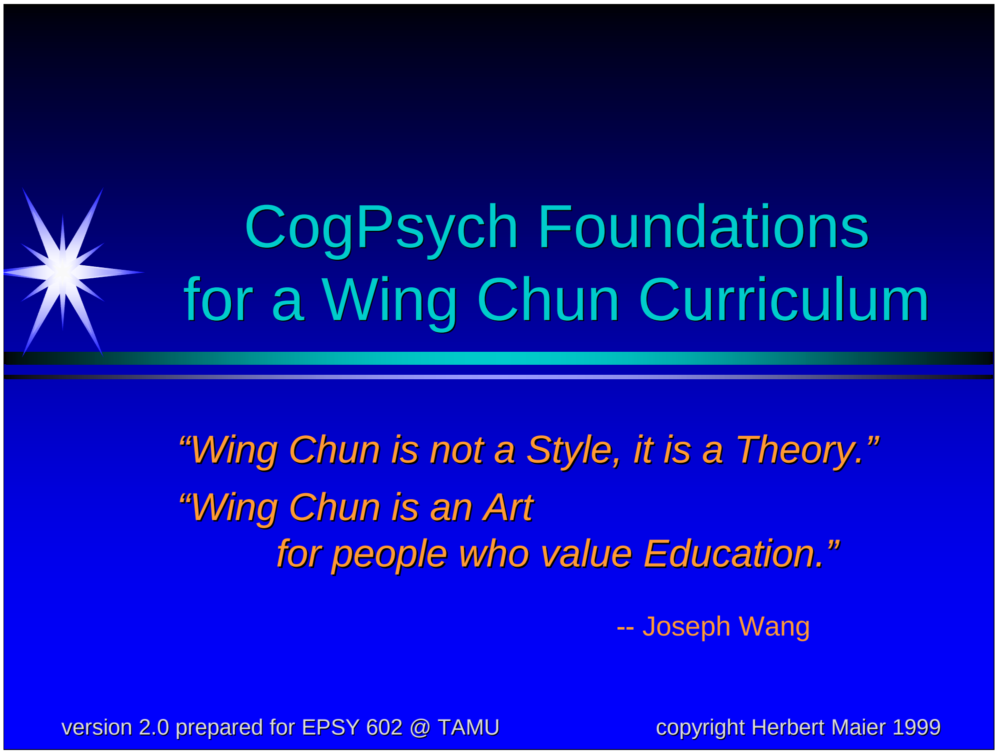
Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review, 9*(2), 185-211.

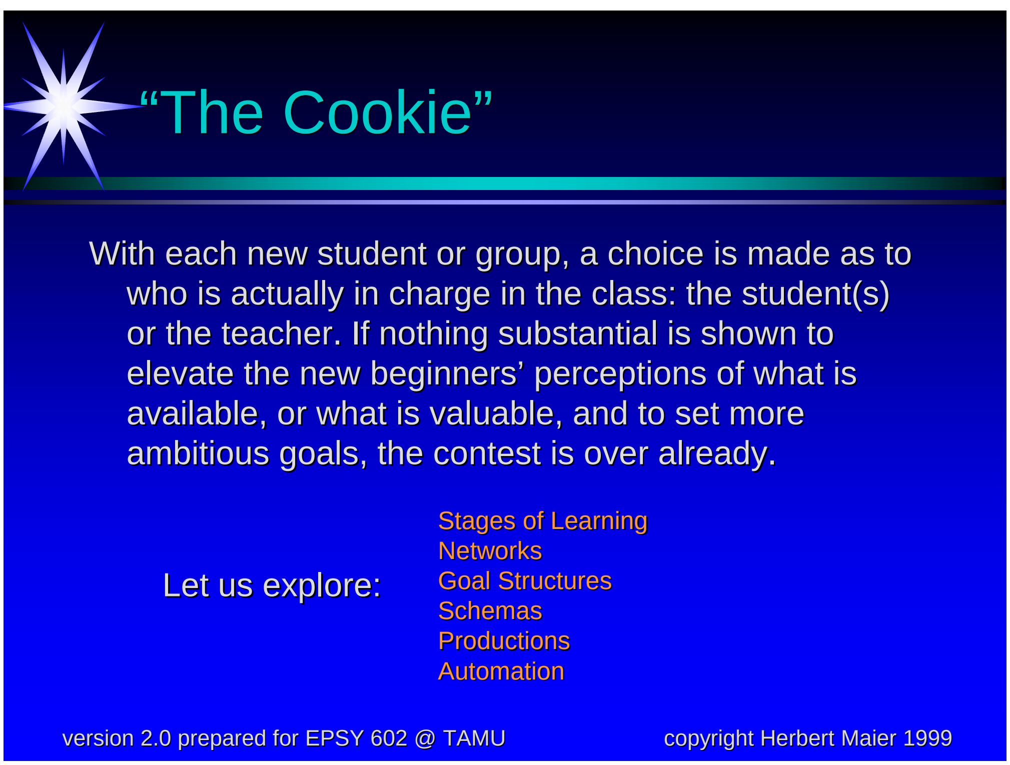
Wyman, B. G., & Randel, J. M. (1998). The relation of knowledge organization to performance of a complex cognitive task. *Applied Cognitive Psychology, 12*, 251-264.

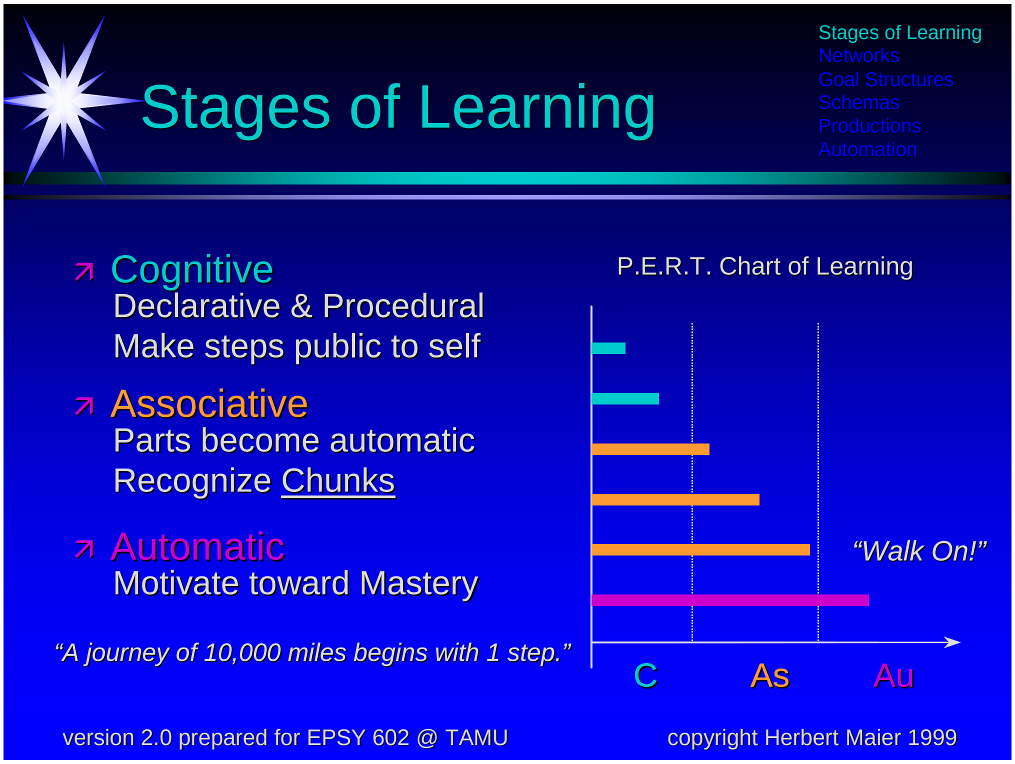
Zsambok, C. E. (1997). Naturalistic Decision Making: Where Are We Now? In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 3-16). Mahwah, NJ: Lawrence Erlbaum Associates.

