Journal of Motor Behavior, 1992, Vol. 24, No. 1, 3–28

The Intentional Spring: A Strategy for Modeling Systems That Learn to Perform Intentional Acts

Robert E. Shaw
Endre Kadar
Mikyoung Sim
Center for the Ecological Study of Perception and Action,
The University of Connecticut

Daniel W. Repperger
Armstrong Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base

ABSTRACT. In motor task learning by instruction, the instructor's skill and intention, which, initially, are extrinsic constraints on the learner's perceiving and acting, eventually become internalized as intrinsic constraints by the learner. How is this process to be described formally? This process takes place via a forcing function that acts both as an anticipatory (informing) influence and a hereditary (controlling) influence. A mathematical strategy is suggested by which such intentions and skills might be dynamically learned. A hypothetical task is discussed in which a blindfolded learner is motorically instructed to pull a spring to a specific target in a specific manner. The modeling strategy involves generalizing Hooke's law to the coupled instructor–spring–learner system. Specifically, dual Volterra functions express the anticipatory and hereditary influences passed via an instructor-controlled forcing function on the shared spring. Boundary conditions (task goals) on the instructor–spring system, construed as a mathematical (self-adjoint) operator, are passed to the learner–spring system. Psychological interpretation is given to the involved mathematical operations that are passed, and mathematical (Hilbert-Schmidt's and Green's function) techniques are used to account for the release of the boundary conditions by the instructor and their absorption by the learner, and an appropriate change of their power spectra.

Key words: adjoint control, intention, learning

Section 1. Introduction

Dynamical learning theory must explain how a task-defined intention becomes internalized by a learner. In learning a new skill, by definition, the task intention must be imposed on the learner from the outside, say, by feedback from an instructor or by trial and error. Operationally speaking, a task-defined intention can be considered to be learned when it comes to act as an intrinsic rather than extrinsic constraint on the learner's perceiving–acting cycle. In other words, imposed constraints must become assimilated constraints so that one's actions are self-controlled. Exploring candidate mathematical techniques for expressing this process of assimilating intention into self-control—what one might call *intentional learning*—is the central issue of this paper.

Our primary aim is to suggest a modeling strategy rather than a model, by means of which assimilation of the task-intention to self-control might be represented formally in models of the same mathematical type. Hence, our goal is not to consider the relative worth of competing models. Rather, it is to provide a generic theoretical framework for treating learning as a problem for dynamical systems theory in general—a framework that has come to be called *intentional dynamics* (Kugler, Shaw, Vincente, & Kinsella-Shaw, 1990; Shaw, 1987; Shaw & Kinsella-Shaw, 1988; Shaw, Kugler, & Kinsella-Shaw, 1990).

Students of movement science recognize that a variety of mechanical models described by ordinary differential equations (ODE) can be used to characterize a broad class of diverse phenomena in the field of self-regulated actions (Berstein, 1967; Hinton, 1984; Kugler & Turvey, 1987; Whiting, 1984). One of the most commonly applied types is the second-order ODE (SODE) commonly used to express Newton's law ($F = m\ddot{x}$) and Hooke's

Correspondence address: Department of Psychology, University of Connecticut, 406 Babbidge Road, Storrs, CT 06269-1020.

law ($F = -kx$). A large variety of models use SODEs to characterize muscle strain or limb movement as dynamical systems, such as springs or the formally similar pendula. Numerous experiments have supported the spring-like behavior of various aspects of the motor system (Bizzi, 1980; Cooke, 1980; Feldman, 1966a, 1966b, 1980, 1986). Related studies have treated action system components as coupled oscillators (Beek, 1989; Beek & Beek, 1988; Kelso & Kay, 1987; Kugler & Turvey, 1987; Schmidt, 1990). Why this variety of dynamical action models? Perhaps, this variety exists because the basic biomechanical, physiological, and psychological processes underlying actions are not yet perfectly understood.

Most models address the problem of coordination at the biomechanical or physiological levels of analysis. A remaining question is whether SODE-based models, regardless of their differences, permit any common techniques for handling the role of intention in learning—a purely psychological level of analysis. Are there general mathematical techniques that the family of SODE models might use to describe how an extrinsically imposed task intention becomes internalized by a successful learner so that intention acts as an intrinsic rather than an extrinsic constraint? An attempt to answer this question in the affirmative is a major focus of intentional dynamics.

