Human–Technology Interaction and Music Perception and Performance: Toward the Robust Design of Sociotechnical Systems

ALEX KIRLIK AND SHIN MARUYAMA

Invited Paper

A common framework for studying perception and performance in both human–technology interaction and music is presented. The framework represents the cognitive challenges faced by both musicians and human operators in technological systems. In the perceptual realm, both must infer distal or covert states from proximally available information (e.g., inferring the emotional meaning of a musical composition; inferring the functional meaning of interface displays). In the realm of action, both must select proximally available actions, or means, to achieve distal ends or goals (e.g., a conductor using hand movements to direct an orchestra, an operator using interface controls to tune an industrial process or respond to a fault). The framework represents these proximal-distal relations and enables quantitative measurement of the degree to which performers adapt to them. The framework is illustrated with a review of music and human–technology interaction research and our own study of the coordination between a professional conductor's hand movements and a concertmaster's bowing actions in the opening of Beethoven's Fifth Symphony. In providing a common theoretical framework for both music and engineering, we hope to enhance prospects for research on group musical performance to inspire novel, robust design models for modern sociotechnical systems.

Keywords*—Conductor–orchestra interaction, human–computer interaction, human–machine systems, human performance modeling, music, musical perception and performance, robust control, robust design.*

I. INTRODUCTION

As demonstrated by the first article in this special issue [1], interesting analogies can be drawn between the performance of musicians, guided by a written score and performing under the direction of a conductor, and human–technology interaction in complex sociotechnical systems. The question remains, however, of whether a suitable conceptual framework can be developed to describe and analyze phenomena from these disparate domains, using a common language and a shared set of theoretical constructs and modeling techniques. Such a framework is required to facilitate cross-disciplinary research and transfer of findings across domains, and to help the music and engineering communities' ability to talk with one another in a mutually intelligible fashion about phenomena of shared interest.

The purpose of this paper is to propose one candidate framework for describing and investigating issues relating to perception, performance, control, and coordination in both engineering and music. While this framework has its origins in early psychological research on adaptive behavior [2], [3], we believe it has a number of properties that should make it attractive to the engineering community. First, the approach is based on a functional perspective on behavior. That is, the focus is on what ends or purposes behavior is achieving, or attempting to achieve, while leaving aside the vast array of subtle and difficult questions pertaining to the inner psychological or neural processes that accomplish these functions. As such, not only is the approach compatible with the functionally oriented mode of reasoning which is one hallmark of engineering, but it is grounded in the analysis of the relationships between observable variables, rather than in theory of unobservable constructs or processes (e.g., attention, representation, memory, decision making, problem solving).

Second, the framework is based on very weak (theoretically speaking) assumptions about these unobservable entities and activities, with two resulting benefits. First, the analysis and modeling approach, being neutral with respect to process, is amenable to describing the behavior of both

Manuscript received February 18, 2003; revised November 10, 2003. The work of A. Kirlik was supported by NASA Ames Research Center under Grant NAG 2-1609. The work of S. Maruyama was supported by a Grant-in-Aid for Scientific Research from the Japan Society for Promotion of Science under Grant 09620.

A. Kirlik is with the University of Illinois at Urbana–Champaign, Urbana, IL 61801 USA (e-mail: kirlik@uiuc.edu).

S. Maruyama is with Indiana University, Bloomington, IL 47408 USA, and also with the University of Tokyo, Tokyo 113-0033, Japan.

Digital Object Identifier 10.1109/JPROC.2004.825881

humans and machines, such as automation and information technology, and especially human–technology interaction. This feature makes the approach particularly attractive from the perspective of analyzing and designing sociotechnical systems, which are typically composed of teams of humans interacting with a wide array of technologies (e.g., a NASA mission control center, a power plant control room, a stock exchange or trading floor).

Additionally, and as will be shown below, once the theoretically relevant variables are identified and measured, there is no practical limit to the range of mathematical or computational techniques that could potentially be brought to bear on the analysis of behavior (e.g., perception, action, coordination, control) or on any limitations on these activities. For example, the framework has already been applied to examine human–technology interaction using techniques as varied as linear regression [4], [5], entropy-based, multidimensional information theory [6], and genetic algorithms [7].

The remainder of the paper is organized as follows. In Section II, we discuss the engineering-oriented motivation for achieving a common theoretical framework for investigating perception and performance in both music and human–technology interaction. This discussion, focusing on the need to develop more flexible and robust sociotechnical systems, motivates Section III, in which the theoretical framework is presented. Section III presents the approach in its historical light, describing and drawing upon the progression of theory in both music and psychology, hopefully demonstrating that the common framework presented in this paper is not merely fortuitous, but is actually grounded in a convergence of thought in both music theory and psychological theory. Section IV then discusses a body of literature illustrating how the theoretical framework has already been applied and extended to investigate perception and performance in both music and engineering, although we know of no preexisting literature that makes explicit connections across these lines. In order to demonstrate how the study of group musical performance might be conducted in a manner that could shed light on engineering problems, Section V presents our own study of the coordination between a conductor's hand movements and a concertmaster's bowing actions in the opening of Beethoven's Fifth Symphony. Section VI concludes the paper, evaluating the prospects for additional studies of group musical performance to yield new insights into designing more effective and robust sociotechnical systems, composed of both humans and sophisticated technology.

II. TOWARD FLEXIBLE AND ROBUST HUMAN–TECHNOLOGY INTERACTION

Inagaki and Stahre [1] have presented the outlines of the "supervisory control" paradigm originating from the work of Sheridan [8], and have noted both similarities and dissimilarities between this paradigm and group musical performance. In keeping with both accepted human factors principles and the practical realities of the design and operation of complex sociotechnical systems (e.g., unexpected faults, environmental disturbances), this paradigm envisions the human operator (or team) as having ultimate decision and control authority over automation. Thus, in the human factors and cognitive engineering communities, the term "supervisory control" is used to describe the situation of a human (or team) monitoring automated control systems (e.g., in manufacturing, process control), managing semiautonomous automation (e.g., a pilot programming a cockpit flight management system), and intervening when necessary to handle unpredictable or rare situations that automation designers could not anticipate [9], [10].

It should be pointed out, however, that the term "supervisory control" has come into prominent use in a second research community as well. In the systems engineering and control theory community, supervisory control theory (SCT) [11], and its applications to computer supervision of complex discrete-event systems have received much attention in the last decade. In contrast to the human factors and cognitive engineering concept of supervision, this community views computers and automation as playing the role of supervisor and seeks, for example, formal methods for synthesizing algorithms for computer supervisory controllers based on detailed specifications of the industrial plants or processes to be controlled.

In this latter community, then, the term "supervisor" is reserved for computer-based automation and the associated algorithms for system control. This approach is increasingly finding application in domains such as chemical process plants [12], factory automation [7], [13], and designing communication networks and railway automation [14]. While SCT has placed emphasis on the analysis of the controlled system, including considerations of observability and controllability, it has failed to address issues that arise from human involvement, which, in most cases, are central to the successful operation of industrial systems in actual practice.

Despite these advances, engineering techniques ignoring human involvement in system control can lead to a wide range of problems in achieving a productive coupling between humans and automation. The desire of many of those coming from an exclusively engineering-oriented mindset to minimize human involvement in system control and, where it is absolutely necessary, to proceduralize worker behavior to the greatest extent possible has, of course, a long legacy going back at least to the ideas of F. Taylor and "scientific management" (see [15] for a historical perspective). In the move toward increased levels of automation and proceduralization in even the most modern work environments, it is often forgotten that humans often bring a capability for creativity and improvisation that, in many cases, represents the only resource for achieving stability in the face of unpredictable events. An example from the activities at the site of the former World Trade Center in New York after the tragic events of September 11, 2001, illustrates this point well:

> Though the firemen who rioted on November 2 did not believe it, when Giuliani gave "safety" as the reason for reducing their presence on the pile, he was completely sincere. This was somewhat counterintuitive,

> since the safety record so far had been extraordinarily good: despite the fires, the instability of the ruins, and the crushing weight of the equipment and debris, not a single recovery worker had been killed, and only a few had been seriously injured. Indeed, discounting possible long-term respiratory problems, the injury rate was about one half that of the construction-industry average. Some people claimed this as a sign of God's favor, but a more mundane explanation was *that the inapplicability of ordinary rules and procedures to such a chaotic environment required workers there to think for themselves*, which they proved very capable of doing [16, p. 103] [emphasis added].

From an engineering perspective, students are taught that process variability is an enemy, a factor to be reduced, if not eliminated, to gain high reliability. But behavioral variability is an essential aspect of human involvement in any system. One key to aiding the design of future sociotechnical systems will be to better understand the tradeoffs associated with tolerating, and perhaps even embracing, some degree of human-induced variability in system performance, in the name of increased robustness [17].

The importance of achieving an adaptive, flexible coupling of humans and technology has been especially stressed by cognitive systems engineers working within the "joint cognitive systems" tradition [18]–[21]. This tradition highlights the importance of considering integrated, human–machine functioning, rather than either the individual machine or human, as the primitive unit of analysis and design. The modeling and analysis approach presented in the following, due to its systems-oriented, functional nature, is fully consistent with this perspective.

When information technology and automation are not explicitly designed to embrace and support human performance, a wide range of problems arise (see [22] for an excellent overview of this issue). However, human factors and cognitive engineering research has not yet been very successful in creating formal methods or models for aiding the design of more effective, and particularly more robust, human–technology interaction (but see [9] and [23]–[26] for a few notable exceptions).

A. *Integration: The Need for a Systems Perspective*

Thus, a look across both the engineering and human factors/cognitive engineering literature suggests that researchers from the two perspectives frame human–technology interaction problems in strikingly different ways, with comparatively little research being conducted at their intersection. The most promising solutions to these problems are unlikely to result from research focusing on either the role of the human or the role of technology, at the expense of the other. Instead, the most promising solutions are likely to arise out of research that focuses on the analysis and design of joint, dynamically interactive, human–technology systems. Performing such research requires techniques for describing human and technological functions in common, or at least mutually compatible, terms.