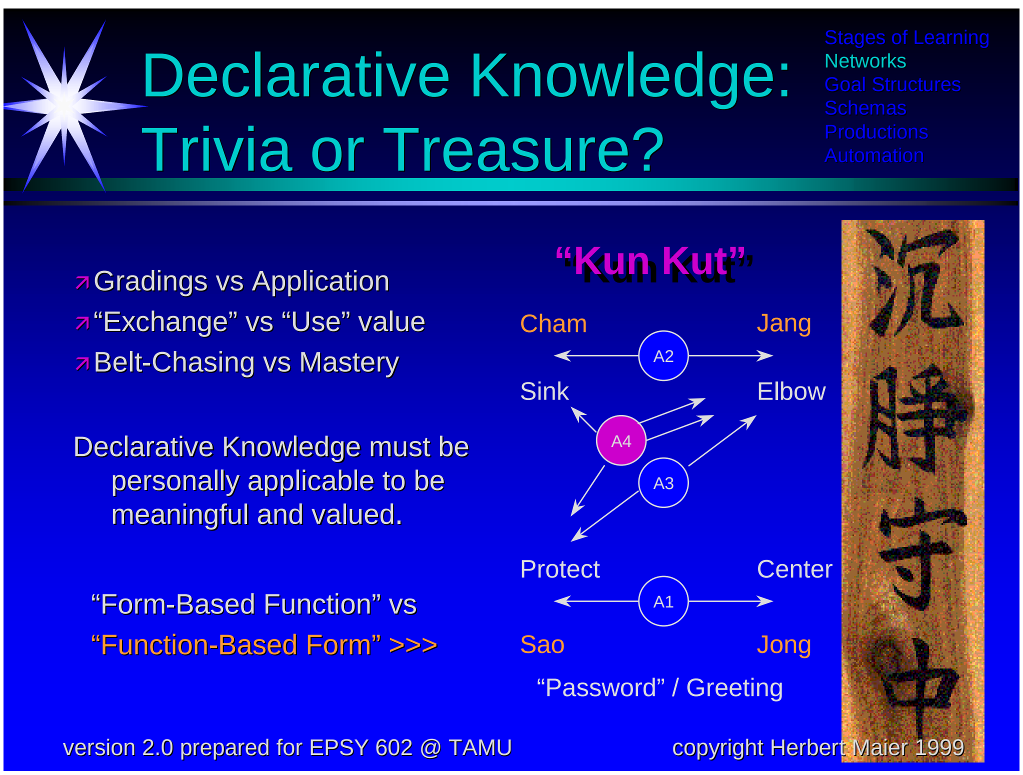
APPENDIX A

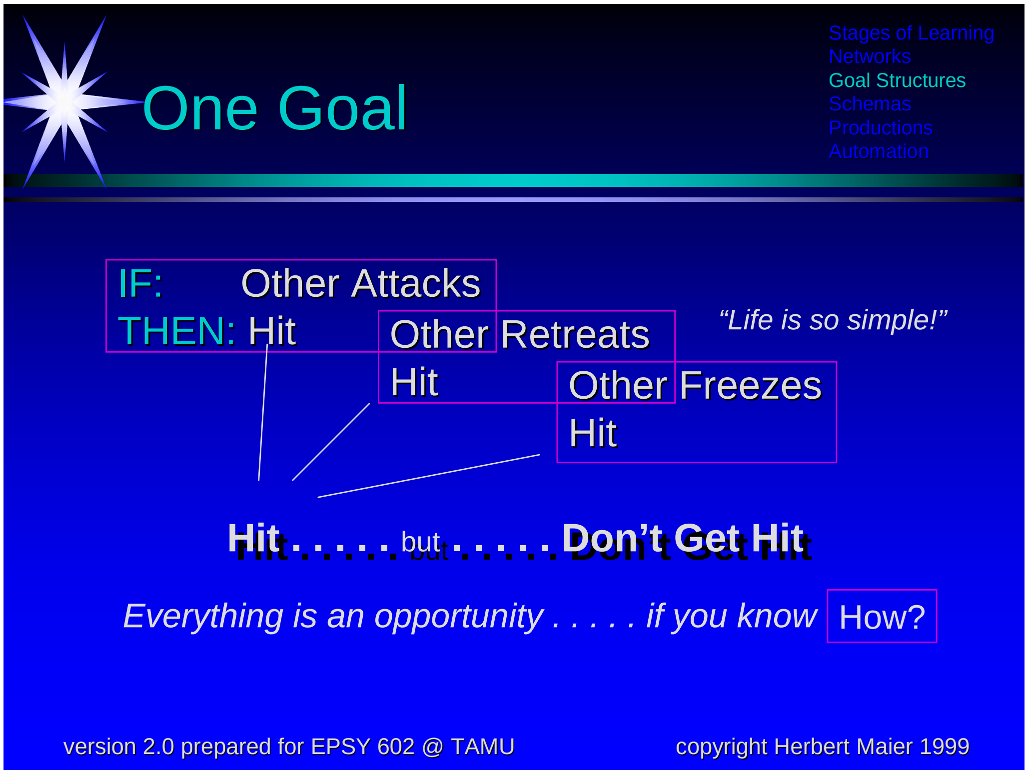
COGPSYCH FOUNDATIONS OF A WING CHUN CURRICULUM

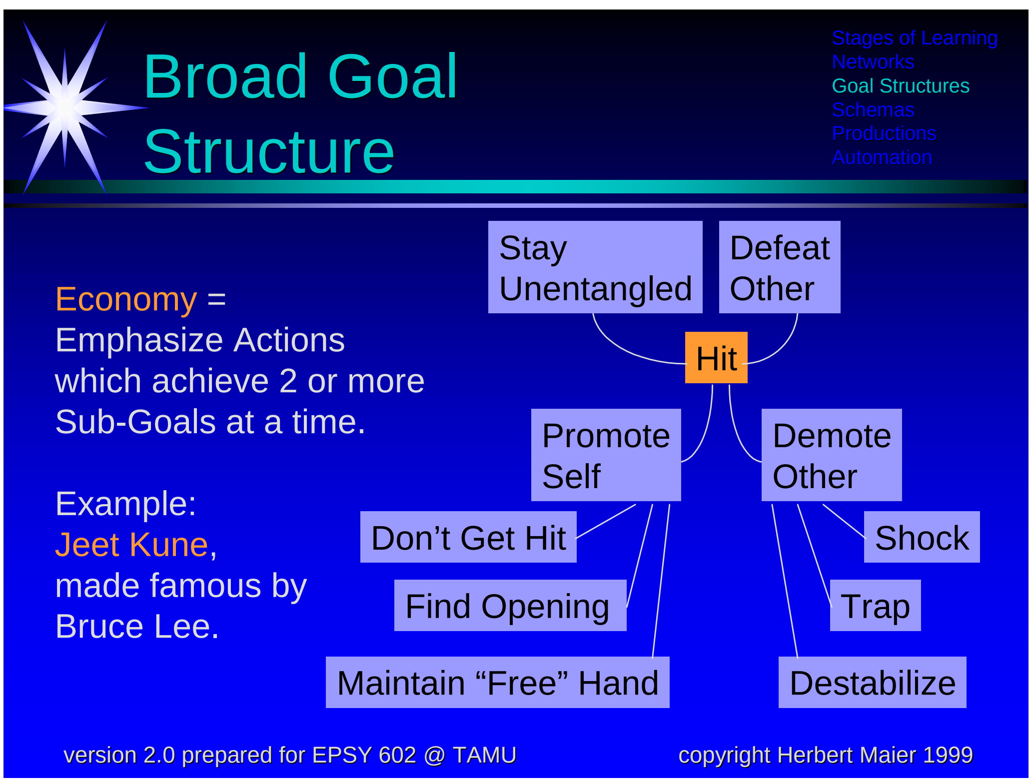


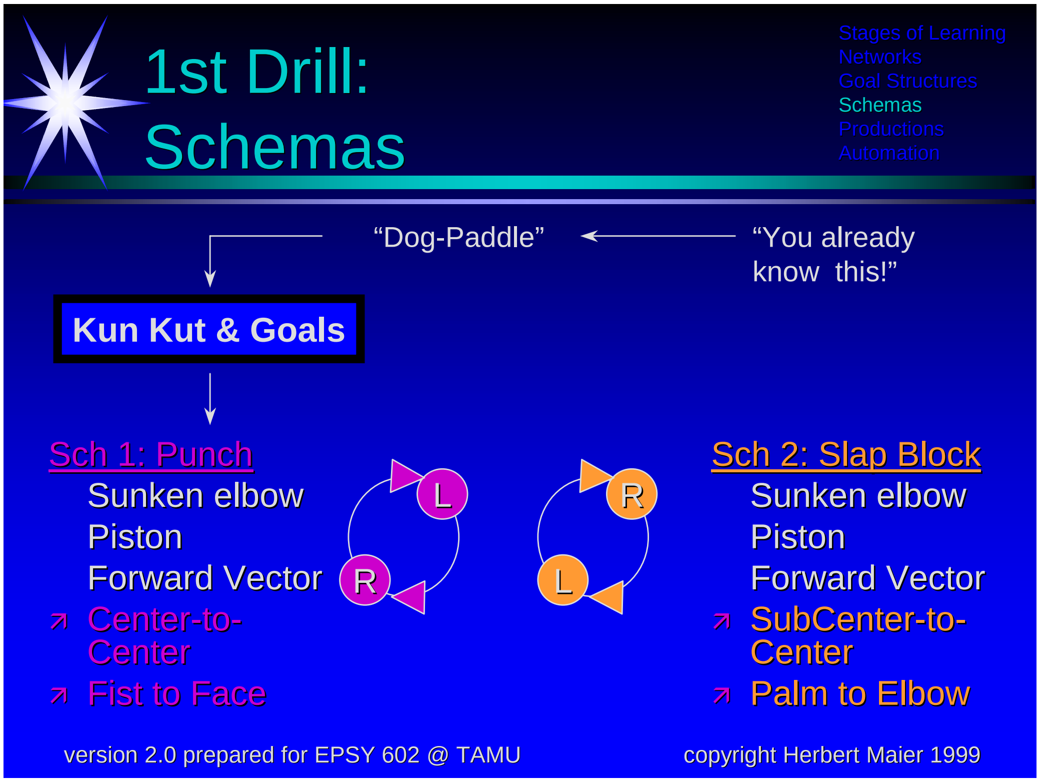


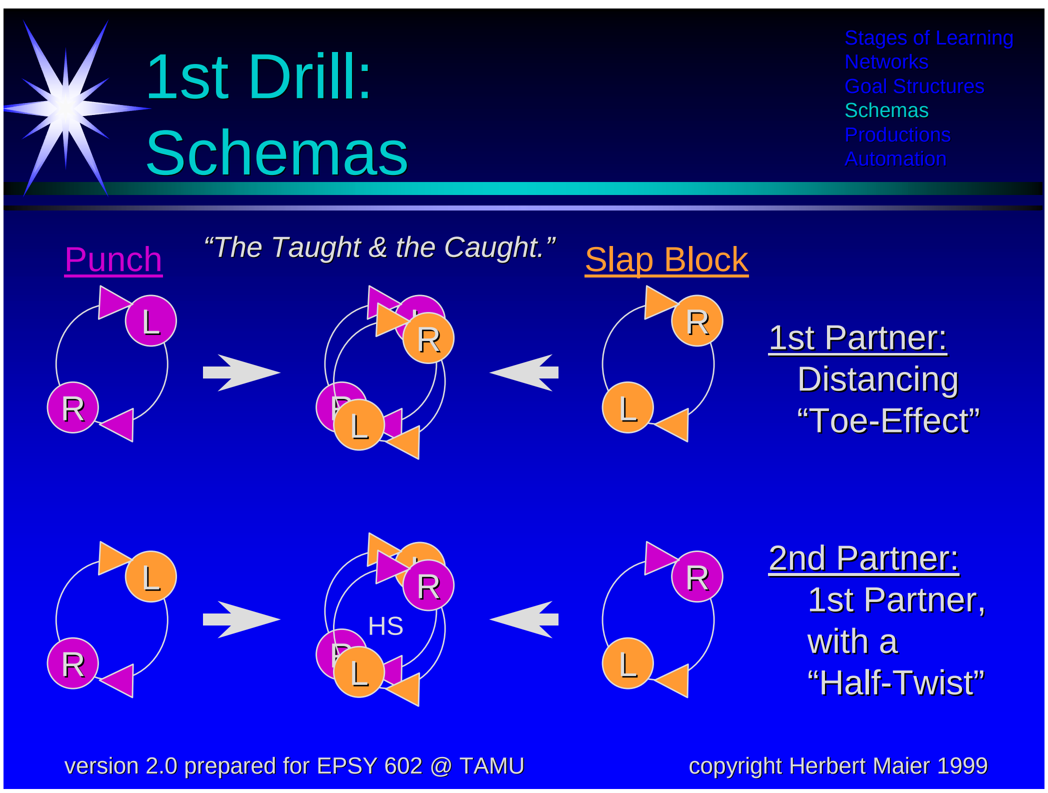

Stages of Learning
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
Cognitive
Declarative & Procedural
Make steps public to self
Associative
Parts become automatic
Recognize Chunks
Automatic
Motivate toward Mastery
"A journey of 10,000 miles begins with 1 step."
P.E.R.T. Chart of Learning
"Walk On!"
C
As
Au
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999

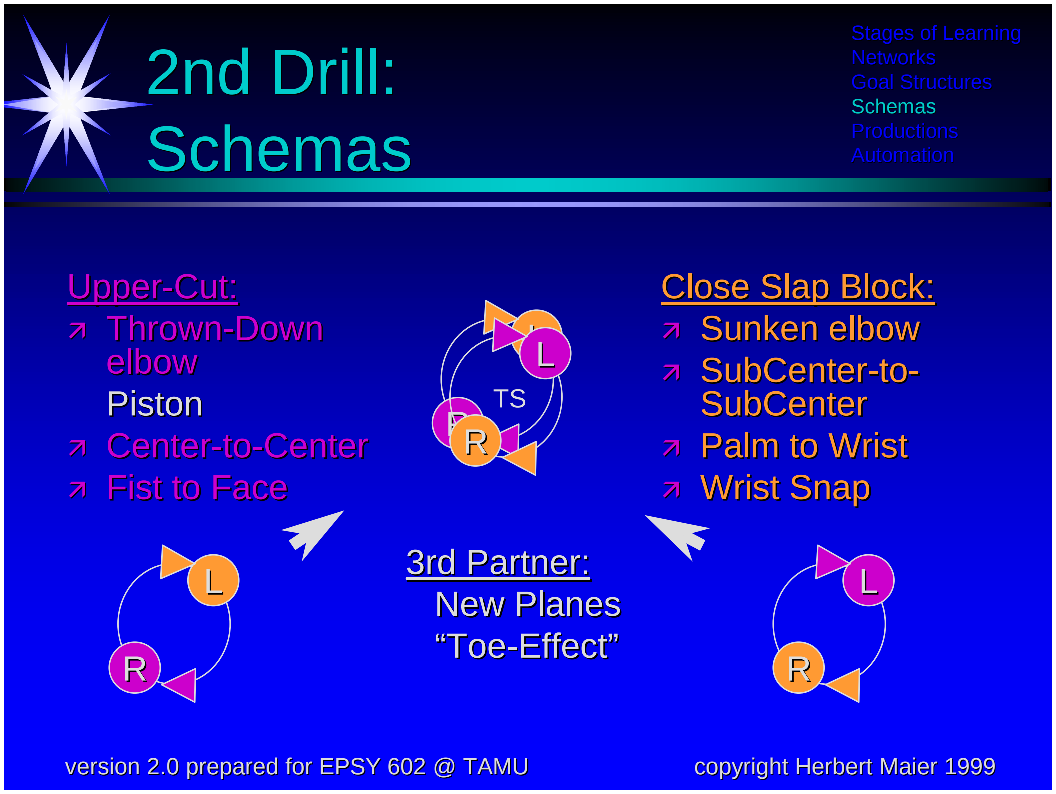

Declarative Knowledge:
Trivia or Treasure?
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
Gradings vs Application
"Exchange" vs "Use" value
Belt-Chasing vs Mastery
Declarative Knowledge must be personally applicable to be meaningful and valued.
"Form-Based Function" vs
"Function-Based Form" >>>
"Kun Kut"
Cham
Jang
A2
Sink
Elbow
A4
A3
Protect
Center
A1
Sao
Jong
"Password" / Greeting
沉静守中
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999