Specifically, the class of dynamical models will be extended to cover intentionally *guided learning processes*. Such processes are required if a system is to adapt to environmental exigencies, to develop requisite skills, and to improve skill accuracy in pursuit of goal-directed behaviors, especially those that achieve success only after repeated attempts. Although most forms of learning may involve repeated efforts, not all do. (For instance, warm-up effects, one-trial learning, adventitious learning, transfer effects, and learning sets, i.e., learning to learn, are examples of learning effects that do not necessarily require repeated efforts. Hence, another requirement for realistic learning theories is to show how all such learning effects might fall naturally under adaptive dynamical models.)

The essence of mathematical modeling goes beyond the arbitrary ascription of properties of natural systems to the formal properties of mathematical equations. For the dynamical model to be semantically appropriate, the formulation of its equations must mirror the intrinsic arrangement of the most significant properties of the natural system. This means that the ascription of mathematical properties to their referent natural properties must not be arbitrary but must be abstractly equivalent (isomorphic) at some critical level of analysis. Whereas legitimate functional criteria exist by which models might be evaluated, such as their predictive power, there are also structural criteria, such as their descriptive power. Scientists often overlook the importance of the latter requirement because its usefulness is less obvious than the former. Although the predictive requirement is seen as pragmatic, the descriptive is often spoken of pejoratively as merely aesthetic. There are notable exceptions to this opinion, however.

Einstein stressed that the physical properties of fields must be taken as real rather than as mere fictions, whereas mechanical theories with the ether assumption should not, even though they may be functionally equivalent. Thus, competing, functionally equivalent mathematical models that are descriptively different are not to be granted the same ontological status (Einstein & Infeld, 1938/1966). In a similar vein, it has been argued vigorously that Turing machines provide a better class of models for mental processes than connection machines because the former have intrinsic formal properties (e.g., compositionality) that better reflect cognitive processing than the latter (Fodor & Pylyshyn, 1986).

Our interest in these issues was heightened when certain semantic shortcomings of ODE-based models became clear to us. ODE models treat as extrinsic and arbitrary key properties of the intentional learning process that, for psychological reasons, should be treated as intrinsic and necessary to this process. (The extrinsic properties of concern are discussed in detail in Section 2.) Functional analysis provides another technique for modeling, what are known as *integro-differential equations* (IDE; Appendix B) that, following Shaw and Alley (1985), may prove potentially more appropriate (also see Newell, 1991). Although ODE models can be constructed that are formally equivalent to any IDE-based models, the two classes of models are not semantically equivalent with respect to their psychological import; nor are they equally convenient formulations to solve. It is fair to say that behavioral scientists have much more experience formulating and solving ODE-based models for the same processes. But this familiarity and convenience does not automatically make the former better semantic models for intentional learning than the latter.

Consequently, for ease of solution, it is often better to formulate a problem initially as an ODE and then, for ease of semantic interpretation, to reformulate it later as an IDE. A method for carrying out these two steps will be outlined and discussed. The best psychological models are those whose mathematics not only make interesting predictions but whose intrinsic structure allows for the most natural ascription of the relevant psychological properties. It seems reasonable, therefore, that one consider incorporating methods (e.g., Sturm-Liouville method) into our modeling strategies for translating ODE models into formally equivalent but semantically distinct IDE models. Before considering the techniques for translating ODEs into IDEs, the claim that IDEs provide a semantically more appropriate strategy for modeling intentional learning processes than ODEs must be justified. Otherwise, why should one bother?

Intentional Learning as a Volterra Equation (IDE)

After reviewing the problems inherent in drawing learning curves to represent learning data, Shaw and Alley (1985) concluded that, regardless of one's learning theory, learning functions necessarily exhibit a set of essential properties. If dynamical models of learning processes are to be formally adequate, they must express the following facts:

1. Learning functions are continuously cumulative, which is formally equivalent to being monotonically increasing functions.
2. The cumulativity is directional in time, being positively directed when constrained by hereditary influences (e.g., reinforcement) and negatively directed when constrained by anticipatory influences (e.g., expectancies).
3. The generic form of learning functions is nonlinear, although learning in the linear range comprises a special but important case. (Here, however, we restrict our approach to treating learning in terms of linear operators.)

Experts typically agree that learning involves a cumulative function of some kind. Cumulativity expresses the fact that the effects of experience are continuously incremental. The continuity assumption expresses the fact that cumulative effects carry over to change later performance. The directionality property expresses the fact that learning is not only positively monotonic—producing a savings in time, errors, effort, or some other measurable quality that might be used to assess successful task performance—but directional in time. Traditionally, learning through reward assumes a past-pending, time-forward influence whereas learning through expectation assumes a future-tending, time-backward influence. By *time-forward* and *time-backward* influences is meant causal versus intentional forms of learning. Mathematically, such time-inverted influences are not at all mysterious; they refer to directional integrals. To appreciate the reciprocal role of hereditary and anticipatory influences, it is necessary to decompose the general form of learning functions into their component variables.