To take just one illustration, consider the many successes of manual control research (e.g., [27]), in enhancing human–machine interaction in vehicular control (e.g., the driven car, the flown plane). To us, at least, a key contributor to the success of manual control research was the availability of common construct languages, such as dynamics, and estimation and control theories, for describing both the human and machine components of these interactive systems. Likewise, the framework presented below has a distinctive systems orientation, as it is also based a common set of constructs for describing both human and environmental aspects of interaction in the realms of both music and engineering.

III. PERCEPTION AND PERFORMANCE IN MUSIC AND ENGINEERING

L. Meyer, in his classic book *Emotion and Meaning in Music* [28], sketched a theory of music grounded in the gestalt theory of perception. Meyer argued that much of what makes music meaningful and emotive arises out of its capacity to prompt, suspend, and sometimes even violate listeners' expectations (what he called "the inhibition of a tendency"). Meyer concluded his analysis with the following observation:

> Finally, it is important to emphasize that a theory of music does not exist in a kind of splendid, irrelevant isolation. If it is to be fruitful, music theory must not only be internally consistent but it must also be consistent with and relevant to concepts and theories in other realms of thought. Thus it is significant that many of the concepts presented in this book have clear counterparts in the theory of games and in information theory. To cite only one instance of this: it seems quite possible to equate the inhibition of a tendency, which of necessity gives rise to uncertainty and awareness of alternative consequents, with the concept of entropy in information theory [28, p. 255].

Meyer's interest in the cybernetic concepts of information and entropy, which allow uncertainty and order to be treated mathematically, arises from the role that listeners' expectations, as well as suspensions and violations of these expectations, play in his theory of music perception. In particular, Meyer conceives of these expectations, or tendencies, as adaptations (learned rather than innate) to the statistical structure of music. In his thinking on these issues, then, the concept of probability plays an important role:

> A sound or group of sounds (whether simultaneous, successive, or both) that indicate, imply, or lead the listener to expect a more or less probable consequent event are a musical gesture or "sound term" within a particular style system.... Ambiguity is important because it gives rise to particularly strong tensions and powerful expectations.... There would seem to be various degrees of ambiguity. A sound stimulus becomes a sound term by entering into probability relationships

> with other sound terms within the style. These probability relationships are of different degrees [28, pp. 45–52].

Given this view of music perception, the problem for Meyer then became to explain why music perception appears to be so regular and highly structured in the face of the richly textured, probabilistic tapestry that is the musical "sound stimulus." And not surprisingly, Meyer turned to a prevailing perceptual theory of his day, gestalt psychology, for the answer. The three central chapters in his 1956 book each describe "Principles of Pattern Perception" in terms of gestalt "laws" such as good continuation, completion, closure, etc. In this view, the listener resolves the uncertainties inherent in the probabilistic stimulation and formulates expectations by structuring the stimulus information in terms of these gestalt principles of "good form."

A. Ecological Perception

While gestalt theories of perception have lost some of their luster since 1956, Meyer came tantalizingly close to another approach to the problem of music perception in his book, in particular, in his discussion of where meaning is to be found in music:

> Meaning, then, is not in either the stimulus, or what it points to, or in the observer. Rather it arises out of what [George] Mead and [Morris] Cohen have called the "triadic" relationship between (1) an object or stimulus; (2) that to which the stimulus points; and (3) the conscious observer [28, p. 34]

While Meyer was intrigued by this relational perspective on the source of musical meaning, he did not further pursue, or try to formalize, this "triadic" view as the basis for his ultimate theory of meaning and emotion in music. However, it is interesting that the same year that saw publication of Meyer's book also saw this "triadic" model of the encounter between a person and the environment put forward as an alternative to gestalt theory, in the psychologist E. Brunswik's book *Perception and the Representative Design of Psychological Experiments* [2]. Brunswik's ecological[1] theory of perception and judgment, which offered a novel solution to perception in a probabilistically textured environment, and one that does not appeal to gestalt principles, is best understood in terms of his "lens model," as depicted in Fig. 1.

Brunswik's lens model makes explicit the three entities that comprise Meyer's "triadic" notion of musical meaning. Music perception or judgment appears on the left side of the diagram, corresponding to Meyer's "observer." Perception or judgment is of what Brunswik called a distal state or event (for Meyer, "that to which a stimulus points"). Finally, in uncertain situations, perception or judgment of the distal state or event is indirect; it is mediated through a reliance on proximal cues (for Brunswik) or the stimulus (for Meyer). The lens model captures all three of Meyer's "triadic" relationships.

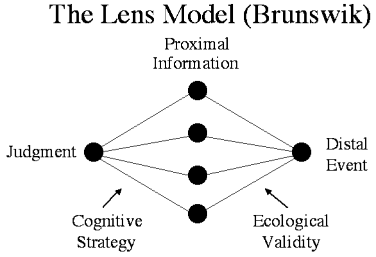


Fig. 1. The lens model represents cognitive adaptation to environmental structure.

Brunswik advanced his lens model of perception as an alternative to gestalt theory, as it solves the problem of perceptual attainment quite differently than in terms of cognitive principles of good form. Instead, for Brunswik, attainment is to be explained by a study of the probabilistic relations that mediate between the observer and the proximal cues or stimulus ("cognitive strategy" in Fig. 1), and also by a study of the probabilistic relations that mediate between the stimulus and the distal object or event ("ecological validity" in Fig. 1). In modern psychological research using the lens model, these relationships are typically described and measured using correlational statistics.

Additionally, linear regression modeling is typically used as a technique for measuring the degree to which the observer's cognitive (or perceptual) strategy is in tune with ecological, cue-criterion relations (i.e., how well the human is adapted to the structure of the environment). In this approach, one regresses human judgments (or perceptions) on the cue set to obtain a statistical model of how the observer weights and combines the cues. One also regresses the distal event (the environmental criterion) on the same cue set to obtain a statistical model of the actual cue-criterion structure that best describes the environment. One then examines the two resulting models to determine how well the observer (in the case of music, a listener) is adapted to the statistical structure of the (musical) environment.

Importantly, by using a mathematical relationship that emerges from the creation of these two regression models, known as the lens model equation, one can decompose human performance into a number of separable components (see [31] for a primer on the methodology of lens model analysis). This equation, for example, allows one to diagnose whether less than perfect human performance is due to imperfect knowledge (i.e., the cognitive strategy for cue selection and weighting), an inability to consistently execute the strategy, inherent environmental uncertainty, and other factors. For a comprehensive description of research

---

[1] The "ecological" approach to psychology, focused on animal–environment relations, is typically taken to be copioneered by both Brunswik and J. J. Gibson, although some of the ideas can be traced back to James' functionalism and Dewey's pragmatism. The central difference is that while Brunswik viewed the environment in largely probabilistic terms, Gibson argued that in natural conditions, information is available to uniquely specify the animal–environment relationship. As such, Brunswikian theory has had its greatest impact on the study of judgment and decision making under uncertainty, while Gibsonian theory has had its greatest impact on the study of perception and action. For more information on both the similarities and differences between the ideas of these two ecological theorists, see [29], [30], and [56].

within this theoretical framework since the publication of Brunswik's classic 1956 book, see [32].

Like Meyer, Brunswik also saw a relationship between his model of adaptation to a probabilistic environment and emerging ideas in cybernetics and the theory of information:

> The reader will recognize that the vicariousness of psychological cues and means which we have come to acknowledge as the backbone of stabilized achievement may be viewed as a special case of receiving or sending messages through redundant, even though not literally repetitive channels.... Even so, there is a long way to go from the rather rudimentary emergence of the concept of cue in cybernetics and in the theory of communication to the more varied and somewhat metaphorical applications that would have to be made to render these considerations really fruitful in psychology. Hitherto most of the efforts to apply these disciplines to psychological problems *have been rather literal minded and have considered the organism rather than the ecology as the prime source of noise and uncertainty* [2, pp. 142–143] [emphasis added].

As reflected in the final sentence above, initial applications of information theory in psychology focused largely on viewing internal perception, cognition, and motor processing in terms of the transmission of information along a noisy communication channel. These applications gave rise to various mathematical human performance laws for relatively simple reaction time and discrete movement behaviors [33]. On the other hand, information theory has, to this point, been of limited value in understanding more sophisticated and flexible cognitive behaviors, such as decision making and problem solving. Many if not most psychological scientists view the heyday of information theory in psychology as already in the past.

Those more interested in human interaction with the external world, with machines, and with other humans may still find uses for information theory for describing *environmental* structure or predictability, especially as it relates to the ease of interactive behavior (see Section IV-C). Probabilistic functionalism is neutral with respect to how data is analyzed after the relevant proximal and distal variables are identified and measured. As shown in the following section, the lens model is a useful framework for modeling perception and performance to support analysis, not only with correlational statistics and linear regression, but also with techniques ranging from information theory to genetic algorithms.

B. Vicarious Functioning and the Role of Action

Before we review applications of probabilistic functionalism in both engineering and music, two aspects of this theory require elaboration.

First, in understanding the process of adaptation within the framework of probabilistic functionalism, the concept of "vicarious functioning" plays a central role. It is a particularly important concept for potentially shedding light on issues such as adaptation, coordination, and even improvisation in the realms of engineering and music. As indicated in the preceding quotation, Brunswik took the "backbone of stabilized achievement" as owing to "the vicariousness of psychological cues and means" [2, p. 142]. In simpler terms, Brunwswik observed that environments typically provide multiple pathways between cues and judgments and between means and ends. Because of these (at least partial) redundancies, there may exist a variety of functionally effective strategies for combining or sequencing these cues and means to yield stabilized adaptation. Thus, Brunswik conceived of adaptation as an ongoing, opportunistic process of "coming to terms" with the cue-criterion and means–ends relations present within a changing environment. As such, he embraced, rather than disavowed, variability as a fact of adaptive behavior.