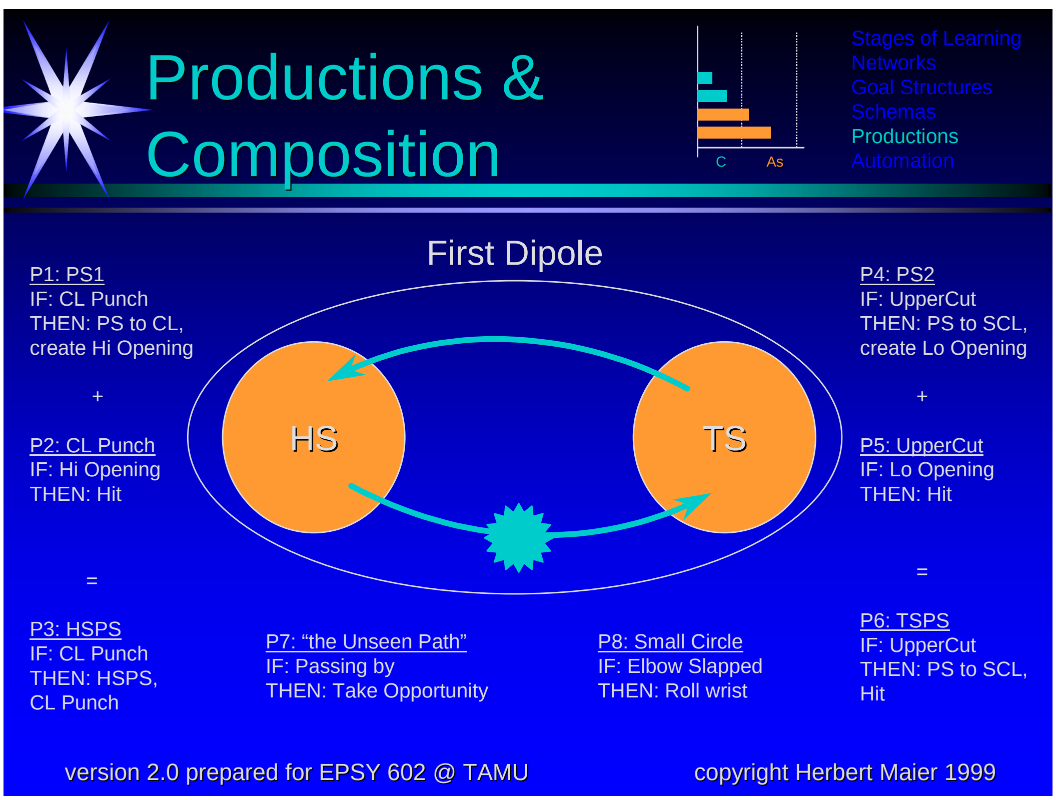

One Goal
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
IF: Other Attacks
THEN: Hit
Other Retreats
Hit
Other Freezes
Hit
"Life is so simple!"
Hit . . . . . . but . . . . . . Don't Get Hit
Everything is an opportunity . . . . . if you know
How?
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999


Broad Goal Structure
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
Economy = Emphasize Actions which achieve 2 or more Sub-Goals at a time.
Example: Jeet Kune, made famous by Bruce Lee.
Stay Unentangled
Defeat Other
Hit
Promote Self
Demote Other
Don't Get Hit
Find Opening
Maintain "Free" Hand
Shock
Trap
Destabilize
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999


1st Drill:
Schemas
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
"Dog-Paddle"
"You already know this!"
Kun Kut & Goals
Sch 1: Punch
Sunken elbow
Piston
Forward Vector
Center-to-Center
Fist to Face
L
R
R
L
Sch 2: Slap Block
Sunken elbow
Piston
Forward Vector
SubCenter-to-Center
Palm to Elbow
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999

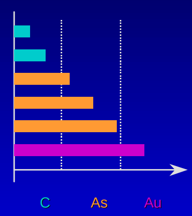

1st Drill:
Schemas
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
Punch
"The Taught & the Caught."
Slap Block
1st Partner:
Distancing
"Toe-Effect"
2nd Partner:
1st Partner,
with a
"Half-Twist"
HS
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999


2nd Drill: Schemas
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
Upper-Cut:
Thrown-Down elbow
Piston
Center-to-Center
Fist to Face
TS
L
R
Close Slap Block:
Sunken elbow
SubCenter-to-SubCenter
Palm to Wrist
Wrist Snap
3rd Partner:
New Planes
"Toe-Effect"
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999

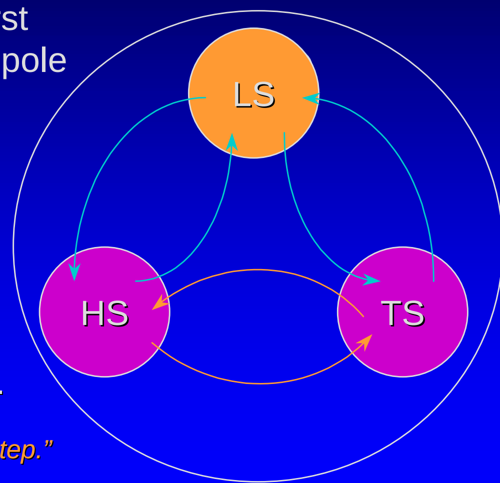

Productions & Composition
C
As
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
First Dipole
P1: PS1
IF: CL Punch
THEN: PS to CL, create Hi Opening
+
P2: CL Punch
IF: Hi Opening
THEN: Hit
=
P3: HSPS
IF: CL Punch
THEN: HSPS, CL Punch
HS
TS
P4: PS2
IF: UpperCut
THEN: PS to SCL, create Lo Opening
+
P5: UpperCut
IF: Lo Opening
THEN: Hit
=
P6: TSPS
IF: UpperCut
THEN: PS to SCL, Hit
P7: "the Unseen Path"
IF: Passing by
THEN: Take Opportunity
P8: Small Circle
IF: Elbow Slapped
THEN: Roll wrist
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999

Complete Sub-System

Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation



First Tripole



- In similar manner, 3rd element is constructed and linked-in.
- Complexity increases quickly. Simplified to show basic structure.

"A journey of 10,000 miles begins with 1 step."

version 2.0 prepared for EPSY 602 @ TAMU

Walking the Long Road

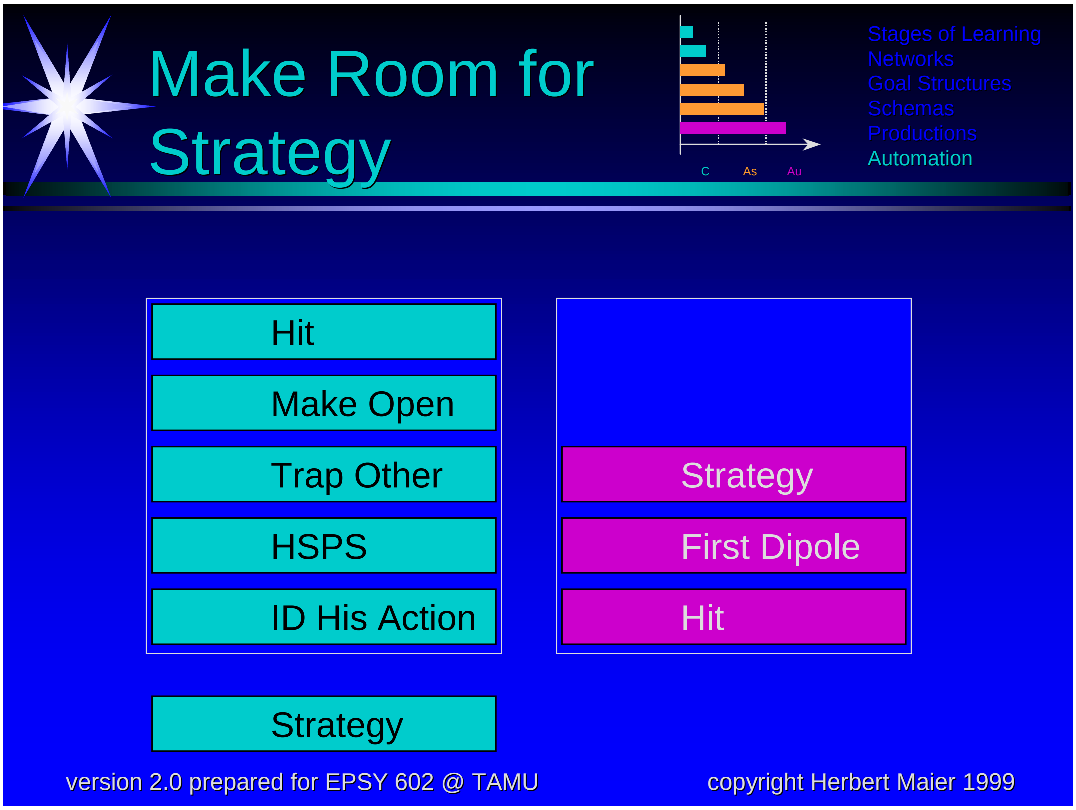
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation

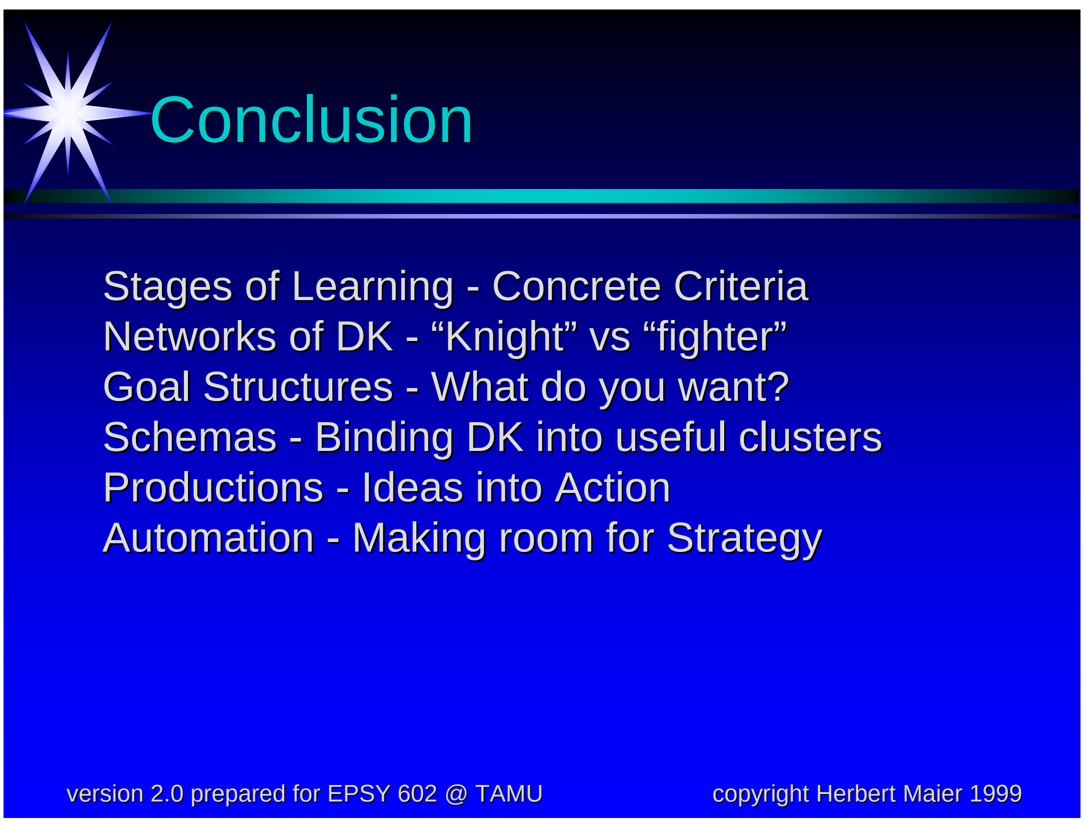
Learning mixes the flour, salt, water & yeast.
Training kneads it into a tough, resilient loaf.
In at least one language, "to eat"
is the same word as "to think on".
There is wisdom in that.

A **well-taught** student is good steel.
A **well-trained** student is a good sword.
Fine craftsmanship deserves
more honor than it gets.