Any function used to represent learning consists of two components: a *state* variable and a *response* variable. The state variable expresses the disposition of the organism to learn, and the response variable expresses the change in the behavior of the system. The state variable can be disposed to facilitate learning, in the sense of "learning to learn" (Bransford, Stein, Shelton, & Owings, 1981) and warm-up effects, or disposed to inhibit learning, in the sense of fatigue effects. The hereditary influence (or time-forward disposition) and anticipatory influence (time-backward disposition) have a formal analogue in the generalized version of Hooke's famous law. The linear form of Hooke's law, $F = -kx$ (see Figure 1), treats the elastic coefficient as if it were a constant, when, in fact, it is a variable whose value changes as a function of frequency of usage (Lindsay & Margenau, 1957). Depending on its material composition, a spring can become harder or softer with usage, as a function of whether the elastic coefficient increases or decreases, respectively. For instance, a metal spring will soften under usage as more and more microfractures occur, infirming its elasticity and sending it further into the plastic performance range. By contrast, as muscle tissue fatigues from overuse, it becomes increasingly hard as a spring. Thus, under this general interpretation, the stiffness parameter, k, is no longer a constant but assumes a variable disposition. The fact that both learning systems and springs are dynamical systems with dispositional parameters that may change with usage suggests that a mathematical formulation that fits one might fit the other equally well. (Caveat: Please note that in applying the generalized spring in this article, for mathematical simplicity we assumed that no hereditary influences on learning are due to fatigue effects.)

In its simple form, Hooke's law expresses stress as a linear function of strain under the fixed initial condition, k = constant. By contrast, the generalized version of Hooke's law expresses stress as a function of strain and a variable k, whose linear or nonlinear course of values defines another function. This means that the generalized version must formulate Hooke's law as a function of variables as well as of another function. A function of a function is called a *functional*. Consequently, our initial strategy is to treat learning as a dispositional functional that may assume a hereditary (time-forward) or anticipatory (time-backward) integral form.

The proper treatment of dispositional functionals proved to be a difficult problem in physics until the great Italian mathematician Vito Volterra introduced the notion of an *integro-differential equation* (Kramer, 1970). The techniques needed to formulate the generalized form of Hooke's law can be found in a variant on

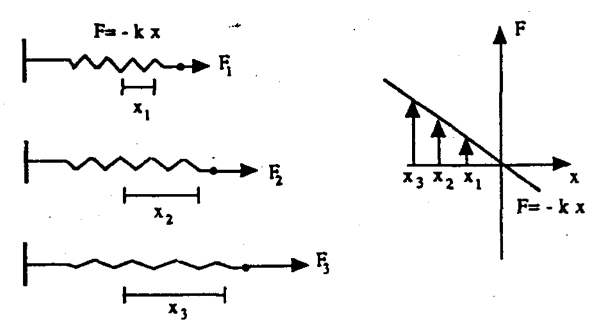

FIGURE 1. Inference to stiffness (k) from stress (F) and strain (x) values.

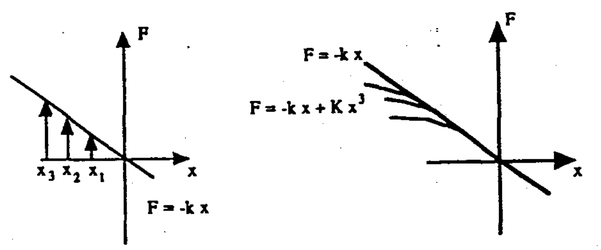

FIGURE 2. Simple Hook's law (left) as contrasted with the generalized version (right), illustrating hereditary changes in stiffness (k).

classical mechanics known as *hereditary mechanics* (Picard, 1907). These techniques have also been used for modeling living systems that depend on hereditary functionals, especially in biomechanics (e.g., Fung, 1981). The most famous is probably the Volterra-Lotka equation for competing ecological systems. Our spring analogue encourages us to borrow the generic integro-differential equation form used in physics to treat learning as being governed by a law of hereditary mechanics, as shown below:

$$y(t) = kx(t) + \int_0^t K(t, s)\, x(s)\, ds. \tag{1.1}$$

The three properties previously ascribed to the learning process now find expression under this Volterra equation. Here y(t) is the *response (behavioral) variable* at time t and x(t) is the *state (dispositional) variable* at time t. The k term represents a scaling value of the state variable, which denotes the initial capacity to learn. (The existence of the kx(t) term makes the linear IDE inhomogeneous.) (For derivation of this IDE, see Appendix B.)