Vicarious functioning, thus, carries with it a notion of robustness, viewing, as it does, adaptation as exploiting both variety and change in both the information (cues) and actions (means) available, in opportunistic and often variable ways, to nevertheless achieve the stable attainment of goals. When an information source or action becomes unavailable, for example, probabilistic functionalism views the human, when operating in a supportively designed environment, as adapting in an online fashion (improvising) novel solutions to problems. This is achieved by exploiting redundancies (Brunswik used the term "causal texture") in proximal–distal relations. E-mail down? Pick up the phone. Cannot find an article you want to cite? Look for another article that probably cites it. The challenge, as well as the opportunity, for engineering design is to use these insights in order to understand and support robust human perception and performance in the design of sociotechnical systems.

Second, it is important to keep in mind that Brunswik intended his lens model to represent performance, or action, in addition to perception (unlike the modern literature in this area, which focuses almost exclusively on perception or judgment). For action, the structure of the lens model still applies, but instead of "cues" one substitutes "means" (or proximal actions), and instead of "distal event" one substitutes "ends" or goals. Note the symmetry between the analysis of both perception and performance in this conceptual framework. Both are understood to be associated with bridging proximal–distal gulfs or relations.

In this scheme, a well-adapted agent is one whose behavior is attuned to *both* the probabilistic relations that mediate between informative cues and perception or judgment *and* the probabilistic relations that mediate between proximal means or actions and the agent's ultimate ends or goals. Studies in this theoretical tradition, however, have typically focused on either the perception or performance problems independently of one another. However, the study of group coordination and improvisation is likely to require the simultaneous study of both perceptual and performance aspects of behavior and, in particular, how the taking of actions to reach goals generates new information sources or cues. Studies in this spirit would then provide resources for closed-loop, dynamic modeling of human–technology interaction (e.g., see [6], [34]–[36].

In the following section, we provide a brief overview of applications of probabilistic functionalism to the study of both perception and performance in both music and human–technology interaction.

IV. APPLICATIONS OF PROBABILISTIC FUNCTIONALISM IN MUSIC AND ENGINEERING

Understanding the functioning of an interactive, and potentially distributed, probabilistic system, whether it be an individual listening to music, the performance of a symphony orchestra, a pilot interacting with cockpit automation, or a NASA mission control team, requires techniques for describing the causal dependencies between the entities or agents within the system that result in (hopefully) adaptive perception and performance. Each agent in such a system, human or machine, might be assumed to have immediately available a set of proximal information sources or cues, as represented in the lens model. From these cues, each agent attempts to either perceive or infer some distal or otherwise directly unobservable, covert state of the environment and attempts to adapt its own behavior in light of the ecological structure relating these cues to that distal state or task criterion.

A. Applications in Music

Perhaps the most broad and sustained program of research investigating music within the framework of probabilistic functionalism and the lens model has been conducted by P. N. Juslin and his collaborators at Uppsala University, Uppsala, Sweden [37]. To illustrate the range of issues addressed by this research group, their recent article titles have included "Cue utilization in communication of emotion in music performance: Relating perception to performance" [38], "Communicating emotion in music performance: A review and theoretical framework" [39], and "Teaching the thing that cannot be taught: feedback-learning of musical expressivity" [40].

Juslin [38] provides a useful, representative example of research within this tradition. This experiment involved asking three professional guitarists to perform short melodies with the goal of communicating emotions of anger, sadness, happiness, and fear to listeners. The goal of this research was to understand both how performers create and listeners perceive various acoustic cues that mediate musician–listener interaction in emotional communication. Juslin identified and measured five such cues and formulated a lens model in terms of Brunswik's theoretical framework. Using multiple regression techniques and the lens model equation, Juslin discovered an interesting array of findings.

Overall, it was found through correlational analysis that performers were generally successful in communicating intended emotions to listeners. Second, regression modeling of the relations between both the performers' intentions and the acoustic cues generated and of the relations between those cues and listener perceptions or judgments revealed that the best fitting cue-weighting patterns were very similar across the two models, indicating a high degree of adaptive "fit" within this musician–listener system.

In addition, through lens model analysis, Juslin found that the cue weighting patterns exhibited by listeners (i.e., "cognitive strategy" in Fig. 1) were more consistent across different melodies than across different performers. This result was due to differences in the way in which the various guitarists used the multiple means available (the generation of acoustic cues) to communicate emotions. Since these cues were partially redundant, different performers could communicate emotions to listeners with equal effectiveness, despite differences in the manner in which they generated cues to communicate these emotions. So, even though listeners' cue weighting strategies differed across guitarists, these listeners were apparently able to pick up on how different guitarists each systematically produced different acoustic cues to communicate the same emotion. This ability resulted in a stable and robust level of adaptation in this study, amid differences in perceptual stimulation. This is a classic illustration of vicarious functioning at work, in that the "causal texture" of the environment could be, and was, exploited in various ways (e.g., different musicians or composers may have different "signature" styles, yet many are similarly effective in communicating sadness or joy).

B. Applications in Engineering

Kirlik provides a comprehensive overview of human–technology interaction research performed within the framework of probabilistic functionalism, including many modern extensions to Brunswik's original theory and modeling [41]. For example, Bisantz *et al.* [4] presented a study using regression-based lens modeling to infer the judgment strategies used by experimental participants in a laboratory simulation of a combat information center (CIC) aboard a U.S. Navy cruiser. Their findings indicated, for example, that suboptimal judgment performance in that context had much less to do with each performer's knowledge about how to weight and combine judgment cues than it did with the abilities of the performers to rapidly and consistently execute these strategies under conditions of time stress and high information load.

Bisantz and Pritchett [5] similarly used lens modeling techniques to examine pilot interaction with cockpit alerting automation in commercial aviation, where automation can be looked at as simply another judgment agent, and can therefore be described in the same *functional* terms as the human pilot. In their study, the alerting automation sounded an alarm to pilots when it predicted a forthcoming conflict or potential collision with another aircraft. The question posed by these researchers was to understand the degree to which unaided, human pilot judgment strategies were either congruent or incongruent with the strategies embodied in the automated alerting system. This study was motivated by practical engineering questions concerning how best to couple humans and alerting automation in aviation contexts. Bisantz and Pritchett's experiment and lens model analysis helped them diagnose the primary source of dissonance

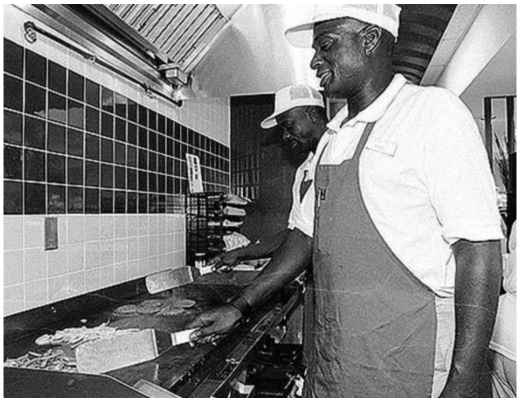


Fig. 2. A short-order cooking work environment.

between human and automated judgments: pilots demonstrated linear-additive cue utilization that focused heavily on a single environmental cue, whereas automation took into account a larger set of cues and used a nonlinear pattern of cue integration. Studies such as this have direct and immediate implications for training and interface design.

To take a final example, a recent study by Rothrock and Kirlik [26] demonstrated that humans are capable of learning and using nonlinear judgment strategies in some human–technology interaction contexts. They, therefore, provided an alternative to the linear-additive regression-modeling approach for identifying these nonlinear (rule-based) strategies within the lens model framework. Their alternative technique for inferring strategies from data on cues and judgments uses a combination of genetic algorithms for rule-based representation and search, and multiobjective optimization for evaluating the fitness (or adaptivity) of a rule set. This research demonstrates that adopting probabilistic functionalism does *not* commit one to any particular technique, such as multiple regression, for modeling proximal–distal relations. For additional details, see [26].

C. Applications in Dynamic, Closed-Loop Interaction With Tools and Technology

Recent decades have seen many advances in the design of perceptually rich (e.g., graphic, multimodal) information displays to support human–technology interaction. These displays narrow the proximal–distal gulf between available information and the distal states or events that need to be known (e.g., system status, fault information) for effective system control [10]. Developments aimed at bridging the proximal–distal gulf in the realm of action, however, have come much more slowly (control devices such as datagloves and gestural interfaces provide exceptions). To motivate the development of action-augmented technological interfaces, (e.g., beyond keyboard and mouse), we studied how performers in relatively less restrictive control environments, allowing much more opportunity and freedom of action, might use this additional freedom adaptively [6], [35].

To do so, we identified a low-technology environment in which interaction with the objects of work was relatively direct and as yet unmediated by information technology. Fig. 2 depicts the type of context studied: short-order cooking in a diner.