It's worth the trip to . . .

version 2.0 prepared for EPSY 602 @ TAMU


Make Room for Strategy
C
As
Au
Stages of Learning
Networks
Goal Structures
Schemas
Productions
Automation
Hit
Make Open
Trap Other
HSPS
ID His Action
Strategy
First Dipole
Hit
Strategy
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999


Conclusion
Stages of Learning - Concrete Criteria
Networks of DK - “Knight” vs “fighter”
Goal Structures - What do you want?
Schemas - Binding DK into useful clusters
Productions - Ideas into Action
Automation - Making room for Strategy
version 2.0 prepared for EPSY 602 @ TAMU
copyright Herbert Maier 1999

APPENDIX B

STRUCTURING RANDOM PRACTICE THROUGH MARKOV PROCESSES

Structuring Random Practice Through Markov Processes

Herbert N. Maier

Texas A&M University

presented at:

Educational Research Exchange

Texas A&M University

2001

Abstract

Well-documented, though counter-intuitive, benefits of contextual interference produced by random practice can be implemented by consistent use of Markov Processes throughout curriculum design, administration, and evaluation. Compatible with current theory in stages-of-learning and in expertise, this method is shown and discussed through heuristic example.

Structuring Random Practice Through Markov Processes

Personal teaching experience led many years ago to the empirical judgment that block practice of skills alone did not provide for real-world utility. A transition to random practice was clearly essential to integrate isolated skills and develop the technical knowledge, analysis and judgment necessary to make any but the simplest skills functional under stress. The sense of fragmentation perceived when in a block-intensive process, and the struggle to "grasp the formless", is a common experience. Bridging this gulf, and structuring without containing is the goal. Learning should begin as it needs to continue; so, the challenge is the seeming paradox of developing a method to model and measure an activity that quickly becomes pseudo-random[1].

Magill and Hall (1990) reviewed research into the influence of contextual interference on motor skill acquisition. They describe an important demonstration by Battig and Schild, based upon master's theses work by Schild, of the counter-intuitive effect of interference increasing retention and transfer performance for verbal skills. Though most work since 1990 is in the motor skill arena, this work has further confirmed that random practice improves long-term retention and transfer over block practice, at the price of poorer performance early during acquisition.

In order to encourage random practice where it may be under-utilized, a design/measurement tool sensitive to this particular quality of activity is proposed. Markov processes are a mathematical tool which has come to be used to model many topics, including physical processes and computer operation. This paper proposes their use in design/modeling, tracking, and evaluating a skill-based learning task. First, Markov processes will be described in general, and in context of a heuristic task, demonstrating the validity of the method using simulate data, based upon several years of empirical usage. Second, aspects of curriculum design including order of instruction, time scales and level of treatment will be discussed, as well as advantages of the detailed task decomposition basic to this method. Literature discussion will include the use of Markov processes in teaching. Comparisons with past studies in random versus block practice, and in expertise, will further the idea with existing theory and research.

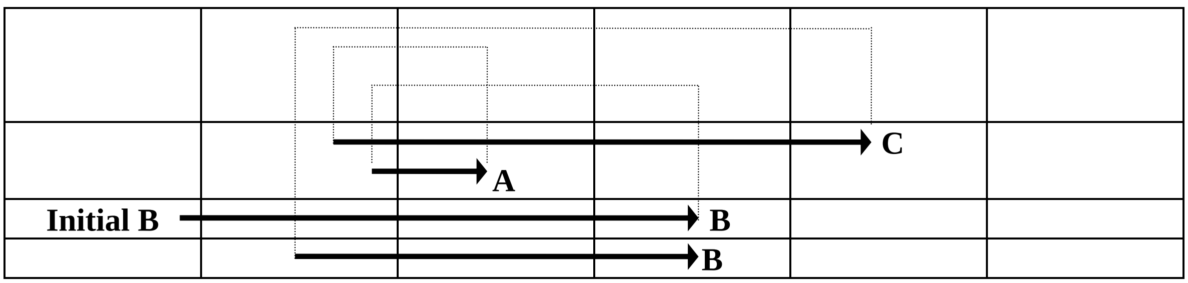
What is a Markov Process?

A Markov Process is a compact method of modeling an array of dependent probabilistic events, which can exhibit a wide, but describable variety of specific instantiations. This array is shown most often as either a transition matrix (Figure 1) or a node-and-link graph (Figure 2). By using a Markov process, a limited view of the dynamics of a system cn be shown. Markovs are used in many fields to model probabilistic events. According to Pfeiffer (1978), a Markov chain is denoted as a sequence in which each event may be conditioned or effected by the one immediately previous, thus giving a simplified but worthy model of a system with some memory. Considering the dognitive effort involved in learning a complex skill, and the fact that almost all attention is directed at getting the next step right, one-back memory provided by a Markov is quite sufficient for present purposes. Each of the basic drills to be discussed will correspond to a state of the system. Since the collection of skills at any time is limited to those learned, this collection or sated space is finite. Any collection greater than 1 provides a selection of initial states, and therefore a probability that one particular initial state will be chosen. Thus even the initial state is dependent upon prior

Initial Probability	->	Activity A	Activity B	Activity C	Total Probability in Row
0.4	Activity A	0.6	0.2	0.2	1.0
0.3	Activity B	0.2	0.7	0.1	1.0
0.3	Activity C	0.1	0.2	0.7	1.0

1.0 Total Initial

a) Sample Markov Transition Matrix



b) Sample Chain

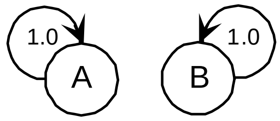
c) Chain shown by arrow in b: B, B, A, C, B

A Markov matrix is enterred through one of its rows, which is selected probailistically. The total of these initial probailities equals 1.0. Use of a 10-sided die is a convenient heuristic as its outcome translates easily into percentages. Each row of the matrix also contains a distribution of probabilities, also totaling to 1.0. Each succeeding throw of the die selects a column to be read in the current row. The nominal value at the top of that column determines the row to be occupied next. Movement through the matrix continues for either a pre-determeind number of throws, or until a particular pre-selcted cell is occupied. The many possible paths to this end-state are called chains. A Markov process is said to have a memory because each event is non-independent of the previous one. Each event selects a row which determines the distribution interpreting the die-throw for the next event. There are different sets of circumstnaces in each row or system-state which influence the outcome of the event at that point in the chain. The matrix contains, in a compact way, all possible chains in the given system.

a) Small arrow shows that the table describes transitions goin in one way only: from the stat listed on the left to the state listed on the top. b) shows one possible chain of events in this system. The string of characters in c) shows the same chain as does the matrix.

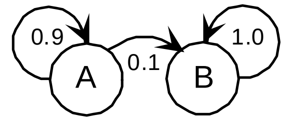
Figure 1.

→	A	B
A	1.0	0.0
B	0.0	1.0



a.

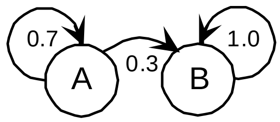
→	A	B
A	0.9	0.0
B	----	1.0



b.

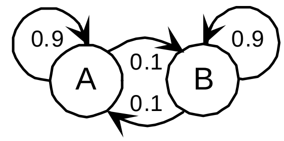
Markov processes are commonly illustrated in a node-link form derived from graph theory. a) in two isolated cyclic drills, numbers show probability of following a particular path, having departed the associated node. b) beginning of network formation with unit-directional linkage.

→	A	B
A	0.7	0.3
B	-----	1.0



c.

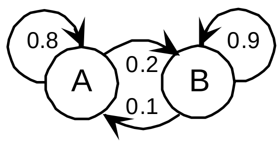
→	A	B
A	0.9	0.1
B	0.1	0.9



d.

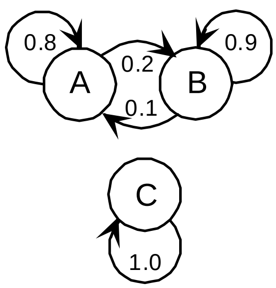
c) Example of probability mass shift to higher frequency of shifting as opposed to staying. d) new B->A link added, with commonly seen deterioration of newest previous knowledge (A->B link).

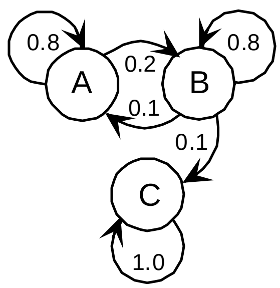
→	A	B
A	0.8	0.2
B	0.1	0.9

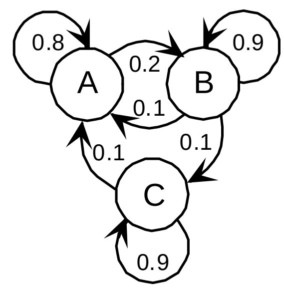


e.

e) Increased stability in the new learning, shown by partial recovery of A->B probability mass lost in progressing from c) to d), along with development of B->A.







→	A	B	C
A	0.8	0.2	----
B	0.1	0.9	----
C	----	----	1.0

f.

→	A	B	C
A	0.8	0.2	----
B	0.1	0.8	0.1
C	----	----	1.0

g.

→	A	B	C
A	0.8	0.2	----
B	0.1	0.8	0.1
C	0.1	----	0.9

h.

f) New drill C added in isolated form, as the earlier ones were. g) new B->C link added. This draws mass away from B, lowering it from 0.9 to 0.8. A new level in complexity has been reached, as B now has 3 choices in where to go. h) new C->A link added. It is now possible to traverse the entire network, though only in limited ways. Dashes in matrix show link that is missing, as opposed to one that is known, but with 0.0 probability.

Figure 2.

conditions. The probabilities of starting in any particular state, or shifting from one state to another, change with training.

In recent years, interest in probabilistic and statistical systems has been accelerated by the availability of computers to do the heavy work of computation. As fields of applied mathematics have grown, dominant concepts have included calculation-intensive ones such as iteration and pseudo-randomness. Iteration means simply that a process's result returns again as its next input. The results of simple iteration based on different starting or seed values ranges from terminally boring to wildly outrageous. Iteration underlies techniques used to model many real-world events, from weather patterns to cardiac rhythms, fluid systems and electronics. Iteration is at the heart of a Markov process.

Instruction as Construction of a Markov Process

It is common in the developmental period of a data-analysis or information theory project to treat the source of data as a "black box", as will be the case here. To make this theoretical discussion concrete however, a heuristic example will be described. The structure shown is in actual use, and will eventually be the subject of an experimental paper. However, it is important to the generalization of this method, and supported by information theory (e.g., (Shaw, 1984)), that actual specific activities being charted do not matter. So long as general reasonableness is observed, specifics must be fairly interchangeable to prove the point.

Selection of the first few skills taught in this method is non-trivial, being multidimensional. The long-term integrity of the structure being built depends upon these first few skills setting task and goal structures for work that will not be encountered until much later in the curriculum. For greatest benefit in a complex field of knowledge, design of this introduction to the system must be iterative, and reverse-engineered from later foals with thought and patience. At the earliest stage of learning, practice in this example consists of one, then two, isolated drills (Figure 2a), each of which provides repetition of a single skill-set. It was found empirically that at least the first two drills are easier if they are close variations of each other. Taking advantage of schema theory in this way eases the burden of earliest learning and shortens the distance of the first shifts. This earliest pre-linked stage is held together by similarity and defined by discrimination. A similar condition exists whenever a new drill is added (e.g., see Figure 2f), but is especially crucial here. At this point, the Markov state space for each drill consists of one state. There are not yet any transition at the drill level, so the probability of returning to the same drill at the end of a cycle is unity (1.0). There is much literature support for this effect under the concepts of schema formation and refinement (e.g., Gagne, Yekovich, & Yekovich, 1993). Schema structure also provides another dimension of connection, while the procedural network is being built and composed.

When the first link skill is taught, the quality of the structure changes dramatically. In this example, the link A->B is chosen, based upon its being the simpler of the two, involving no new skills, but only noticing an opportunity that is passed up every time A is repeated (Figure 2b). de Bono (1970) discussed this situation excellently (p. 42). Now the possible options are to start in B and remain there, or to start in A and remain there until a conscious effort or an external force initiates the link-skill, and a shift is made. As the effort becomes manageable, the link is used more often. Because the total probability-mass must still add up to 1.0, the mass of probabilities redistributes away from repetition toward the link (Figure 2c).

At this point, B acts as what is called an absorbing or trapping state, since the system can not depart it. Learning the second link-skill makes possible a shift from B to A. Figure 2d shows a typical beginner-level balance of repetition and shifting. It is not unusual to see earlier skills drop off at these times, considering the short-term cognitive demands of the newest skill. Through this process, links are shown to be an important category of knowledge in themselves, equivalent to nodes. At this early stage of learning, several repetitions are being spent contemplating and “setting-up” the shift. This mental imaging, part of the burden of simultaneous performance and planning, and thus a source of contextual interference, is a vital part of learning and of training.