Perhaps the best way to appreciate this formal analogy between springs and learning systems is to draw specific parallels. For instance, in the case of the generalized spring, k is a constant representing its initial stiffness at a given time before being pulled again; or in the case of learning, k is a constant representing the initial amount of learning of the system at a given time before additional practice. K(t, s) is an operator that represents an integral transform, called the *coefficient of hereditary influence.* K is a function of additional strains or practice, which adds incrementally and continuously to the constant k over the interval from s to t.

Furthermore, as will be shown later, if the interval is temporally inverted so that the integral transform reads K(s, t), then the influence of this operator is said to be *anticipatory* rather than hereditary, and to be an IDE of the anticipatory rather than hereditary type. Hence, in the linear case, the symmetry of the integral transforms, K(t, s) = K(s, t) (i.e., Green's operator) opens up the possibility that hereditary and anticipatory influences might have temporally dual effects that balance each other. Such dual IDEs are said to be *self-adjoint* and provide the basis for a theory that incorporates information detection as a dual to energy control whenever a perceptually controlled behavior successfully approaches an intended goal. This provides the formal means by which to treat action learning as a hereditary influence and perception as an anticipatory influence that together comprise reciprocal components of the perceiving–acting cycle (Shaw et al., 1990). The possible existence of an inner (scalar) product invariant between these two temporally dual processes provides evidence that perception, as information detection, and action, as resource control, might be lawfully related. This point is discussed later more fully.

Addition and multiplication of operators can be defined in a straightforward manner so that they constitute an operator algebra (Golden & Graham, 1988). Thus, stated algebraically in operator notation, the rather complicated IDE is approximated by y = k + Kx, where K is the multiplicative factor whose value changes as a function of the discrete number of pulls of the spring or the amount of practice (Figure 2). But, because the effects of strains or practice could, in principle, be continuous rather than discrete, it is necessary to integrate over the interval (t, s) in which the pulls or practice took place. Here, s is the next pull of the spring or learning opportunity and t is the last pull or learning opportunity. Thus, (t, s) is the interval comprising the distribution of strains or practice affecting the changing dispositional (e.g., stiffness) curve, as represented by the appropriately designated K under the definite integral. This complex dispositional variable changes in units of strain, x(s) ds (Appendix B). If the hereditary influence of K should be nonlinear rather than linear (Equation 1.1), then the integral transform term would have to be changed to reflect this fact. (The required alterations to the formula may be found in Shaw & Alley, 1985; Appendix B; or Tricomi, 1957.)

In summary, hereditary influence and hereditary laws provide an appropriate way to model action learning—construed as the causal effects of learning on the intentional control of action as a function of past experience (e.g., traditionally identified with reinforcement). Anticipatory Volterra equations provide an appropriate way to model perceptual learning—construed as a function of anticipations of future goal states (e.g., traditionally identified with expectancy). As noted, the functionals that express the law of anticipatory influences on learning are mathematical duals of the functionals that express the law of hereditary influences on learning. This duality relationship is manifested in the inversion of the order of terms specifying the interval over which Equation 1 is defined with respect to the interval over which Equation 2 is defined, as follows:

$$x(t) = ky(t) + \int_0^t K(s, t)\, y(s)\, ds. \tag{1.2}$$

That these equations define functionals rather than functions is clear from the fact that they provide mappings from functions to functions, that is, from $x(t)$ to $y(t)$ or vice-versa. Because Volterra equations provide a formulation for functionals, a single IDE can express an infinite set of ODEs (Kramer, 1970). (See Appendix B.)

The Role of Instruction in Learning

Our approach assumes that learning is a complex dynamical process by which two initially uncoordinated processes, information detection and action control, eventually become sufficiently coordinated to accomplish intended acts. Organisms may learn in their environmental contexts with or without the benefit of instruction: Instruction might originate from an instructor who provides information about the selection of goals and the means for obtaining them, including the avoidance of thwarts. Or, instruction might take the form of self-instruction without intervention or guidance from another party—the learner himself sets his own goals, procures his own means, and avoids thwarts, without outside assistance. The traditional view asserts that learning is governed by laws objectively applied to the learning situation. By contrast, this new view argues that the learner must learn the laws that govern the coordination of perception and action. Intention initializes the laws with respect to the particular learning situation and reinitializes them from trial to trial or over change in learning situation.