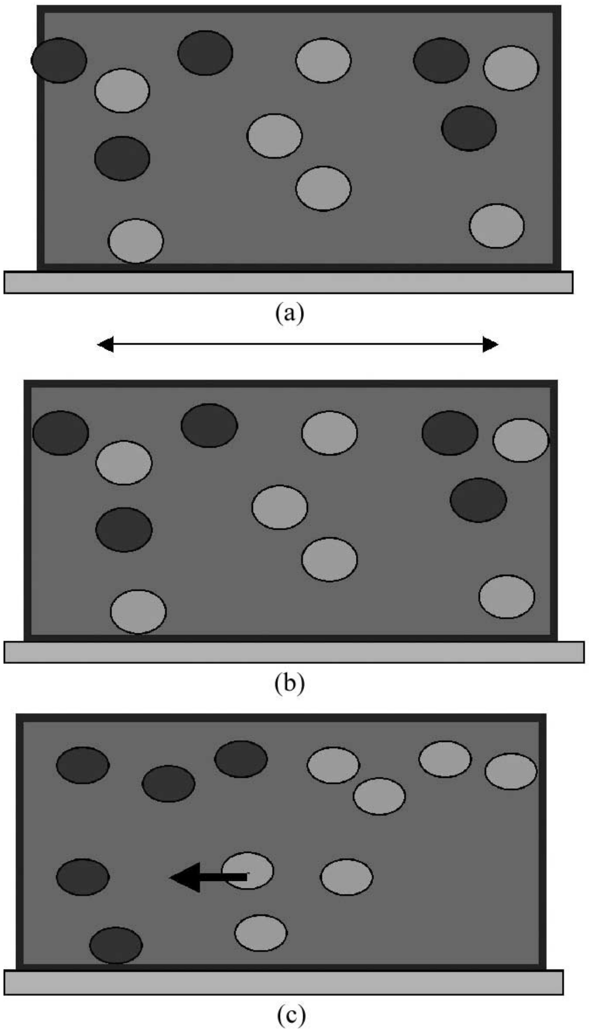


Fig. 3. The three observed short-order cooking strategies. (a) The random or brute force strategy. The grill contains no information about the doneness of the underside of the meats, nor how long each should be cooked. (b) The position control strategy. Meats to be cooked well done are placed to the left, mediums are in the center, and rares to the right. The grill, however, contains no information about the doneness of the undersides of the various meats. (c) The position + velocity control strategy. Meats to be cooked well are placed toward the rear and far right edge of the grill. Mediums are placed in the middle (front to back) and less far to the right. Rares are placed to the front and even less far to the right. All meats are intermittently moved to the left over time, and flipped halfway in their journey across the grill. Note that the all the information necessary to perform the task has been made observable in the grill layout.

We observed three different cooks using three different strategies for ensuring that each piece of meat (hamburgers) placed on the grill were cooked to the specified degree of doneness (rare, medium, or well). Fig. 3 graphically depicts the three strategies observed.

Fig. 3 depicts how more sophisticated strategies for the placement and movement of meats result in cognitive economies for the cook, in terms of reduced demands upon memory and internal modeling. Fig. 3(a) depicts the situation where the cook randomly places the meats on the

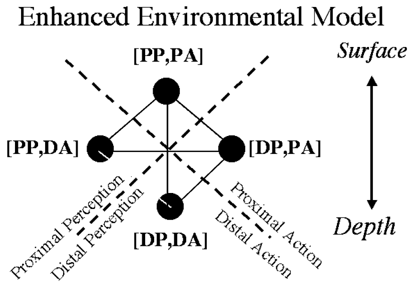


Fig. 4. The environmental components of Brunswik's lens model, augmented to describe the proximal–distal status of variables in the realms of both perception and action.

grill and uses no consistent policy for moving them, with the result that this cook's external environment contains relatively little task-relevant information. Fig. 3(b) depicts a strategy in which the initial placement of meats results in environmental information useful for knowing the degree to which each piece of meat should be cooked, thus eliminating the demand for the cook to keep this information in memory. Under the most sophisticated strategy, the position + velocity control strategy depicted in Fig. 3(c), the cook's placement and movement strategies allow one to consider the dynamic grill system itself as a fully observable, external model of the otherwise hidden dynamics and states of this interactive system. In this case, everything that the cook needs to know about the task is proximally available from the grill itself, and as such, the meats signal their own completion when they arrive at the left boundary of the grill.

To model this situation, we developed a technique formally extending Brunswik's lens model to the realm of action, in addition to perception. This model, depicted in Fig. 4, includes resources to describe whether a particular environmental variable (e.g., the doneness of the underside of a piece of meat) is either proximal (directly perceivable, directly manipulable) or distal (must be inferred, must be manipulated by manipulating intermediary variables).

In Fig. 4, [PP,PA] variables are proximal with respect to both perception and action: given an agent's perceptual and action capacities, their values can be both directly measured and controlled. [PP,DA] variables can be directly measured by the agent but cannot be directly manipulated. [DP,PA] variables, on the other hand, can be directly manipulated but cannot be directly perceived. Finally, [DP,DA] variables can be neither directly measured nor controlled. See [6] for a more complete presentation.

Each set of n such variables at a particular time t represents the state of the dynamic system, or a particular value of the state vector. Over time, which can be considered discrete to simplify the analysis, one can record a progression of the state variable values and treat this data as one would a contingency table, that is, as samples from a multivariate probability distribution. One can calculate the entropy of the distribution of any particular state variable a as $H(a)$, using the standard Shannon measure [42]

$$H(a) = -\sum_{a} p(a) \log_2 p(a).$$

Similarly, one can treat the entire state vector in a multivariate sense and calculate the entropy $H(S)$ of the dynamic system as a whole, thus, $H(S) = H(a, b, c, \ldots n)$ [43], [44]. As Conant [44] showed, this approach allows one to measure various dependencies within the system, for example, *conditional entropies* and *information transmission* values.

One useful application of these ideas is to measure the degree to which a particular subsystem of S corresponding to the (observable) proximal perceptual variables (cues) in S is informative about the complementary subsystem of S composed of the (unobservable) distal variables in S. This allows one to determine the degree to which the demands imposed by a task can be met merely by perception, rather than by also placing demands on memory or internal modeling. To make calculations such as these, one uses the concept of *conditional entropy*.

The system S can be partitioned into arbitrary subsystems $S1, S2, \ldots S_N$. In the cooking case, for example, we partitioned S into two mutually exclusive and collectively exhaustive subsystems SP and SD, corresponding to the proximal perceptual and distal perceptual variables in S, respectively, for a given agent. The concept of conditional entropy allowed us to measure the average amount of uncertainty about a particular subsystem of S that remains for one who knows the values of a complementary subsystem of S. In our cooking example, we can write $S = (\mathrm{SP}, \mathrm{SD})$ to indicate that the system S is made up of the subsystems SP and SD. We can write $H_{\mathrm{SP}}(\mathrm{SD})$ as the conditional entropy of SD given SP: that is, the average uncertainty about the variables in SD given knowledge of the variables in SP.

Conveniently, it is known that for any partitioning of S into subsystems $S1$ and $S2$

$$H_{S1}(S2) = H(S1, S2) - H(S1) = H(S) - \mathrm{H}(S1).$$

Thus, $H_{S1}(S2) = 0$ if the variables in subsystem $S2$ are completely determined by the variables in subsystem $S1$. The above result can be generalized to any number of subsystems for an arbitrary partitioning of S using an entropy "chain rule." For our purposes, we sought to measure $H_{\mathrm{SP}}(\mathrm{SD}) = H(S) - H(\mathrm{SP})$ which indicates, given a particular trajectory of the variables in the system S over time, the degree to which the surface variables (perceptible from the grill) are informative about the (covert) depth variables. Given actual streams of behavioral data corresponding to each of variable, and a particular partitioning of system variables into SP and SD, we can readily compute a scalar measure of the degree to which proximal variables specify distal variables.

In [6], we used this method (based on the enhanced lens model) along with techniques based on entropy and multidimensional information theory to analyze strategy differences in the short-order cooking study. The results of that analysis are depicted in Fig. 5. As the figure demonstrates, while the overall variability (entropy) of the system controlled by the most sophisticated (position + velocity) strategy is highest, a cook using this strategy has almost no uncertainty remaining

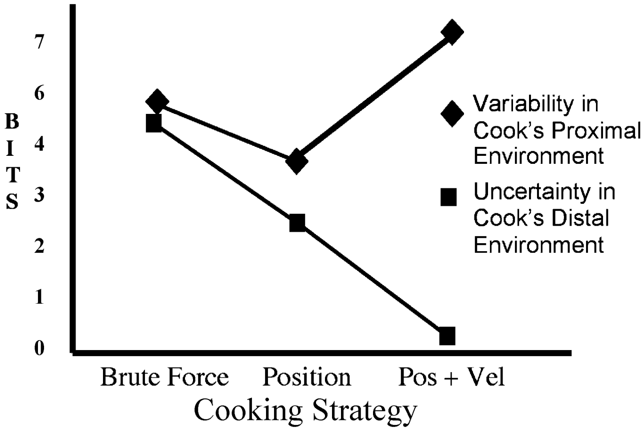


Fig. 5. Results of the short-order cooking analysis based on the enhanced version of Brunswik's environmental model (see text for details).

regarding the values of distal environmental variables relevant to the task.

Freedom of action allowed this cook to literally become his own real-time interface designer, in *creating* lawful relationships existing between perceptually available variables and variables that were once distal, and, thus, had to be either inferred, remembered, or generated by internal modeling by the other cooks. The improvisational and cognitively efficient behavior of this cook was enabled by the fact that the system design allowed him to have relatively free, unmediated control over the many variables in his work environment. We now consider an even less mediated situation involving adaptive control and coordination.

V. A STUDY OF CONDUCTOR–CONCERTMASTER COORDINATION IN A SYMPHONY ORCHESTRA

To this point, we have hoped to describe the flexibility and utility of Brunswik's probabilistic functionalism for the study of perception and performance in both music and human–technology interaction. As stated in the introduction to this paper, one of our central motivations is to enable a shared framework to enable cross-disciplinary research to be conducted and findings to be shared. More specifically, however, we believe that it is possible that the study of group musical performance, such as orchestra performance under the direction of a conductor, may provide valuable insights and new concepts or models for the design of more adaptive and robust sociotechnical systems.