The next addition is a new drill, referred to as C (Figure 2f). C, as were A and B, is introduced as an isolated drill, with new skills integral to its performance. Whereas A and B were closely related, C is the beginning of divergence. It seems to work well on an empirical bases to teach new drills as self-contained units of block practice, incorporating them into the flowing random practice network as soon as general familiarity is attained. In this way, attention is focused on one level of work at a time, but kept in motion and fully occupied.

Following the tnem of making every step in learning as small as possible, one available link between B and C (B->C) is taught first (Figure 2g). In order to facilitate immediate exploration through and familiarization with the entire network as it grows by removing any trapping or absorbing states, directionality of each link must be considered. This level of complexity is sufficient to illustrate the value of the design process including the graph-theory representation of the system as well as the matrix. Thus, the next link added travels from C to A (C->A, see Figure 2h). Adding the C->A and A->C links (not shown) completes the basic structure. This is only the beginning of the training system, but sufficient to show the direction.

Order of Instruction and Level of Treatment

The order of construction of a system is a significant factor in teaching: 1) to emphasize the underlying connectedness of a knowledge field, and thereby to deal with the doubts of “what is this for” or “what does this relate to”, and 2) to maintain sufficient stability and functionality at any given point in learning to facilitate and encourage the practice that leads to skill development.

Initial task analysis, based as always upon a set of priorities and values, provides definitions of skills A, B, and C above. Further task analysis of each of these skills implies another level of treatment, in that each task is decomposed into smaller sub-tasks—a process to which reasonable limits must be observed. Selection of a primary scale involves several factors. Stepping to a smaller scale than that shown in Figure 2, each of these drills is a system or chain in itself. State Aa and Ab are shown in Figure 3b to make up complete cycles and Markov processed in themselves. When put together to form the seed for A, there is actually a branching, and therefore a probability mass distribution even at this smaller scale. This line of thinking is consistent with theory involving compositions of skills (Anderson, 1995).

Initial exposure to the field could be at the individual “atomic” action level instead of the drill-level. But interest is most fragile at the beginning stages of learning.

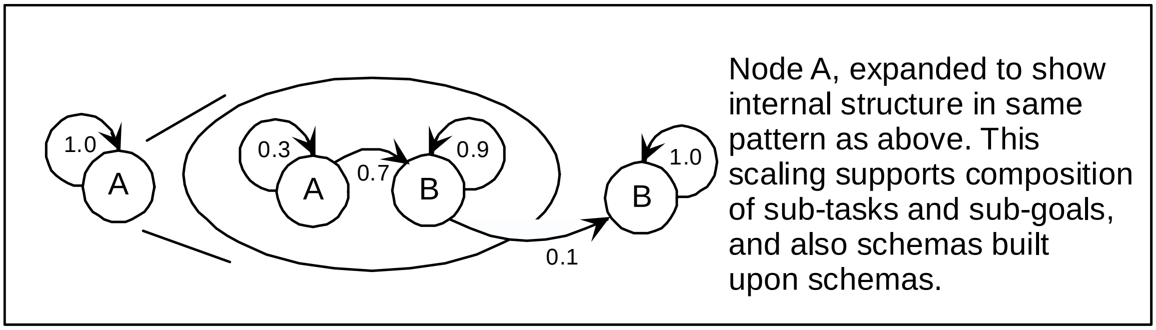


Figure 3.

Full solution of a simple problem gives a sense of validity to the student's effort. Drill-level is also the level at which most decisions will be made for some time. New sub-skills must be learned as new drills are constructed, so scale will be shifted frequently. Taking one cycle of a pattern as the unit of measure (1.0) allows fractional and multiple repetitions of that cycle to be discussed conveniently. These considerations are quite compatible with Anderson's cognitive theories of automation and composition, and Schmidt's (1988)motor schema theory.

It has been shown in Figure 3 that once a network as small as two or three skills is constructed, attention is shifted up another level by focusing on strings of skills, in Markov terms called chains (Figure 1c). In cognitive terms, composition of skills describes this chain becoming a unit, which can be chained at the next level, and so on. This encourages the consideration of "sentencing" and syntax, and the application of skills to more realistic problems. It also directs attention to branch points, and therefore to judgment and foresight. Prioritization of chains used as teaching examples can be based upon high total intended probability-masses, indicating most commonly encountered sequences, or upon important branch points which may represent major issues in the field. As the chains are shown to cross, merge and split, forming a describable network, an appreciation of the integrity of the structure is established. These points deserve being made explicit in the teaching. Related points are picked up again under Expertise and Stages of Learning.

Markov as Evaluation Tool

What has just been used to design and implement a curriculum can also be sued to evaluate skills as they develop. Statistics needed to generate the probabilities used in the matrices and in the graphs can be gathered from video footage, computer files or other suitable records of a pre-determined period of performance. Such a metric can act as a constantly-evolving “fingerprint” of a student as training progresses. Carrying forward the same structure in evaluation as in design makes possible “testing to the teach” rather than teaching to the test”.

The use of the Markov matrix as a performance metric is illustrated by the shift of probability mass from the diagonal (repeating) cells outward, increasing the variation and pseudo-randomness of practice. This shift is hypothesized to represent a profound step in development, though discussing its cause or mechanism is not the purpose of the present paper. However such a shift may be explained, its net effect certainly implies less effort dedicated to concentration and set-up or contemplation time required, making more resources available to judgment. An outward shift of probability mass also “flattens” the matrix be reducing its highest values. This can extend to a figure-ground reversal (Figure 4). As “traffic” moves “off the highway” and out “into the neighborhood”, single and even

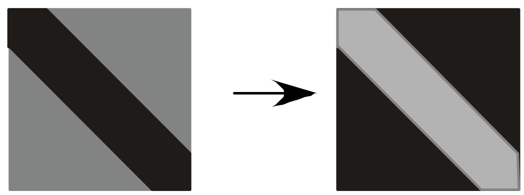


Figure 4.

fractional expressions of each state (mentioned above in terms of Figure 3) become common, and the artifice of node and link begins to disintegrate.

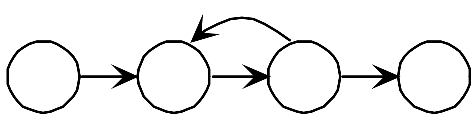
The purpose of this paper is to explore the possibilities of a tool at its most generic level. Specific applications would be likely to place specific targets on probability-mass values rather than encourage unlimited shifting. Some of these target values may well be 0.0 to indicate a shift that is undesirable; in such a case, a measured value above 0.0 would indicate a misconception. Such targets might be either criteria-based or norm-based. Criteria-basing might use initial designs based upon a teacher’s estimation of desirable distributions, considering both relative value and relative difficulty of each link. Norm-based standards would follow methods used in other studies of expertise or knowledge structure to measure representative field experts.

Thus a Markov process provides quantifiable measurement of highly complex behavior. This allows considerable study and concrete discussion, such as 1) comparison of two persons at one moment in time, and 2) comparison of one person at two time-points, showing development, including pruning and possible forgetting or tunnel-vision.

Related Research

Past Uses of Markov Processes in Teaching

Only one published use of Markov processes in terms of designing or evaluating a learning situation has been found. A Markov process has been used to describe and measure the learning of mathematical model generation, thus demonstrating the credibility of modeling a specific activity with the goal of obtaining measures of the difficulty of steps in a learning process (Voskoglou, 1994, 1995a, 1995b). Since Voskoglou is a mathematics educator, the process studied was that of students learning to create and use mathematical models for solving problems. Testing occurred with 2 groups of 20 students from different departments. Results showed that the second and third states in the chain of five (Figure 5) had the highest difficulty for both groups, though to differing degrees. More students either quit or repeated steps at these points than at other steps in their problem-solving process. The metric Voskoglou generated involved the mean number of shifts made in the five-state chain, indicating the number of times a step failed and was re-done. This generated a differential score for the two groups in which the lower score showed greater facility in the learned skill.



A 4-node markov chain. Note directionality of links. Note also unitary start and end points, which distinguish a chain from a network.

Figure 5.

Voskoglou's work is similar to this paper's focus in measuring success in learning, but differs in that it focuses on measuring level of difficulty of specific nodes in a defined chain as compared to incorporation of multiple alternative—"depth" versus "breadth".

Contextual Interference and Practice Schedules

It has been found empirically that in skills of any complexity, a short initial phase of block practice is usually necessary, just to establish a minimal performance level sufficient to withstand the additional challenge of pseudo-randomization. Late block practice periods are also often rewarding in terms of focusing on a sub-skill, or increasing initial understanding. What is at issue here is the investment of the majority of practice time, its definition and carrying-through of purpose.

Recent literature agrees with Magill and Hall's (1990) review mentioned earlier. Literature consistently indicates that random practice variously enhances retention, transfer, recall, and/or recognition at the expense of short-term performance during skill-acquisition. It also shows that early losses in random condition performance during the acquisition phase come to approximately equal block condition performance by the end of the acquisition phase. The mechanism, as discussed by Li and Wright (2000), seems to be contextual interference (CI), in that changes defining random practice act as interference, increasing the cognitive load during practice. As discussed by Wulf (1991), differing theories on mechanism of influence from this CI involve either multiple and variable processing increasing elaboration, or forgetting between matching repetitions requiring more restructuring of the solution (implying abstracted, as opposed to literal memory). Shea and Morgan (1979) recommend teaching a number of skills at each session, consciously exchanging early acquisition performance in favor of robust retention and transfer.

Shea, Kohl and Indermill (1990) tested "rapid force production" and "impact of increased practice" with general results similar to other studies, but showing a possible benefit in starting in block practice and shifting to random practice, even in a task inherently simple, when it is made more complex by short inter-task interval. "Very early in practice, it is difficult for subjects to determine the appropriate strategies when faced with random contexts" (p. 153). This is taken to agree with what is said above in connection with Figure 2c about the burden of simultaneous performance and planning. Shea, Kohl and Indermill suggest that this may also support 1) a schema structure of memory, especially as the mechanism of benefit in acquisition practice "may change as a generalized motor program is developed and subsequently reined" (p. 148)., and/or 2) a stages-of-learning model, especially in what Fitts calls the cognitive stage, or Adams calls the verbal-motor stage. This is discussed further below.

Sources of CI are a study in themselves. Sources of CI from complexity in a task include multiple-movement sequences (dance), detailed movement (golf swing), and compressed time (martial arts). These main effects and/or interactions between them may differentially influence factors such as the switch from block to random practice noted by Shea, Kohl and Indermill. Other sources of CI include the size of the pool of movement trained at one time, which influences the number of variations encountered before a repetition, also noted by Shea, Kohl and Indermill. More will be added in the next section.

Shea (2000) suggested that excessive initial block practice can not only lose its benefit, but become an actual detriment through an increased need for later restructuring. This is taken to agree with the present thesis that in the interest of long-term gain, learning should begin as it needs to continue, with minimal block and maximal random practice, building a skill-structure not only more complex, but probably also more receptive to further growth and change.

Experimentally, random practice means shifting between limited specifics in an order unpredictable to the student. True pseudo-randomness is not really the issue as experimental work maximizes effect size by maximizing frequency of change within the constraint of keeping repetitions of various conditions equal for statistical purposes. These constraints remove an important aspect of randomness which is its periods of repetition. Current practice also limits study to closed skills. Markovs can provide a map-like plan and record of freer practice, extending study to open skill (described below), even mapping both participants simultaneously.

Current experimental work also seems limited to randomizing self-contained units of activity with no continuity between them. The proposed method would emphasize connectivity, actually increasing contextual interference by giving no inter-task interval at all—in fact, replacing it with a defined transition-skill which also demands attention.

Shea and Morgan's recommendation to diversify sessions is compatible with the present proposal both in selecting a level above the "atomic" for major focus, and in placing priority on shifting smoothly between several skills, and allowing quality of single skills to progress more gradually.