Instruction orginating inside or outside the learning system may be a source of either hereditary or anticipatory influences on the successful achievement of a goal. As pointed out, hereditary influences are past-pending; they manifest themselves as the cumulative effects of record-keeping or postattunement—postdictive changes in the initial conditions that act on the current (detection or control) state of the learner. Typically, but not exclusively, one identifies such hereditary influences with reward or reinforcement. The mechanism for hereditary influences is (negative) feedback.

By contrast, anticipatory influences are future-tending; they manifest themselves as the cumulative effects of expectancies or preattunement—predictive changes in informational sensitivities or control constraints—that act on the current state of the learner. The mechanism for anticipatory influences is feedforward—the design or reinitialization of a system or device as a function of its anticipated use. (See Newell, 1991; Shaw & Alley, 1985.)

Let (self-)instruction be interpreted as feedforward or feedback constraints on information or control. The feedforward information pertains to the setting of goals and the feedback control to the obtaining and executing of means to accomplish the goal (including the avoidance of thwarts). By what strategy do learners become successfully instructed to carry out goal-directed activities?

Circular Causality of Perception and Action

Perception, construed as the detection of *goal-specific information*, is the primary source of anticipatory influence on the control of *goal-directed actions.* Action, reciprocally construed as the causal control of coordinated behavior, is the primary hereditary influence on the detection of goal-specific information. In ecological psychology, anticipatory and hereditary influences are assumed to enter into a *circular causality* rather than a linear causal chain. This circular causality is called the *perceiving–acting cycle* and is postulated to be the smallest unit of analysis for psychological theories. Hence, in the final analysis, the dynamical interplay of perception and action provides a single complex object of scientific study rather than two individual processes that might be studied in isolation.

In mathematical control theory, the concepts of *observability* and *controllability,* respectively, make explicit some of the core intuitive content behind the ideas of anticipatory and hereditary influences on goal-directed action that are guided by detection of goal-specific information. The possibility of using the dual mathematical concepts of observability and controllability (Kalman, Engelar, & Bucy, 1962) will play a central role in our development of a theory of intentional learning, as discussed in Section 3.

The Role of Intention as an Operator in Learning

Of what adaptive value are aimless actions or undirected perceptions? A safe assumption is that adaptive value for a living system accrues from intended actions that are guided to a successful end by the directed detection of relevant information. To do so, the system must engage the perceiving–acting cycle under the coordinating influence of an intentional rule for the perceptual control of action (Gibson, 1979; Shaw, 1987; Shaw & Kinsella-Shaw, 1988; Shaw et al., 1990). This intentional rule specifies the lawful conditions under which observable goal-specific information can enter into a circular causality with controllable acts so as to satisfy an intended goal. Consequently, this is what successful learners learn. As abstractions, observability and controllability are mathematical duals (i.e., mutual and reciprocal) whenever their cyclic interplay successfully leads to goal satisfaction over a minimal path (Kalman et al., 1962).

The circular causality of the perceiving–acting cycle means that goal-specific information acts, first, as antecedent constraints on action consequences and, subsequently, as the consequences of antecedent action constraints. This reciprocal interplay continues iteratively to improve the system's goal-directedness and hence the specificity of the goal information. The interweaving of action control constraints (controllability) and information detection constraints (observability) repeats in a cyclic manner until the intended goal is reached or the attempt to do so is aborted.

Intention, as addressed here, is not merely a philosophical metaphor nor a phenomenological construct, but denotes a system operator. This *intentional operator* mathematically represents the act that selects goal parameters and, thereby, specifies the circular logic by which information and control constraints mutually reciprocate. These goal parameters constitute the relevant boundary conditions imposed by intention on the perceiving–acting cycle. Goals require specification in terms of kinematic and kinetic parameters. The former are called *target parameters* and the latter *manner parameters.* Together these characterize the actor's initial, ongoing, and final state with respect to the intended goal state. Target parameters denote the time, distance, and direction to contact the target; the manner parameters denote the impulse, work, and torque forces required to move the system to the target on time, over the requisite distance and direction, respectively. Target parameters comprise the specific information to be detected (the observability criterion) and manner parameters the specific control to be applied (the dual controllability criterion). Being a target connotes more than just being a physical object with an objective spatiotemporal location. Being a target connotes the selection of an intended final state (defined with respect to objects or target locations in space–time) from among the set of all possible final states (given the initial conditions and relevant natural laws). Hence, the target is more aptly defined as a kinematically specified family of goal paths in space–time.