To this end, we now present an initial study of the development of conductor–concertmaster coordination during the rehearsal and recording of Beethoven's Fifth Symphony, as performed by the Tokyo Mozart Players [45]. Before describing that work in detail, we want to make explicit that we have not formally analyzed the data collected from that study in terms of any of the statistical or mathematical models discussed previously in this paper. Those models make relatively heavy demands on the availability of data, and the study we present below, being one of the first of its kind ever performed, did not result in the collection of sufficient data for formal analysis and modeling. However, after describing the study, we do provide a sketch of how conductor–concertmaster coordination could possibly be formulated in terms of probabilistic functionalism and lens modeling.

A. Functions Involved in Conducting an Orchestra

What, exactly, is a conductor doing while directing a group musical performance? A variety of answers to this question have been given, and many basic techniques have been identified as required skills that a student conductor must learn to acquire. Student conductors are typically taught to master a set of systematic skills, which usually focus on how to show precise beat patterns with their hands and control an orchestra efficiently. Although a few textbooks on conducting have been published, they focus almost exclusively on the visual depiction of beat patterns [46], [47]. For instance, when a conductor shows a triangle with his hand, this gesture means three beats. Braem and Bräm [48] have produced a catalogue of the expressive gestures which are generally used by orchestra conductors to create particular musical effects. As suggested below, however, a conductor typically communicates with players through a broader and more varied range of bodily movements than those described in [48].

For example, Maruyama and Furuyama [49] analyzed videotaped data of rehearsals by a professional conductor and orchestra, and identified a variety of functional variations which could not be understood in terms of standard textbook accounts. Leyden [50] investigated the conducting patterns (mainly, beat patterns) demonstrated by A. Toscanini, one of the greatest conductors of the last century. He found that Toscanini used unique gestures to communicate beat patterns also quite distinct from the gestures described in textbooks. For these reasons, an observational study was performed to gain better insight into conductor–orchestra interaction.

B. Observational Study

As described in Maruyama [45], videotaped data of an actual performance by a professional conductor and orchestra was collected. We chose observational methods because a conductor's body movements in actual practice are more various and dynamic than descriptions of conducting practice found in the music literature. P. Boulez, generally considered to be one of the most important conductors and composers of our times, also pointed out the complexities that can be understood only when a conductor faces an actual orchestra, focusing particularly on educational problems:

> I've always taught it [conducting] with an orchestra, with players... That's because I think trying to teach conducting by means of a piano, or two pianos at the most, is absolutely useless. First, you don't obtain the real sound, and second you're not working with a large group, which is very important. Third, it doesn't correspond to the psychology of conducting orchestra [51].

Considering the above, we decided to collect data while observing a professional orchestra without any experimental restrictions. We believe that field observations of this type, though they sacrifice experimental control, are a crucial first

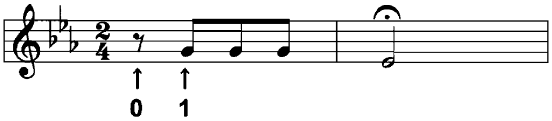


Fig. 6. The opening theme of Beethoven's Fifth Symphony (1808). **0** = the eighth rest (beat point by the conductor's hand stroke); **1** = the first note (start of the concertmaster's bowing action).

step toward identifying and distilling the central issues involved in any aspect of human performance, and especially performance as complex as group musical performance. All rehearsals and a recording session were videotaped with the agreement of the conductor and the players. The researchers provided no special instructions or other information during data collection.

C. The Conductor

R. Nuamjiri (1964-), an energetic Japanese professional conductor, was observed. Numajiri studied conducting under S. Ozawa, K. Akiyama, and T. Odaka at the Toho Gakuen School of Music. In 1990, he won the international conductor's competition at Besançon, France, and since 2000, he has been a principal conductor of the Tokyo Philharmonic, the major professional orchestra in Japan. Numajiri has also conducted many of the worlds' other major orchestras, such as the London Symphony and the Berlin Symphony.

D. The Orchestra

The Tokyo Mozart Players were organized by R. Numajiri as a professional chamber orchestra composed of mainly young, talented musicians in 1995. The Tokyo Mozart Players continue to give various concerts and produce CD recordings to the present day.

E. Data Information and Apparatus

Data were collected for three days. The first two days were spent as rehearsals and the third day was a recording session. Both conductor's and concertmaster's bodily actions were recorded using high speed cameras at 120 Hz (Victor GR-DVL-700). These data were analyzed with a computer-based motion coding system (DigiMo-Swallo2001).

F. The Music

The music performed during observation was the Fifth Symphony composed by L. van Beethoven. This is one of the most popular pieces classical music, especially the opening theme of the first movement (Fig. 6). However, this theme is actually extremely hard to play due to its unique rhythmic structure. Thus, this opening provides a good case study for examining how well the conductor and the players became coordinated. Initially, we focused on the dynamic process of relative timing between the conductor's body movements and the concertmaster's bowing action in preparing to play this opening theme.

G. Data Collection and Analysis

Detailed motion coding was performed by two coders who were members of a university student orchestra, and who had played under a professional conductor's direction. All measurements of timing were based on the time code of videotaped data. The following temporal intervals served as the focus of the analysis: 1) the gestural initiation and duration of the conductor's hand stroke movements; 2) the relative timing between the conductor's actions or gestures taken to indicate beat points, typically consisting of right hand movements (the concertmaster's bowing start point was considered as the time of the initiation of sound produced); and 3) the induction timing of the concertmaster's bowing, which was a preparation phase prior to the initiation of bowing. The opening theme was repeatedly performed; seven trials on day 1, six trials on day 2, and 11 trials in the recording session. However, data from the third trial on day 1 was excluded from analysis due to flaws in data collection.

H. Structure of the Conductor's Hand Stroke

The conductor conducted with a baton on days 1 and 2 of rehearsals. However, in the recording session, he conducted all trials without a baton. The outward forms of the conductor's hand strokes initially appeared to be extremely variable. However, subsequent analysis clearly revealed two kinds strokes among the three days. Fig. 7(a) depicts the *vertical* stroke with a baton, observed on days 1 and 2. Fig. 7(b) depicts the *horizontal* stroke without a baton, observed only in the recording session. The outward differences between these strokes were quite dramatic. When conducting was performed in the horizontal style, the conductor displayed especially energetic movements, as if the upper part of his body were strongly stretched out with both arms and hands flung to the side.

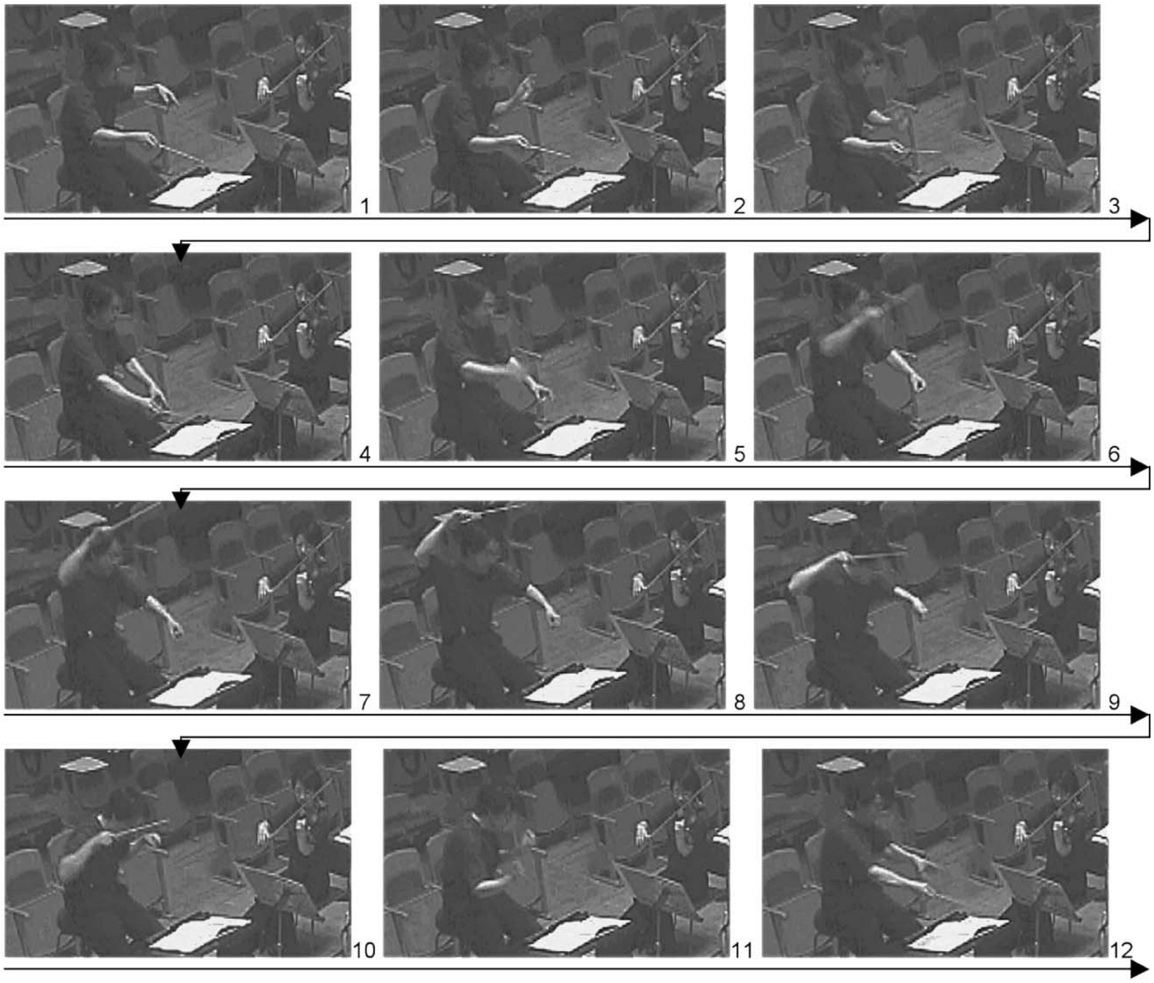
I. Progressive Change of the Eighth Rest Value

To examine potential adaptation between the behavior of the conductor and concertmaster, the interval between the end point (the beat point) of the conductor's hand stroke corresponding to the position of the eighth rest (indicated by "0" in Fig. 6), and the concertmaster's bowing start point (indicated by "1" in Fig. 6) was measured in all performances. Any increased level of structure in these durations may reflect increased coordination in conductor–concertmaster interaction over practice.