Expertise and Stages of Learning

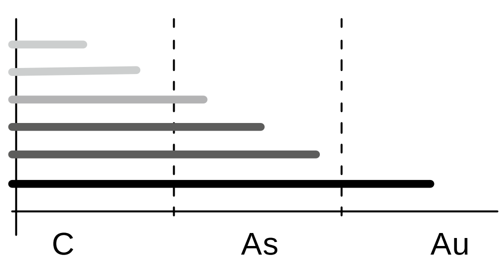
Chi, Glaser and Farr (1988) emphasized the importance of specialized structures of field knowledge, pre-considered into chunks and patterns, for easy recognition. They refer to Chase as saying that “expert drivers can generate a far greater number of secondary routes…than novice drivers” (p. xvii). This thought of having been there before goes back to de Bono’s point mentioned earlier that achieving a position of “hind-sight” improves the vision. The matrix and graph representation of a Markov process provide arithmetic and visual tools that can support the laborious, time-consuming, and difficult to communicate labyrinth of knowledge that an expert develops over as much as 10,000 to 20,000 hours of study. Chi et al. also emphasize strong self-monitoring skills, which come under meta-cognition.

Allard and Starkes (1991) extend the expertise topic to the motor realm in which CI and practice schedules are currently being most studied. In discussing motor-skill expertise, they assume any requisite physical ability and focus on cognition. Their distinction between open and closed skills is important. The skills tested in CI studies are all closed skills. These are easier to test for various reasons, and, being lower in CI, may actually be better evidence, since open skills contribute the added CI of 1) an opponent acting simultaneously, and of 2) generating an environmental effect, not just accuracy or speed of performance. They also note that open sports “seem to require cognitive skills more often associated with ‘pure’ cognitive skills such as chess or physics” (p. 137), and that in the balance between doing and knowing, closed skills such as dance or gymnastics lean toward doing and open skills lean toward knowing. This would seem to suggest that both CI testing, and also the benefits of an organizational/tracking tool like a Markov would show even greater effect in the open skill realm than in the closed skill realm. Neuroscience supports this distinction; difference in brain size and structure between prey and predator species reflects the greater cognitive demands of the hunter.

Related to the study of expertise, there are multiple models of “stages of learning” describing the changes undergone in development. All such must account in some incremental way for recognized novice-expert differences. Anderson’s (1995) 3-stage model, based upon Fitts, consists of cognitive, associative, and automatic stages. Though these stages are to be considered as general ranges in a smooth process, the involvement of difference reduction, operator sub-goaling, and production rules is quite detailed. Dreyfus and Dreyfus’s (1986) observationally-based model is 5-stage, consisting of novice, advanced beginner, competence, proficient and expert. However stages are broken down, models agree that critical differences include the reduction of effort and conscious involvement required at higher levels.

A point of speculation would be whether the acquisition phase described under practice schedules is the same as the cognitive stage, or the novice stage, or whether it extends further. It is here suspected that Markov processes have the most to contribute in these early stages of learning, when skills are being acquired, sorted out, and organized. All students must go through those early stages in which the learning curve is steepest. Some find the cost-effort greater than the reward at some point and slow or quit study. Anything which reduces the confusion and frustrations during this steep learning-curve period may be of value. The fact that every field-of-knowledge population has a pyramidal shape implies that most workers in the field are, at least part of the time, at some sub-expert level of function. This paper’s intention is to focus on the early stages, at the bottom of the pyramid, where so many people stop advancing. If beneficial there, perhaps specific or abstracted habits develop in the usage of the method will still be found of value when the individual reaches more independent stages of development.

Verbal traditions of a 4-stage model, stating that 10 repetitions bring memorization, 100 bring learning, 1000 bring understanding, and 10,00 bring mastery are echoed in formal research such as that done by Anderson (1995), or Chase and Simon (1973). This brings up two points of interest: 1) Markov-structure practice may achieve accelerated accomplishment of high numbers of repetitions, especially in link-skills, by efficient management of practice time, 2) using the 10-100-1000-10000 log-scale as a loose metric in stepping up through the levels of the method encourages the adoption of a Program Evaluation and Review Technique (PERT) diagram of stages of learning (Maier, 1999) (Figure 6). PERT became a standard tool of “fast-tracking” in architectural construction in the 1970’s, organizing the many trades involved to an optimized timeline. Elemental skills at a more advanced stage while more complex skills are being introduced is much closer to real-world progression, in which people must keep many responsibilities met simultaneously. Its cost is an increase in planning and monitoring, but if done well, its benefit is faster more integrated work. It is



Idealized PERT Chart showing simultaneous development of related skills at different levels of advancement. C = cognitive, As = Associative, Au = Automatic.

Figure 6.

visibly compatible with recommendations above to de-emphasize block practice with its short-term, small-scale focus in favor of a more integrated system approach to teaching.

Conclusion

The academic nature of Voskoglou's learning task coordinates with Battigs's language work in that both are considered predominantly intellectual and contrasts with the predominance of motor skill tasks in recent CI research. Allard and Starkes point out commonality of these two areas, at least in terms of the high cognitive demands of open sports. Neuroscience show greater brain size and complexity in predators. All this gives evidence of the broad terrain that can be influenced by work in a pivotal area like that viewed through tools like Markovs.

Markovs are useful in the investigation of primary activities, and also the ephemeral transitions which connect them. This focus on links as an important category of knowledge in itself may be the greatest single contribution of the method. The specific use of Markovs in mapping conceptual connectivity as reflected in action choices may clarify decision points more than methods like self-reporting. A major point is their potential to measure and compare actual application of knowledge and training by students under pseudo-random conditions. Markovs can serve the meta-cognitive purpose of drawing attention to possible missing or inaccurate knowledge. This may help especially in the early stages of learning. Markovs may also be able to accelerate the development of rapid pattern recognition, as well as a form of distant transfer.

This theoretical exploration gives ample reason to investigate the potentials of Markov processes in a variety of specific learning situations, and in light of several desired applications. It is hoped that this will stimulate and support use of structured pseudo-random practice and also conceptual mapping both as design elements for teaching and as meta-cognitive tools in learning and practice.

Footnotes

1The colloquial use of the tem “random” is misleading. Perfect non-correlation is just as mythical as perfect correlation. Efforts in software design to create a true random number generator have not yet succeeded. There are always a seed value, from a non-random source, and a process. Seeking randomness is thus rather like chasing the horizon or seeking a perfect vacuum. Hays is quoted by Thompson (1996): “there is surely nothing on earth that is completely independent of anything else” (p. 27). This paper uses the more accurate term “pseudo-random” except in reference to earlier works using the colloquial term.

References

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance, and other domains. In K. A. Ericsson & J. Smith (Eds.), *Toward a General Theory of Expertise: Prospects and Limits* (pp. 126-152). Cambridge: Cambridge University Press.

Anderson, J. R. (1995). *Learning and Memory: An Integrative Approach*. New York: John Wiley & Sons, Inc.

Chase, W. G., & Simon, H. A. (1973). Perception in Chess. *Cognitive Psychology, 4*, 55-81.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The Nature of Expertise*. Hillsdale, NJ: Lawrence Elrbaum Associates.

de Bono, E. (1970). *Lateral Thinking: Creativity Step by Step*. New York, NY: Harper and Row.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind Over Machine: The Power of Human intuition and Expertise in the Era of the Computer*. New York: The Free Press.

Gagne, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The Cognitive Psychology of School Learning* (2nd ed.). New York: Addison Wesley Longman, Inc.

Li, Y., & Wright, D. L. (2000). An assessment of the attention demands during random- and blocked practice schedules. *The Quarterly Journal of Experimental Psychology, 53a*(2), 591-606.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9*, 241-289.

Maier, H. N. (1999). CogPsych foundations for a wing chun curriculum [PowerPoint presentation]: unpublished.

Pfeiffer, P. E. (1978). *Concepts of Probability Theory* (2nd Revised ed.). New York, NY:

Dover.

Schmidt, R. A. (1988). *Motor Control and Learning* (2nd ed.). Champaign IL: Human Kinetics Publishers, Inc.

Shaw, R. (1984). *The Dripping Faucet as a Model Chaotic System*. Santa Cruz, CA: Aerial Press.

Shea, C. (2000). Meeting.

Shea, C. H., Kohl, R., & Indermill, C. (1990). Contextual interference contributions of practice. *Acta Psychologica, 73*, 145-157.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology*, *5*(2), 179-187.

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, *25*, 226-230.

Voskoglou, M. G. (1994). An application of Markov chain to the process of modeling. *International Journal of Mathematical Education, Science and Technology*, *25*(4), 475-480.

Voskoglou, M. G. (1995a). Measuring mathematical model building abilities. *International Journal of Mathematical Education, Science and Technology*, *26*(1), 29-35.

Voskoglou, M. G. (1995b). Use of absorbing Markov chains to describe the process of mathematical modeling: a classroom experiment. *International Journal of Mathematical Education, Science and Technology, 26*(Sept/Oct), 759-763.

Wulf, G. (1991). The effect of type of practice on motor learning in children. *Applied Cognitive Psychology, 5*, 123-134.

APPENDIX C

MEASUREMENT OF INITIATIVE IN HIGH-SPEED TACTICAL DECISION-MAKING

as presented at:
NDM6: Naturalistic Decision Making 6 Pensacola Beach, Florida, May 15-17, 2003

Measurement of Initiative in High-speed Tactical Decision Making

Herbert N. Maier, M.S.[1]
Katherine H. Taber, Ph.D.[2]
1- TactiCog[SM] and Department of Educational Psychology, Texas A&M University
2- Departments of Radiology, Psychiatry and Behavioral Sciences, and the Herbert J. Frensley Center for Imaging Research, Baylor College of Medicine

Schmitt and Klein (1996) discussed seizing the initiative, restricting the enemy's options and freedom of action, forcing situational development.

INTRODUCTION

This is a preliminary report on a method-development study of decision making in the martial art, Wing Chun. Wing Chun is shown to require macrocognitive (Klein, Klein and Klein, 2000) activity from the interaction of time-pressure and management of attention between tasks of response to the partner (uncertainty) and planning/re-planning one's own actions against a changing situation. The observed activity is cyclic. Each response action must be completed in less than 1/2 a cycle (~ 0.45 second), or fail utterly.

This urgency tends to preempt investment in building long-term advantage. Planning done as a background task during one or more full cycles pays off in pushing the urgency of judgment and response onto the partner, impacting his investment in planning. If he is close to his maximum capacity, this can push him into dysfunction, concretely demonstrating the aphorism that: ***"the best defense is a good offense".***

Martial arts are exercises in iterative high-stakes, time-constrained decision making under great uncertainty. In working at a closer range than better-known martial arts such as Karate or TaeKwonDo, practicioners of Wing Chun rely upon overtly-stated principles of physical, logistical, and cognitive organization to push the opponent to overload by imposing complexity and time-pressure. The art's close range is qualitatively unique in that two or more objectives are usually accomplished at once, versus one objective in longer-range, more "traditional" martial arts.

The activity observed here is a drill-set which has a tightly-constrained set of actions and options, occurring at a high rate of speed, limiting reflective thought, and precluding verbalization. The actions are physical, therefore directly observable.

A dimension of expertise which still appears to be mostly unmeasured is that of breadth of knowledge: what are the available options, which are feasible under different conditions, and are they actually available to the practitioner under pressure?

The use of a smaller system as a surrogate model to study processes that also occur in larger systems is well-established in biology. In observing the nervous system of the sea slug Aplysia, Kandel (2001) retained the emergent properties of neuron and synapse while viewing few enough to track their activities. Studying a small empirical decision making system about which a great deal is already known could provide similar insight into larger situations.

METHODS

Participants

12 males, age 18-24--the all-male sample is representative of the preponderance of males in the martial arts. Being undergraduate college students probably impacted their seriousness and focus. Participants had 2-6 semesters of moderately-supervised practice with the drill-set being observed, during the model's development. Their background consisted of 2-4 hours of instruction in the drill-set each semester from the designer, and irregular practice under the faculty advisor of the club, who is also a student. These far-from-ideal training conditions would be expected to reduce effect-size of training, providing a challenging test for the proposed measurements.

Procedures

Data-collection: Each participant partnered with 10 other members of the sample for 2-minute periods, video recorded for later data-extraction. Prior to data-extraction by slow-motion viewing on a personal computer, 2 pairs of participants were selected as approximating the upper and lower quartiles in experience (Dyad-1: more-experienced, Dyad-2: less-experienced). Individuals were labeled L or R based on left/right position in the camera's view, and are referred to as 1L, 1R, 2L, and 2R. In Dyad-1, 1L had 6 semesters involvement, 1R had 4. In Dyad-2, 2L had 4 semesters involvement, 2R had 2.