Likewise, a manner of approaching a target connotes more than momenta values along an arbitrary goal path. Rather manner connotes, in addition, the selection of the kinetic mode of resource allocation at each choice point along a target trajectory. Hence, an intended manner selects from among the set of all possible goal paths an individual goal path that satisfies certain intended kinetic criteria. Taken together, target and manner parameters individuate (or finalize) a kinetically determined mode of approach to a kinematically specified target from all physically possible space–time trajectories (allowed under the prevailing initial conditions and natural laws). More briefly: An intention selects an informationally specified final condition, initializes, and reinitializes the governing control laws at each choice point on the way to realizing the designated final condition. Systems that can do this may be said to operate in an *intentional mode.* In general, systems must learn to do this.

Successful attainment of the goal means that the intention operator has become attuned through learning to direct the appropriate interplay of the system's perceiving and acting until the goal coparameters have been nullified. What does this mean when formulated in terms of ODEs? In this article, an ODE is considered to be in *homogeneous* form if it can be rewritten with a zero on one side of the equals sign and no terms on the other side that vary explicitly as a function of time only. For example, in the second-order case, an ODE can be written as $p(t)\ddot{y} + q(t)\dot{y} + r(t)y = 0$. (Please note: the sense. For instance, $\dot{y} = F(y/t)$ also is considered to be in homogeneous form in a different contest. For a comin homogeneous form in a different context. For a comparison of the two senses, see Boyce & DiPrima, 1986, p. 88 and p. 111.) Recall that a *boundary condition* is a requirement to be met by a solution to a set of differential equations on a specified set of values of the independent variables. In terms of an ODE characterization of the system, this means that all boundary conditions have become homogeneous. The details of this argument are discussed later. Because these goal-specific boundary conditions are solutions to the dynamical equations used to model the perceiving–acting cycle, the ODEs used take on homogeneous form if and only if the system reaches the goal selected by its intention (Shaw & Kinsella-Shaw, 1988; Shaw et al., 1990). Being homogeneous simply means that none of the boundary conditions for the ODE are extrinsically time-dependent; that is, they are nonautonomous.

The central concept of mathematical duality is discussed next.

A Word About Motivation: Duality in the Spirit of General Analysis

Typically, a function is considered to be a single-valued mapping between two sets of numbers; or, as in measurement, between a set of geometric properties (e.g., length and area) and numbers; or, as in topology, between two sets of objects. But this is not the most general notion of function. The American mathematician E. H. Moore and the French mathematician Maurice Fre'chet studied functions so general that they could define mappings between sets of abstract entities that need not be numbers, geometric properties, nor topological objects. From this study, called *general analysis* (Kramer 1970), Fre'chet (1925) concluded that the existence of analogies between fundamental features of various theories implies the existence of a general abstract theory that includes the particular theories and unifies them with respect to those fundamental features.

In this spirit, consider a *duality*, a function so general that, under the framework of ecological psychology, it putatively reveals a fundamental analogy between organisms and their environments, affordances and effectivities (discussed later), perception and action, and detection of information and the control of action (Shaw & Turvey, 1981; Shaw, Turvey, & Mace, 1982; Turvey & Shaw, 1979). Because our motivation is in the spirit of general analysis, what is offered is a formal analogy that we believe holds over all learning situations rather than a model for learning as such. Any learning model, however, might profit by incorporating and exploiting these abstract properties.

A mathematical duality D is a symmetrical mapping that establishes a special (nontransitive) correspondence between one structure, **X** (e.g., a series of information detection states), and another structure, **Y** (e.g., a series of action control states), so that for any function f that maps from **X** to **Y**, there exists another function g that maps from **Y** to **X**. Furthermore, a duality between the structures is inherently nontransitive, for if there exists another function that putatively carries the image **Y** into another structure **Z**, so that $\mathbf{X} \rightarrow \mathbf{Y}$, $\mathbf{Y} \rightarrow \mathbf{Z}$, then **Z** must be isomorphic with **X**.

When D is equal to its own inverse, it is said to establish a *double* dual between **X** and **Y**. The ecological approach espoused here postulates a doubly dual relationship between the values of **X**, taken as informationally specified environmental properties, and corresponding values of **Y**, taken as organism-determined action control states. We call the operation $f: \mathbf{X} \rightarrow \mathbf{Y}$ *perception*, and the perceived environmental properties that support goal-directed activities *affordances*. Likewise, we call the inverse operation $g: \mathbf{Y} \rightarrow \mathbf{X}$ *action*, and the organism's capabilities for realizing such affordance supported actions *effectivities*. The double duality, $D: \mathbf{X} \rightarrow \mathbf{Y}$, composed from the operations f and $g = f^{-1}$, designates a system of affordance–effectivity constraints specific to an organism–environment system, or *ecosystem*, in which the organism, construed functionally as an effectivity system, is essentially successfully adapted to its environment, construed functionally as an affordance structure, or *econiche*. The abstract duality function therefore reveals that affordances (the values of perceptual functions) and effectivities (the values of action functions) are formally analogous under the commutativity property of this symmetric mapping. The notion of an intention operator formally exploits this symmetric mapping. (The details of this duality function are spelled out in Shaw et al., 1990.)