The average and standard deviation of these eighth rest values are shown in Fig. 8. Interestingly, the longer eighth rest values were generated in the recording session. However, the associated standard deviations of these times gradually decreased over practice.

J. Overall Temporal Course of Events

Fig. 9 illustrates the overall course of events observed in this study as a set of bar graphs. The length of each bar indicates the duration of each action. For instance, it can be seen that, on the first trial on day 1 [Fig. 9(a)], the conductor



(a)

Fig. 7. (a) *Vertical stroke*—Conducting sequence observed on day 1.

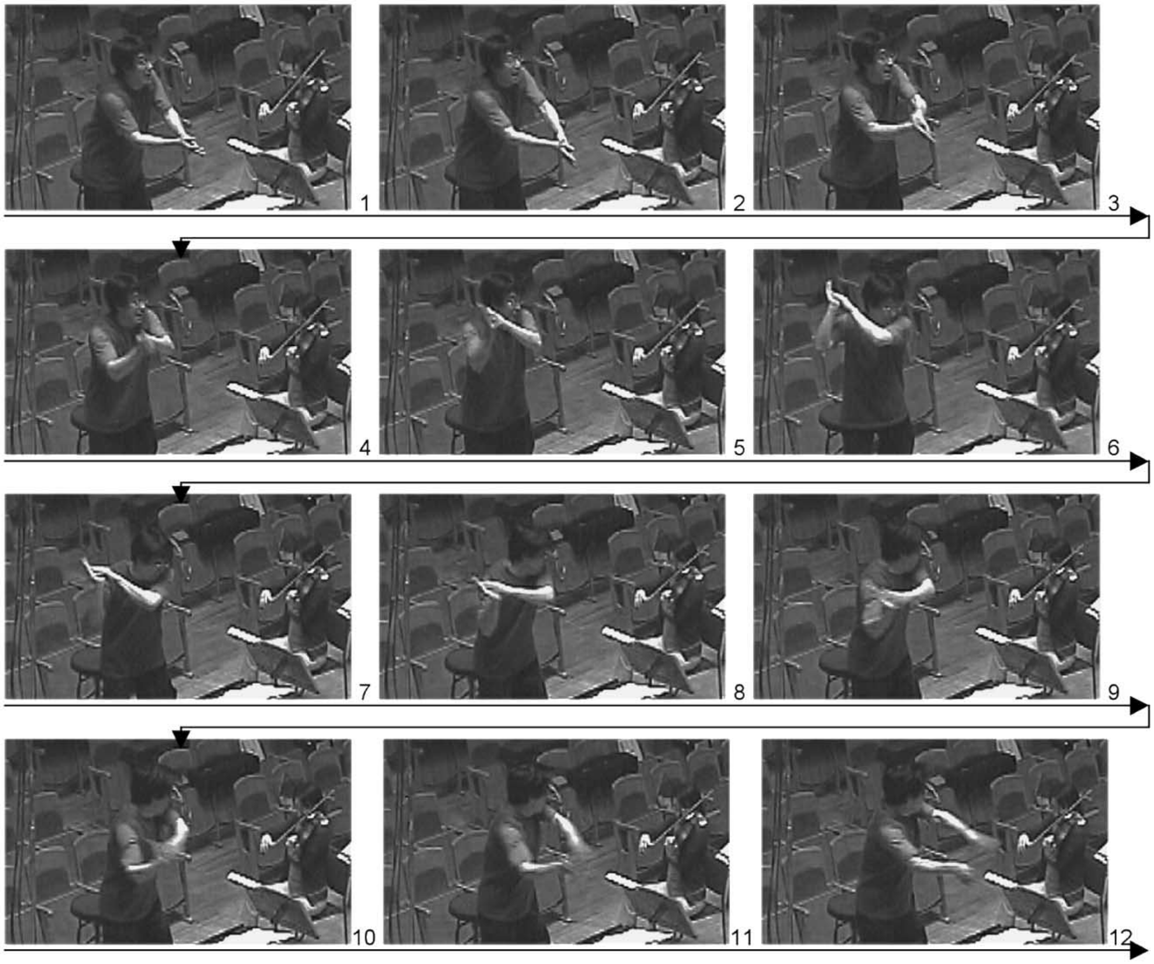
first began to move his left hand and, after a few milliseconds, his right hand followed his left hand. The four landmark temporal intervals discussed above can be understood as the length and the relative positions among the bars in each trial. It should be pointed out that, especially in the recording sessions, the relative positions among elements in each bar, and the lengths of each bar [e.g., Fig. 9(c)], tended to be more regular than in the two previous days of practice.

K. Interpretation

Detailed motion coding revealed that the conductor established and maintained a relatively stable timing structure in all his hand strokes, regardless of their highly varied outward forms, again, an illustration of vicarious functioning at work. The durations between the conductor's hand strokes and the concertmaster's bowing actions were gradually regulated over rehearsal. Although little can be firmly concluded on such a small set of observational data, some findings do suggest that on days 1 and 2 of rehearsal the conductor and the concertmaster were engaged in a process of developing a well-regulated timing interval, required to achieve a consistent and specific musical intention for the recording session.

That is, the larger time intervals (the longer values of the eighth rests) observed consistently in the recording session may reflect both the conductor's and concertmaster's musical intentions and an emergent feature of their mutual explorations. However, as these data indicate, the time interval between the beat point indicated by the conductor's hand stroke and the concertmaster's bowing start in the recording session became *longer* than in the previous two days. In naive conceptions of coordination, the degree of interpersonal coupling is reflected by synchronicity [52]. Under this reading, one might naively conclude that coordination between the conductor and the concertmaster was the lowest in the recording session.

This interpretation must be suspect, because the recording session followed two days of rehearsal, and the conductor indeed said that "the opening part is very good" during the recording session. In short, we suggest that interpersonal coordination should be analyzed in terms of the consistency of temporal structure, perhaps measurable by entropy or some related means, rather than in terms of the degree of synchronicity between people's behaviors. This interpretation is supported by the idea that rhythmic activity is directed by a meaning or a content which guides a person to generate



(b)

Fig. 7. (*Continued.*) (b) *Horizontal stroke*—Conducting sequence observed in the recording session.

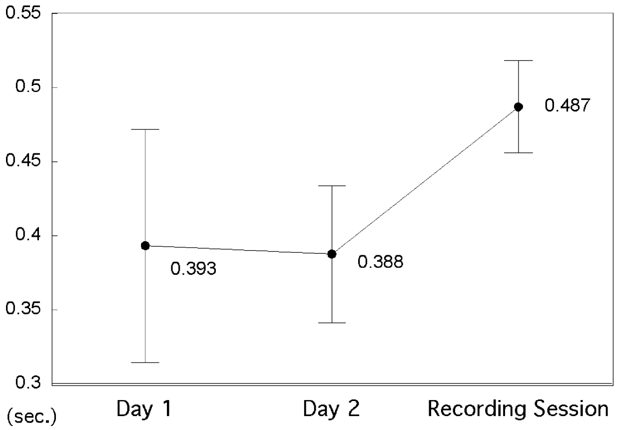


Fig. 8. The progressive change of the eighth rest values: the time interval between the beat point shown by the conductor's hand stroke (point **0** in Fig. 6) and the concertmaster's bowing initiation (point **1** in Fig. 6).

a comfortable and common tempo for rhythmic movements [53, p. 408]. In this research, the conductor and the concertmaster appeared to find a specific timing solution which was guided or constrained by the specific task requirements posed by the opening of Beethoven's Fifth Symphony.

L. Toward a Lens Modeling Formulation

Although the data collected in this initial study were sparse, we do see prospects for interpreting musical phenomena such as conductor–musician coordination from the perspective of probabilistic functionalism. In the present case, doing so would require identifying and coding the gestural movements of both conductor and concertmaster as a set of perceptually available cues, which are presumed to form the basis for communication and coordination. Akin to the Juslin [38] study described earlier, analysis and modeling would then focus on the manner in which the conductor (akin to Juslin's guitarists) created gestural cues to communicate intentions, most likely related to timing, but possibly also with respect to emotion or other stylistic goals. Analogously, one would analyze the concertmaster's judgment policies for selecting and weighting these cues in action selection and timing. Noting that the concertmaster's actions create perceptible cues for the conductor, it may even be able to provide a dynamic, closed-loop account of coordination strategies in the spirit of the short-order cooking study described above.