Table 1: An incident was tallied each time a participant was seen blocking a punch. In this illustrative example, drill 1 continues until participant 1R transitions to drill 6 in incident 8. Participant 1L follows this transition in incident 9. Drill numbers match those in Figure 1C.

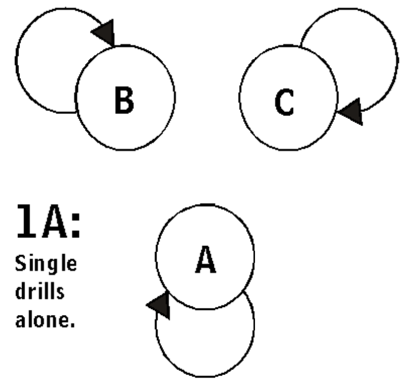
1L		1R	
Incident #	Drill	Incident #	Drill
5	1	6	1
7	1	8	6
9	6	10	6

OBSERVED ACTIVITY

The first author's decades of teaching in the field provided access to and understanding of the required knowledge-base. The activity demonstrates elusive aspects of training supported by published research in cognitive load (Sweller, van Merrienboer, & Paas, 1998), practice schedules (e.g. Brady, 1998), and knowledge organization (e.g. Schvaneveldt, 1990; Wyman & Randel, 1998). It is consistent with situational awareness literature as it intrinsically includes simultaneous foreground, background and secondary tasks.

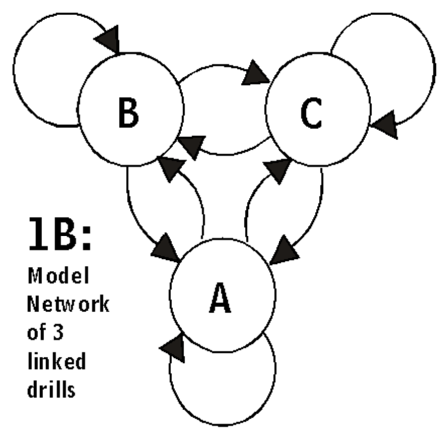
Generating Time Pressure

Three basic martial arts drills (Figure 1A), each with a block and a punch, differ in the specific block and punch given. Each complete cycle, from one partner blocking to the same partner blocking again, takes approximately 0.90 second (Table 2). Each single drill can continue indefinitely with partners alternating roles of punching and blocking.



Generating Uncertainty

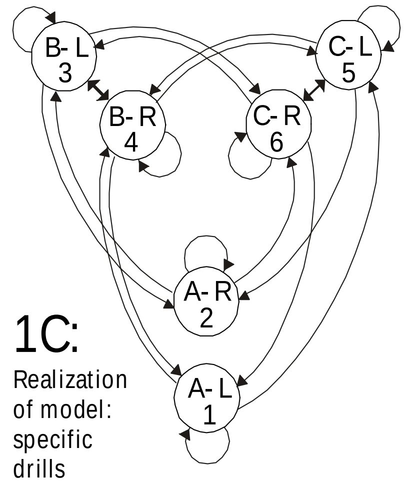
Transitions linking these drills (Figure 1B) turn the isolated drills into a network of continuous action. Each partner has three choices at each cycle: continue in the same drill, or transition to one of the other two drills. Both partners are encouraged to transition between drills as often as manageable, forcing the other partner to respond correctly to the transition while also planning his own transitions. Since neither partner knows what the other partner is going to do, the drill-network structure is an empirical example of "random practice" (Magill & Hall, 1990; Brady, 1998)



Generating Complexity

Additional complexity (Figure 1C) is introduced to the generic model by the specific drills used:
  * each drill has left- and right-hand versions, doubling the number of nodes
  * mechanical / tactical constraints prune the network, eliminating some links (Figure 1D).

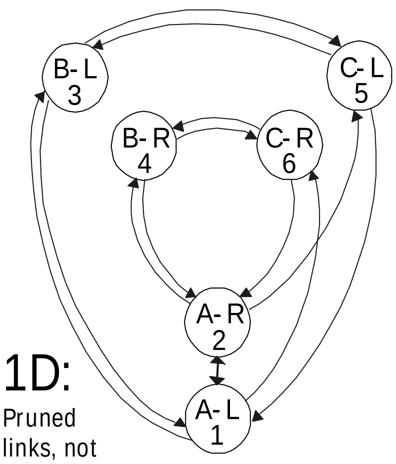
This traditional model is consistent with motor-learning literature (ibid.) concerning practice schedules, in which random practice is shown to produce greater retention and transferability than block practice.



1C:
Realization of model: specific drills

Footnote: Errors in Learning

This presentation postpones the issue of incorrect learning which occurred, generating activity in some of these pruned-out links. The participants frequently "got hit" (though the non-contact safety rules of the activity allowed them to ignore that fact). As transitions performed and therefore work accomplished, these errors were included in our analysis of quantities of activity. Their incorrect learning is credited to supervision inadequate to this particular group's needs. It did, however, generate highly informative patterns that will be dealt with in a separate article (in preparation).



1D:
Pruned links, not to be used.

RESULTS & DISCUSSION - DRILL DATA

The discovery process generated by simple observation of people in their natural activities is familiar to researchers in Naturalistic Decision Making.

Stage 1: Summary Statistics

Two dyads are sufficient to show that a range of distributions can be expected. But the very characteristics being sought preclude conventional statistics such as chi-squared due to lack of independence in the data.

Table 2: Incidents counted in each drill, in a two-minute period.

Drill ID	A		B		C			
	1	2	3	4	5	6	Total	Mean Cycle-Time
Dyad-1	71	54	49	71	7	12	264	0.93 sec
Dyad-2	41	41	76	99	13	0	270	0.88 sec

The data must be heavily dependent in order to display the emergence of an iterative system with memory from the interaction between two partners.

Stage 2: Markov Transition Matrices

Literally thousands of different samples would provide exactly the incident-counts shown in Table 2. As one of us (Maier, 2001) pointed out, Markov matrices improve on this time-blindness by viewing two-incident sequences. In focusing on transitions, this procedure was the first breakthrough in showing quantitatively what Figure 1B and 1C show in-principle.

Matrices tally changes in the state of a system by comparing consecutive pairs of states, in this case, which drill a partner-pair is doing. The matrix-row represents the drill performed first, the matrix-column, the drill performed second.

Dyad-1

		A		B		C		Drill Totals	Total Transitions
		1	2	3	4	5	6		
A	1	60	6	0	0	0	6	72	12
	2	0	48	0	0	6	0	54	6
B	3	4	0	41	3	1	0	49	8
	4	7	0	4	59	0	0	70	11
C	5	0	0	2	5	0	0	7	5
	6	0	0	2	4	0	6	12	8
Drill Totals		71	54	49	71	7	12	264	
Total Transitions		11	6	8	12	7	6		50

Dyad-2

		A		B		C		Drill Totals	Total Transitions
		1	2	3	4	5	6		
A	1	34	6	0	1	0	0	41	7
	2	5	33	2	0	1	0	41	8
B	3	1	0	67	7	1	0	76	9
	4	1	2	5	90	1	0	99	9
C	5	0	0	2	1	10	0	13	3
	6	0	0	0	0	0	0	0	0
Drill Totals		41	41	76	99	13	0	270	
Total Transitions		7	8	9	9	3	0		36

Table 3: Markov transition matrices. In conventional form, counts are translated into percentages. Incident-counts and transition-totals are more useful here.

A quick visual appraisal shows that large regions of the activity are not being practiced, and that these fall into clear patterns. Dyad-1's clusters of activity show a strong habit of circulating from A to C to B and back to A, but never C to A, or A to B. Dyad-2 shows a lack of activity between A and C, making both dead-ends, and potential traps.

Dyad-1 has shifted more activity than Dyad-2 from the diagonal outward into transitions (50 transitions vs. 36), though their experience-level would hope for more. But the proportion of the matrix explored is no greater (12 available cells vs 11).

The most informative feature is in the distributions of activity across each row, showing the branching pattern after each cycle. For example, on leaving Drill-1, Dyad-1 transitioned only to Drill-2 or 6, and on leaving Drill-2, transitioned only to Drill-5.

So, Markov Matrices produce information useful to a trainer/evaluator, on the extreme detail level of individual transitions.

Each dyad's deficiencies offer predictability and vulnerability to an opponent—a lack of preparedness to cope with the unpracticed transitions, too few options ready-to-hand. Endsley & Robertson (2000) wrote: "Less effective pilots appeared to apply the same strategies in all cases rather than matching their strategy to the situation." p.352

Stage 3: Event-Series Analysis

Voskoglou (1994, 1995a, 1995b) showed (similar to Figure 2), using Markov processes to observe problem-solving in mathematics, that transitions are still total counts, failing to differentiate large numbers of possible samples in terms of longer sequences. These longer sequences are essential in describing human thinking and behavior (i.e. composition and production sequences), indeed, any system with memory.

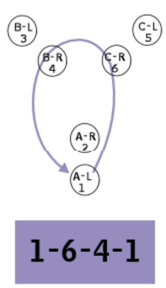
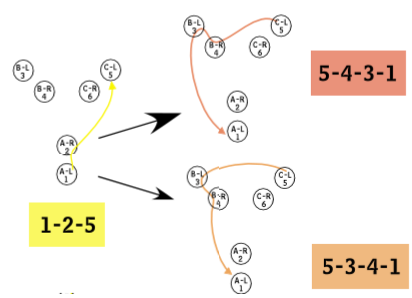


Figure 2: Our data confirms Voskoglou's point: Markov matrices are not practical in distinguishing longer sequences like these exhibited by Dyad-1. Even in such a tiny option-set as this one, such resolution is essential. As in Table 3, Dyad-1 departs Drill-1 only to 2 or 6, and departs 2 only to 5. Neither behavior nor recognition is probaility-based, but pattern-based. Analysis must respect this.



For this preliminary report, performance-time variance of individual events was filtered out, in order to focus on ordinality and sequencing. This reduced an actual time-line to an event-line.

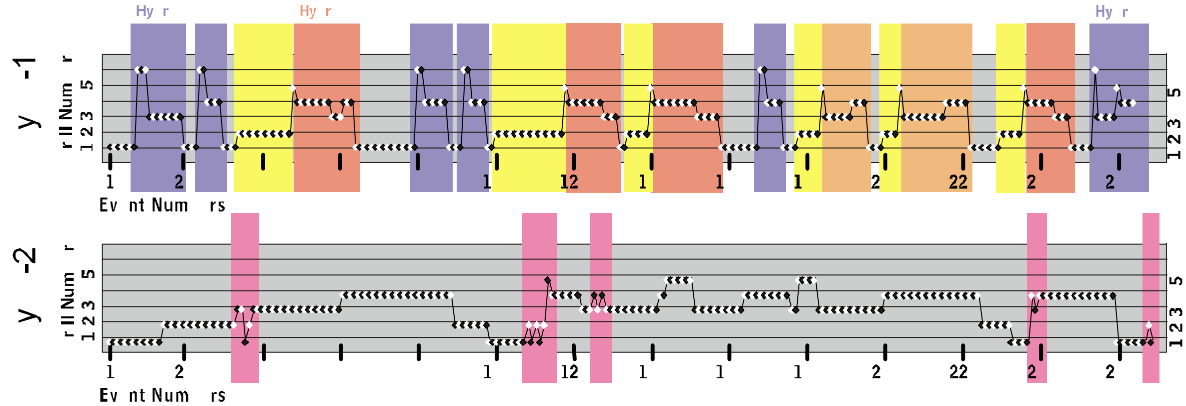


Figure 3: Event Series. Black dots indicate left partner, white dots indicate right partner.

The question of Dyad-1, Drill-6, shown but unexplained inTable 3, is clear in Figure 3, where most Drill-6 activity is in sequences of 2, because 1-R is initiating the transitions. Similarly, Dyad-1's Drill-1 to Drill-2 to Drill-5 transitions, are clearly shown as bottlenecks and harbingers of stereotypical sequences described in Figure 2.

Dyad-2's record is quite different--Table 3 shows a quantitative difference in transitions (36:50), but can not show the qualitative effect of concentrating those transitions more densely. Longer periods of block repetition are broken by intense sequences of consecutive transitions. These sequences are only possible because partners 2-L and 2-R are participating more equally than are partners 1-L and 1-R, raising the importance of recording who is initiating the transitions.