Thus, one sees that as duals, perception and action are inverse mappings of each other. The former takes values from the environment and maps them into the organism variables, whereas the latter does the inverse mapping. To be successful in performing an act, the environment must afford the informational and causal support for the act. In addition, the organism needs to possess the specific effectivity for carrying out the act. Intention must couple the approximate effectivity controls with the affordance information and sustain their coordination. For an actor to grasp an object requires both that the "graspability" of the object be an affordance property for that organism, and also that the actor has a "grasper" effectivity whose parameters are mutually compatible (dual) with the parameters of the object to be grasped. Whenever the organism is successfully guided by perception through a series of felicitious regulatory acts (e.g., muscular adjustments) that achieve the intended goal (e.g., grasping an object), the two functions necessarily become doubly dualized. The double dualization of these functions can be progressive rather than immediate, however. As described earlier, under the appropriate intentional selection of a goal, perception and action act together mutually, and on each other reciprocally, as a closed loop. In this way, learning is the appropriate intention to learn the lawful operation that progressively tightens this closed loop over each cycle until the goal is reached. An index of the success of this learning is the degree to which there exists, over the course of task performance, an inner product invariant between information and control. (This will be shown later.)

To illustrate the role of the perception–action duality and to make clear the semantic distinction between ODE-based models and IDE-based models with respect to their differing psychological import, it will be useful to develop a concrete (if hypothetical) example.

The Intentional Spring Task

Overview. The task requires the subject to learn to imitate the way an instructor stretches a given spring. The spring has two identical handles, so that learner and instructor can hold onto it simultaneously (Figure 3). It is also fixed at both ends, with the handles attached to its center. A vertical partition separates the instructor and the learner. The learner is blindfolded and wears earplugs, so that all learning must be haptically rather than visually specified (i.e., by feeling how the spring is stretched). A slot in the partition guides the hands so that the spring has but one unrestricted spatial (horizontal) degree of freedom.

In general, the kinematic degrees of freedom allow the instructor to choose an infinite variety of forcing functions to move the spring. The spring can be moved

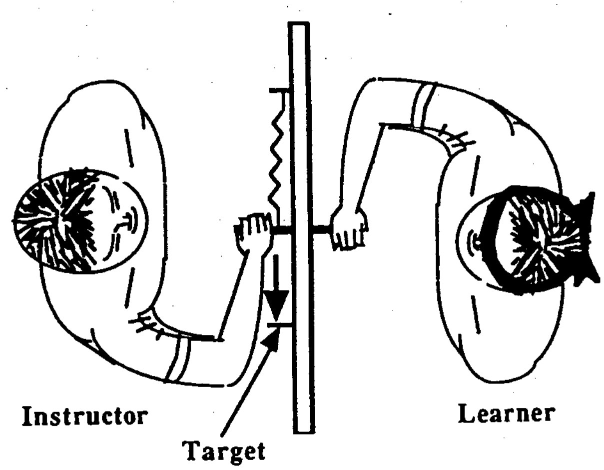

FIGURE 3. Top view of the social tracking task.

at any speed for any number of times in either the leftward or rightward direction, but, for the sake of illustration, in this task, movement is assumed to be a simple smooth stretching to a target point and releasing. The forcing function applied to the spring by either learner or instructor conveys a pair of dual perspective-dependent functions: an *L-informing function,* defined as what the learner feels as the spring is moved by the instructor. Namely, the learner detects an *I-controlling function,* defined here as the force applied by the instructor but in a reversed direction (assuming both participants, who face each other across the partition, grip the spring with the same hands). Conversely, when the learner attempts to move the spring, then an I-informing function is defined dually for the instructor by the L-controlling function. It is important to note, therefore, that this task is defined over a dual pair of dual functions—that is,

> (I-controlling functions that are dual to L-informing functions) are dual to (L-controlling functions that are dual to I-informing functions).