While coding gestural cues in social interaction may indeed be difficult, we note that Gillis and his colleagues

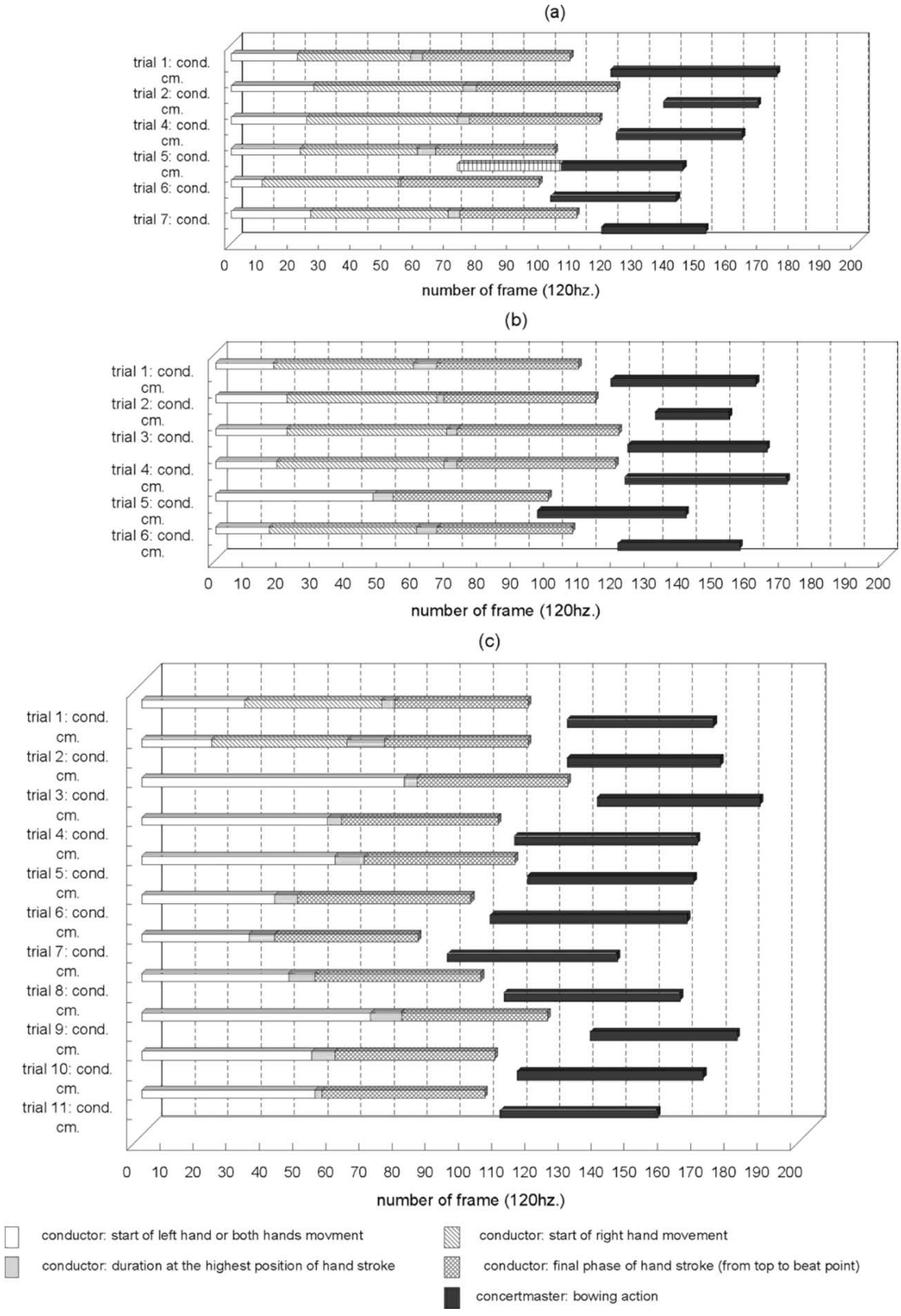


Fig. 9. The overall patterns of conductor's and concertmaster's actions and interactions ("cond": conductor; "cm": concertmaster). (a) Day 1. (b) Day 2. (c) Recording session.

[54] have done so with reasonable success in their studies of face-to-face rapport, for which a dictionary definition is "relationship or communication, esp. when useful or harmonious." Working within the framework of probabilistic functionalism, these researchers collected a video archive of 120 male and female college students paired in both

cooperative and mildly adversarial contexts. Using multiple human coders, they found that they could reliably assess more than 70 cues from these videos, including information such as smile frequency and duration, sitting distance, the frequency of encroachment and withdrawal movements, and so on. The goals of this research were to understand how and if postencounter self-reports of rapport or lack of rapport (by both males and females) could be understood via probabilistic functionalism.

Using lens model analysis based on observing and coding interactions between male and female dyads, Gillis and his colleagues found that over half of the variance in participants' self-judgments of rapport could be reliably be predicted by a mere five cues: 1) frequency and duration of silences; 2) amount of female gesturing; 3) physical distance maintained between dyads; 4) number of postural shifts; and 5) the racial mix of the dyad. Another notable finding was that, when novel observers of the videotaped interactions were asked to predict the level of rapport reported by the dyads being observed, their judgments were significantly compromised by overweighting what the researchers called "animation" cues, such as the frequency of smiling and the extent of gesturing. These animation cues, while highly salient to observers, actually had very little ecological validity (in predicting rapport), and a subsequent experimental intervention, explicitly instructing observers to ignore such cues, had no effect on performance.

Research in social interaction such as this suggests that there is potential for the modeling and analysis of musical coordination, a type of rapport, from the perspective of probabilistic functionalism. We look forward to bringing these methods to bear on the study of group musical performance in future research.

VI. CONCLUSION: IMPROVISATION AND ROBUSTNESS IN MUSIC AND HUMAN–TECHNOLOGY INTERACTION

We have presented a theoretical framework providing a common language and sets of techniques amenable to investigating perception and performance in both human–technology interaction and music. We truly hope that this framework fosters increased cross-disciplinary research and, especially, improves prospects for increasingly robust design of sociotechnical systems. Group musical performance is notoriously robust and adaptive, and may, therefore, serve to inspire novel design models for sociotechnical systems.

We conclude with a brief discussion of improvisation, because we believe that it is a topic where engineers may have much to learn from those in the music domain. As discussed in early sections of this article, the goal of reducing process variance in purely technological systems has proven useful and cost effective. As a technique for designing human work, however, it results in brittle human performance and stressful working conditions. Data showing that even the most intelligently designed procedures and instructions are insufficient for adequately supporting human performance are not hard to find [10]. Still, no good theory or model of human performance in engineering systems yet exists that gives due regard to the knowledge an operator (pilot, physician, etc.) needs to know *when* to apply a procedure, *when* not to, or *how* to tweak or tailor an existing procedure to meet the demands of a novel situation.

We know that metacognitive and improvisational abilities are key contributors to the robust and adaptive operation of sociotechnical systems, although we have not yet been able to pin down exactly what these abilities are, how to train them, or how to design interfaces to support them. Until we have such a theory, it will continue to prove difficult to defend against the ever-increasing use of proceduralization and automation in sociotechnical systems.

While it may be a difficult task to defend the need to allow the human operator latitude for responsible improvisation in many technological systems, improvisation has always played a central role in the theory and practice of music performance. Indeed, most would find it a laughable notion to think that music performance could be enhanced across the board by eliminating opportunities for improvisation. Consider what Meyer had to say on this point:

> The musical relationships embodied in a score or handed down in an oral tradition do not fix with rigid and inflexible precision what the performer's actualization of the score or aural tradition is to be. They are indications, more or less specific, of what the composer intended and what tradition has established. The performer is not a musical automaton or a kind of musico-mechanical medium through which a score or tradition is realized in sound. The performer is a creator who brings to life, through his own sensitivity of feeling and imagination, the relationships presented in the musical score or handed down in the aural tradition which he has learned [28, p. 199].

Meyer goes on to devote two chapters of his book to the crucial role that "deviations" from the written score and oral tradition play in creating high-quality musical performances, noting that it hardly speaks well of a performer to say that he or she "merely played the notes" or played "mechanically."

Currently, however, our experience is that design and training in many sociotechnical systems proceed all to often as if "doing it by the book" or working "like a machine" were admirable qualities. Experienced human operators know otherwise, and in their better moments, so do engineers, researchers and practitioners in human–technology interaction. Investigations of the constrained liberation underlying musical performance may hold promise for the development of a theory of responsible improvisation that could have significant social value (e.g., see "The etiquette of improvisation" by noted sociologist and jazz pianist H. Becker [55]). We have presented a theoretical framework that may allow scientific investigations of these issues as they naturally occur in musical contexts. Whether research such as this will yield insights into important engineering problems, however, remains to be seen.

ACKNOWLEDGMENT

The authors would like to thank P. Juslin, K. Vicente, and five anonymous reviewers for valuable comments on a previous version of this paper.