Dyad-1 and Dyad-2 thus show different aspects of the learning expected under the networked-drill practice structure. Dyad-1 shows composed sequences that run with high reliablility, and needs to train for more variety. Most of Dyad-2's performance shows a halting exploration of bounded regions of the activity; however, they intermittently sample the speed (if not the variance) of interaction which the instructional design builds up to.

The step to event-series analysis also made visible the empirically well-known importance of individual differences.

Event-series are valuable in spotlighting predictability of stereotyped actions, the significance of which would depend on the situation. Within a team, it could indicate reliability—a cohesiveness of procedure. In combat or other competition, it could be a point of high vulnerability.

Webster defines initiative as "energy or ability displayed in initiating something."

DEVELOPMENT OF CAIO SCALE

The CAIO scale was developed specifically to quantify the personal dynamic of initiative between two people as an emergent aspect of a competitive interaction.

C: (continue) expected block for same drill

A: (accomodate): expected block for drill partner had just transtioned to

I: (initate) a block which transitions to a different drill

O: (overload): failure to respond effectively

Ordinality in this scale is based upon motor-learning (practice schedule) research (e.g. Magill & Hall, 1990; Brady, 1998) which hypothesizes that block practice (continuing) is the least demanding, and that cognitive demand increases with increasing randomness (more transtions initiated), as noted earlier.

The order of A, and I assumes that causing a transition is more demanding than responding, because this task must be managed in addition to keeping up with input from the partner. O is placed at the top of the scale on the presumption that an excessive cognitive load, shown by confusion, is higher than loads which are managed successfully. O was coded based on a participant non-responding to an action (freezing) or giving one or more inappropriate responses (flailing).

RESULTS & DISCUSSION - CAIO SCALE

Summary Statistics

Dyad-1 initiated more (51 times versus 33 times) and overloaded less (1 versus 9) than Dyad-2 (Table 4). Thus the CAIO scale captures some aspects of the differences in expertise between these pairs.

It is seen here that Dyad-2 (initiate ~3:2) is better-matched than Dyad-1 (initiate ~4:1). The consistency of the more experienced dyad and the more experienced partner within each dyad initiating more and overloading less is suggestive.

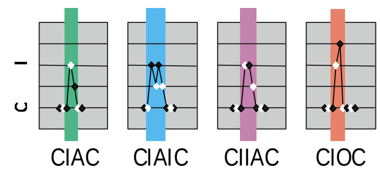
Pair	Person	C	A	I	O
1	L	89	31	11	1
	R	81	11	40	0
2	L	105	11	13	6
	R	105	7	20	3

Table 4: Incidents interpreted through CAIO.

Dyad-1 has grasped only 36% of the initiative opportunities available in their 2 minutes, and Dyad-2 only 23%, based upon the potential for every event to be an initiation.

CAIO Series Analysis

As with drill data, the clear contrasts in CAIO summary statistics are enriched by viewing the order and clustering in a series. The difference in frequency and pattern of initiations and overloads becomes much easier to appreciate.



Four basic sequences are seen.

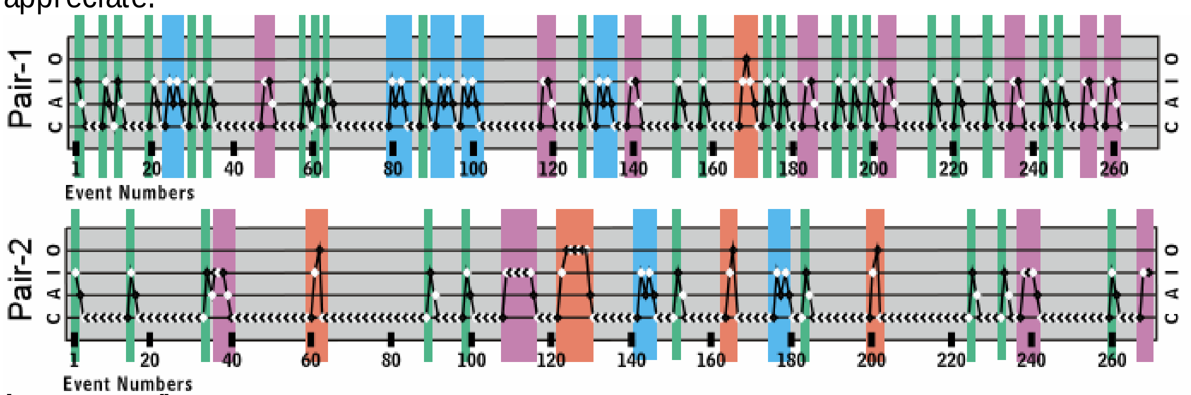


Figure 4: CAIO Series for Dyads 1 & 2: As with event series shown in Figure 3, longer sequences are highly informative. Black dots indicate left partner, white dots indicate right partner.

Sequences seen in the CAIO series of each pair reveal considerable differences in the dynamic emerging in each pair. In Dyad-1, 1-R's (white dot) dominance in initiating is seen in both CIAC and CIAIC sequences. It is, however, noteworthy that 1-R does not extend the CIIAC sequences that occur, even though all 8 incidences involve the same drill-sequence.

1-L initiates only in response to 1-R. This may be effective counter-fighting, but is not what was instructed. 1-L is coping with greater incoming activity than either member of Pair-2, but is doing so by significantly reducing his cognitive load through not initiating. 1-L exhibits the only overload CIOC in Dyad-1's series.

In contrast, Dyad-2 shows lower levels of initiation overall and their I-incidents are more clustered, for example at incident-numbers 238-240 (3 consecutive I's) and incident number 109-114 (7 consecutive I's). A third such episode at incident-number 123 caused a discrepancy between Dyad-2's total initiations between Table 3 (36) and Table 4 (33) by generating a cascade of overloads beginning at incident number 124, seen in Figure 4. This moment of total disarray in both partners includes very game but ineffective attempts to recover through initiating a different drill.

CONCLUSIONS

It has been shown that:

1) breadth of domain-knowledge accessible under pressure

2) level of initiative taken

are measureable in ways that provide detailed, concrete feedback for further learning and practice. Acceptance of overload incidents in practice is important both to stretch capacity, as well as to acclimate to the experience and learn recovery.

The CAIO scale **shows promise for revealing dimensions of macrocognition:**

1) management of attentional resources

2) response to overload.

It provides a way to analyze an emergent property of each individual within the interpersonal dynamic of the dyad. The method's adaptability is indicated by different possible interpretations in different situations (competition vs. team).

The networked-drill structure **increases pressure by**

1) concentrating approximately 60 "decision incidents" into each minute

2) maintaining a constant flow of relevant time-critical demands.

The discontiuity emerging from these factors builds the fundamental habit of shifting urgency onto the opponent. It also describes an expert model in high detail.

Endsley & Robertson (2000) described 1)teams who prepared only one plan as being suseptible to Level-2 SA errors, and 2)the importance of pilots training in a broad range of sets of situational cues. They mention the need for preflight preparation, and shifting from passive to comprehending to proactive in thinking.

Explicit learning of transitions **redefines the "learning unit". It:**

1) builds a smoothly integrated knowledge-set from the first lesson

2) provides a kernel for integrating future lessons.

Link knowledge is needed for quick flexible response and judgment/decision-making: defining options at each decision-point, including degrees-of-freedom, parameters, and critical values (crucial to recognition in vague circumstances of incomplete, contradictory information).

FUTURE DIRECTIONS

1) Networked-drills have generations of real-world application in certain martial arts. This structure should be tested in both training (skill-development) and diagnostic/qualification (skill evaluation) roles, in various activities.

2) The CAIO scale and its relationship to the traditional arousal curve.

3) The complexity and dynamic stability of an unfolding scenario makes sequence a vital aspect of any study. Time-series analysis could reveal the evolution of situation awareness, communication, and error recovery during an activity. It has been used successfully in social sciences (e.g. Gottman, 1981).

4) Though only hinted at here, the cohesive and explicit conceptual structure of Wing Chun underlies this study.

a) personal communications with US Army instructors strongly support elements of the art as models of fundamental infantry tactic. This should be explored.

b) any model of expertise introspective enough to verbalize the application and limitations of its own theoretical structure should be examined, recorded and compared to others.

The above aspects should be tested under good conditions to validate the method, and poor conditions to test its resilience and introduce additional questions that will certainly appear under the varying conditions of actual use.

This approach could offer a comprehensive view of the macrocognition of initiative in a framework amenable to research (Woods, 2002).

REFERENCES

Brady, F. (1998). A theoretical and empirical review of the contextual interference effect and the learning of motor skills. Quest, 50, 266-293.

Endsley, M. R., & Robertson, M. M. (2000). Training for situation awareness in individuals and teams. In M. R. Endsley & D. J. Garland (Eds.), Situation Awareness Analysis and Measurement (pp. 349-365). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gottman, J. M. (1981). Time-series analysis: a comprehensive introduction for social scientists. New York: Cambridge University Press.

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. Science, 294, 1030-1038.

Klein, D. E., Klein, H. A., & Klein, G. (2000). Macrocognition: linking cognitive psychology and cognitive ergonomics. Paper presented at the 5th International Conference on Human Interactions with Complex Systems, Urbana-Champaign.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. Human Movement Science, 9, 241-289.

Maier, H. N. (2001, March 21). Structuring random practice through markov processes. Paper presented at the Educational Research Exchange, Texas A&M University.

Schmitt, J. F. M., USMCR, & Klein, G. A. (1996). Fighting in the fog: dealing with battlefield uncertainty. Marine Corps Gazette, 62-69.

Schvaneveldt, R. W. (Ed.). (1990). Pathfinder Associative Networks: Studies in Knowledge Organization. Norwood, NJ: Ablex.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. Educational Psychology review, 10(3), 251-296.

Voskoglou, M. G. (1994). An application of Markov chain to the process of modeling. International Journal of Mathematical Education, Science and Technology, 25(4), 475-480.

Voskoglou, M. G. (1995a). Measuring mathematical model building abilities. International Journal of Mathematical Education, Science and Technology, 26(1), 29-35.

Voskoglou, M. G. (1995b). Use of absorbing Markov chains to describe the process of mathematical modeling: a classroom experiment. International Journal of Mathematical Education, Science and Technology, 26(Sept/Oct), 759-763.

Woods, D. O. (2002). GUTs or no GUTs (Grand Unified Theories): does/can/should cognitive engineering have G.U.T.s? Paper presented at the 46th Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD.

Wyman, B. G., & Randel, J. M. (1998). The relation of knowledge organization to performance of a complex cognitive task. Applied Cognitive Psychology, 12, 251-264.

VITA

Herbert N. Maier
19130 Oak Canyon Dr.
Conroe, TX 77385
hmaier@tacticog.com
sifu@yellowrose-winghcun.org

Education

M.S.	Educational Psychology, Texas A&M University, College Station, TX	2000
B.S.	Architecture, University of Houston, Houston, TX	1978

Teaching Experience (Major)

Trainer, Science Fair Judging, Sci://Tech	2002-present
Martial Arts Coalition, Mesa State College, Grand Junction, CO	1999-present
Director of Training, Yellow Rose Wing Chun Association	1986-present
Maier Academy of Martial Arts, Spring TX	1984-88
Instructor, various martial arts programs	1978-84

Conference Presentations

Maier, H. N., & Taber, K. H. (2003). Measurement of initiative in high-speed tactical decision making. Naturalistic Decision Making 6, Pensacola Beach, FL.

Maier, H. N., & Stricker, A. (2001). Software-driven sociogram generation for small-group analysis. Human.Systems 2001: Exploring the Human Frontier, Clear Lake City, TX.

Maier, H. N. (2001). Structuring random practice through Markov processes. Educational Research Exchange, Texas A&M University.

Maier, H. N. (1998). Working the centerline: defining the problem with hands-on-thinking. American Creativity Association, Houston, TX. April 22-25, 1998.

Maier, H. N. (1998). Attack the centerline: keep your balance under pressure. American Creativity Association, Houston, TX. April 22-25, 1998.

Accepted

Maier, H. N. (2004). Thinking Outside the Box(er) ©, Hands-on Problem-Solving. American Creativity Association, Houston TX.. April 1-3, 2004.

Honors and Awards

Phi Kappa Phi, National Honor Fraternity
Kappa Delta Pi, Educational Honor Fraternity