REFERENCES

[1] T. Inagaki and J. Stahre, "Human supervision and control in engineering and music: similarities, dissimilarities, and their implications," *Proc. IEEE*, vol. 92, pp. 589–600 , Apr. 2004.
[2] E. Brunswik, *Perception and the Representative Design of Psychological Experiments*. Berkeley: Univ. California Press, 1956.
[3] E. C. Tolman and E. Brunswik, "The organism and the causal texture of the environment," *Psychol. Rev.*, vol. 42, pp. 43–77, 1935.
[4] A. Bisantz, A. Kirlik, P. Gay, D. Phipps, N. Walker, and A. D. Fisk, "Modeling and analysis of a dynamic judgment task using a lens model approach," *IEEE Trans. Syst., Man, Cybern. A*, vol. 30, pp. 605–616, Nov. 2000.
[5] A. Bisantz and A. Pritchett, "Measuring the fit between human judgments and automated alerting algorithms: a study of collision detection," *Hum. Factors*, to be published.
[6] A. Kirlik. (1998) The ecological expert: acting to create information to guide action. *Proc. 4th Symp. Human Interaction With Complex Systems* [Online]. Available: http://computer.org/proceedings/hics/8341/83410015abs.htm
[7] S. A. Reveliotis, M. A. Lawley, and P. M. Ferreira, "Polynomial complexity deadlock avoidance policies for sequential resource allocation systems," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1344–1357, Oct. 1997.
[8] T. B. Sheridan, *Telerobotics, Automation, and Supervisory Control*. Cambridge, MA: MIT Press, 1992.
[9] ——, *Humans and Automation*. New York: Wiley, 2002.
[10] K. J. Vicente, *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Erlbaum, 1999.
[11] P. J. Ramadge and W. M. Wonham, "Supervisory control of a class of discrete event processes," *SIAM J. Control Optimizat.*, vol. 25, no. 1, pp. 206–230, 1987.
[12] M. Tittus, "Control synthesis of batch processes," Ph.D. dissertation, School Electr. Comput. Eng., Dept. Control Eng., Chalmers Univ. Technology, Göteborg, Sweden, 1995.
[13] Z. A. Banazak and B. H. Krogh, "Deadlock avoidance in flexible manufacturing systems with concurrently competing process flows," *IEEE Trans. Robot. Automat.*, vol. 6, pp. 724–734, Dec. 1990.
[14] G. Decknatel and E. Schnieder, "Modeling railway systems with hybrid petri nets," in *Proc. Hybrid Dynamical Systems: 3rd Conf. Automation of Mixed Processes*, J. Zaytoon, Ed., 1997, pp. 309–315.
[15] R. Kanigel, *The One Best Way: Frederick Winslow Taylor and the Enigma of Efficiency*. New York: Penguin, 1999.
[16] W. Langewiesche, "American ground: Unbuilding the World Trade Center, part three: The dance of the dinosaurs," *Atlantic Monthly*, vol. 290, pp. 92–126, Oct. 2002.
[17] Santa Fe Inst., "A founding program in the study of robustness," *Santa Fe Inst. Bull.*, vol. 15, no. 2, pp. 1–9, 2000.
[18] J. M. Flach and J. Rasmussen, "Cognitive engineering: designing for situation awareness," in *Cognitive Engineering in the Aviation Domain*, N. B. Sarter and R. Amalberti, Eds. Mahwah, NJ: Erlbaum, 2000, pp. 153–180.
[19] E. Hollnagel, "Modeling the orderliness of human action," in *Cognitive Engineering in the Aviation Domain*, N. B. Sarter and R. Amalberti, Eds. Mahwah, NJ: Erlbaum, 2000, pp. 65–98.
[20] E. Hollnagel and D. D. Woods, "Cognitive systems engineering: new wine in new bottles," *Int. J. Man–Machine Stud.*, vol. 18, pp. 583–600, 1983.
[21] D. D. Woods, "Paradigms for intelligent decision support," in *Intelligent Decision Support in Process Environments*, E. Hollnagel, G. Mancini, and D. D. Woods, Eds. New York: Springer-Verlag, 1986, pp. 153–174.
[22] R. Parasuraman and V. Riley, "Humans and automation: use, misuse, disuse, abuse," *Hum. Factors*, vol. 39, no. 2, pp. 230–253, 1997.
[23] K. Akesson, S. Jain, C. Yuan, and P. Ferreira, "Hybrid computer-human supervision of discrete-event systems," presented at the IEEE Conf. Robotics and Automation, Washington, DC, 2002.
[24] A. Degani and M. Heymann, "Formal verification of human–automation interaction," *Hum. Factors*, vol. 44, no. 3, pp. 378–396, 2002.
[25] A. Kirlik, "Modeling strategic behavior in human–automation interaction: why an 'aid' can (and should) go unused," *Hum. Factors*, vol. 34, no. 2, pp. 221–242, 1993.
[26] L. Rothrock and A. Kirlik, "Inferring rule-based strategies in dynamic judgment tasks," *IEEE Trans. Syst., Man, Cybern. A*, vol. 33, pp. 58–72, Jan. 2003.
[27] D. T. McRuer and H. R. Jex, "A review of quasilinear pilot models," *IEEE Trans. Hum. Factors Electron.*, vol. HFE-8, pp. 231–249, Sept. 1967.
[28] L. B. Meyer, *Emotion and Meaning in Music*. Chicago, IL: Univ. Chicago Press, 1956.
[29] J. J. Gibson, "Survival in a world of probable objects: review of Egon Brunswik, 1956, perception and the representative design of psychological experiments," *Contemporary Psychol.*, vol. 2, no. 2, pp. 33–35, 1957.
[30] A. Kirlik, "On Gibson's review of Brunswik," in *The Essential Brunswik*, K. R. Hammond and T. R. Stewart, Eds. New York: Oxford Univ. Press, 2000, pp. 238–242.
[31] R. Cooksey, *Judgment Analysis: Theory, Methods, and Applications*. New York: Academic, 1996.
[32] K. R. Hammond and T. Stewart, *The Essential Brunswik: Beginnings, Explications, and Applications*. New York: Oxford Univ. Press, 2001.
[33] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2000.
[34] M. D. Byrne and A. Kirlik, "Integrated modeling of cognition and the information environment: Closed-loop, ACT-R modeling of aviation taxi errors and performance," Univ. Illinois, Urbana, Inst. Aviat. Tech. Rep. AHFD-02-19/NASA-02-10, 2002.
[35] A. Kirlik, "The design of everyday life environments," in *A Companion to Cognitive Science*, W. Bechtel and G. Graham, Eds. Oxford, U.K.: Blackwell, 1998, pp. 702–712.
[36] A. Monk, "Modeling cyclic interaction," *Behav. Inf. Technol.*, vol. 18, no. 2, pp. 127–139, 1999.
[37] P. N. Juslin and H. Sloboda, *Music and Emotion: Theory and Research*. Oxford, U.K.: Oxford Univ. Press, 2001.
[38] P. N. Juslin, "Cue utilization in communication of emotion in musical performance: relating performance to perception," *J. Exp. Psychol.: Hum. Percept. Perform.*, vol. 26, pp. 1797–1813, 2000.
[39] ——, "Communicating emotion in music performance. A review and a theoretical framework," in *Music and Emotion: Theory and Research*, P. N. Juslin and J. Sloboda, Eds. Oxford, U.K.: Oxford Univ. Press, 2001, pp. 309–337.
[40] P. N. Juslin, A. Friberg, and E. Schoonderwaldt, "Teaching the thing that cannot be taught: feedback-learning of musical expressivity," in *Enhancing Musical Performance: A Resource for Performers, Teachers, and Researchers*, A. Williamson, Ed. New York: Oxford Univ. Press, to be published.
[41] A. Kirlik, Ed., *Adaptation in Human–Technology Interaction: Models, Methods and Measures*. New York: Oxford Univ. Press, to be published.
[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[43] R. C. Conant, "Laws of information which govern systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 240–255, Apr. 1976.
[44] W. J. McGill, "Multivariate information transmission," *Psychometrika*, vol. 9, no. 2, pp. 97–116, 1954.
[45] S. Maruyama, "What information is explored and shared in the orchestra performance?," presented at the 11th Int. Conf. Perception and Action, Storrs, CT, 2001.
[46] E. A. H. Green, *The Modern Conductor*. New York: Prentice-Hall, Inc., 1961.
[47] M. Rudolf, *The Grammar of Conducting*. New York: Schirmer Books, 1950.
[48] P. B. Braem and T. Bräm, "A pilot study of the expressive gestures used by classical orchestral conductors," in *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, K. Emmorey and H. Lane, Eds. Mahwah, NJ: Erlbaum, 2000, pp. 143–167.
[49] S. Maruyama and N. Furuyama, "Functional variations and organization of expressive gestures produced by a professional orchestral conductor," presented at the 1st Int. Congr. Gesture: The Living Medium, Austin, TX, 2002.
[50] N. F. Leyden, "A Study and analysis of the conducting patterns of Arturo Toscanini as demonstrated in kinescope films," Ph.D. dissertation, Columbia Univ., New York, 1968.

[51] P. Boulez and J. Vermeil, *Conversations With Boulez: Thoughts on Conducting*. Pompton Plains, NJ: Amadeus Press, 1996. transl. from French by C. Naish.

[52] F. Bernieri and R. Rosenthal, "Interpersonal coordination: behavior matching and interactional synchrony," in *Fundamentals of Non-Verbal Behavior*, R. S. Feldman and B. Rime, Eds. New York: Cambridge Univ. Press, 1991, pp. 401–432.

[53] P. N. Kugler and M. T. Turvey, *Information, Natural Law and the Self-Assembly of Rhythmic Movement*. Mahwah, NJ: Erlbaum, 1987.

[54] J. S. Gillis and F. Bernieri, "The perception and judgment of rapport," in *The Essential Brunswik*, K. R. Hammond and T. R. Stewart, Eds. New York: Oxford Univ. Press, 2000, pp. 380–384.

[55] H. Becker, "The etiquette of improvisation," *Mind, Culture, Activity*, vol. 7, no. 3, pp. 171–176, 2000.

[56] J. J. Gibson, "Survival in a world of probable objects: review of Egon Brunswik, 1956, perception and the representative design of psychological experiments," in *The Essential Brunswik*, K. R. Hammond and T. R. Stewart, Eds. New York: Oxford Univ. Press, 2000, pp. 244–246.

Alex Kirlik received the Ph.D. degree in industrial and systems engineering from Ohio State University, Columbus, in 1989.

He was with the Georgia Insititute of Technology and has held visiting positions with Stanford University, Yale University, Haskins Laboratories, and NASA Ames Research Center. He is currently Associate Professor of aviation human factors, psychology, and mechanical and industrial engineering (affiliate), and Member of the Beckman Institute, at the University of Illinois, Urbana–Champaign. His research focuses on understanding and supporting cognitive aspects of human–technology interaction.

Shin Maruyama received the B.S. degree in human science from Waseda University in 1996 and the M.S. degree in education from the University of Tokyo in 2000.

From 2000 to 2003, he was a Student in the Graduate School of Education, University of Tokyo, and a Research Fellow of Japan Society for the Promotion of Science. Since 2003, he has been a Research Scholar, Department of Psychology, Indiana University, Bloomington. His research interests are in the development and acquisition of expert motor skills, and especially orchestral conducting. Most recently, he is working toward applying dynamic systems approaches to achieve qualitative descriptions of the dynamics underlying conducting, through the kinematic analysis of the conductor's body movements.