This system of dual functions characterizes, in the spirit of Moore and Fre'chet's general analysis, a fundamental analogy holding between experiment and theory, as exemplified under the ecological framework: The concrete scheme for the experimental paradigm, called *the social tracking paradigm* (Smith & Smith, 1987), is formally analogous to the abstract scheme for mathematically describing instructing and learning functions, called the *self-adjoint control paradigm* (Repperger & Shaw, 1989; Shaw & Repperger, 1989a, 1989b).

The task may be thought of as a perturbation-reduction task. Here, successful learning is achieved where the learner feels no perturbations introduced by the instructor as corrective feedback. Or, dually, successful instruction is achieved when the learner feels no perturbation introduced by the learner—as defined against the instructor's expectation. The task is over when the learner's attempts require no corrective (felt) feedback from the instructor for three successful repetitions. At such time, one may conclude that the subject has succeeded in absorbing the intentional operator that has been successfully passed by the instructor through the shared forcing function on the spring. Ideally, a final state of perfect learning is achieved if the learner's informing function eventually becomes dual to the instructor's controlling function so that the learner's controlling function is dual to that informing function. In this way, the instructor's task intention is passed to the learner through the double dual described above.

Instructions A: Task goal. The learner is apprised of how the apparatus is set up, blindfolded, ear-plugged, and given the task instructions. For convenience, we have broken the instructions into sections that correspond to the mathematical methods to be discussed later on. The subject is told that his/her task is to learn to anticipate the way an instructor pulls the spring so that he/she might eventually come to imitate that performance autonomously.

The learner is told to begin the learning process by holding lightly and passively to the spring handle. As the instructor pulls the spring, the learner's arm is pulled through the motion that is to be learned. Consequently, the subject is told to pay close attention to the manner and the target of the instructor's pulling. Specifically, this means paying attention to the frequency of oscillation and the target length to which the spring is brought to rest by the instructor at the end of each pull.

Instructions B: Learning goal parameters. As the task goal becomes clearer, the learner is to gradually assume control of the spring, and, conversely, the instructor is to gradually relinquish control. After doing so, however, the instructor is to maintain a light grasp on the spring to monitor the learner's progress and to correct departures (perturbations) by the learner from the criterion performance pattern (that is, from the task pattern the instructor learned prior to the task).

As soon as a perturbation is detected, the instructor is to resume control. Consequently, if the learner should feel any perturbations while pulling, control of the spring is again to be relinquished to the instructor. This assuming and relinquishing of control alternately by the learner and the instructor continues until perturbations are eliminated or reduced to some low criterion level.

Examples of this kind of active instruction are numerous, for instance, learning from an instructor to dance or to swing a golf club. In such cases, the instructor initially leads, with the learner attempting to follow. The learner is then allowed to lead while the instructor monitors his/her attempt to anticipate the criterion sequence of movements. The instructor intervenes to recapture the lead only when there is a need to compensate for the perturbations. If the learner gets hopelessly lost in major perturbations, however, then the instructor may intervene to reinitialize the task (i.e., to abort the current trial in favor of starting a new trial).

Instructions C: Achieving adaptive autonomy. The learner's task is to learn to match the instructor's manner of pulling within the prescribed target range and to terminate pulling at the same place and in the same manner that the instructor does. This task is successfully completed when the instructor's intention is assimilated by the learner. At such time, the learner will feel no counterforces (perturbations) from the monitoring instructor.

Trials. Trials are self-paced. A single trial is defined as the period from the moment the learner begins monitoring the spring to the moment the learner attempts to bring the spring to rest at the target. Clearly, from the perspective of the learner, each trial has a monitoring (informing) phase and a test (controlling) phase and a evaluating (informing) phase. The informing and controlling phases are 180° out of phase across the participants—an expression of the duality of their roles in the experiment.

Feedback and feedforward information. Two kinds of task-specific information are given and received by each participant: *Directive* feedback information is given by the active participant to the passive participant through the controlling function, and *corrective* feedback is felt by the passive participant through the informing function. Negative feedback from the instructor to the learner is detected through felt corrective counterforces; no felt counterforce is positive feedback that the trial is going well and that the learner should continue the performance. Performance (feedback) information passed from learner to instructor is corrective on instruction, whereas instructional (feedforward) information from the instructor to learner is directive on performance.

When the learner is attempting to practice his/her control of the spring, both the learner and the instructor are told to pay close attention to the corrective feedback provided by their respective but dual informing functions. Corrective feedback is operationally defined for the learner as the felt mismatch (perturbation) of the L-controlling function with the (I-produced) L-informing function, whereas corrective feedback is dually defined for the instructor as the felt mismatch (perturbation) of the very lightly applied I-controlling function, which shadows the (L-produced) I-informing function.

Clearly, corrective feedback for the learner during his/her attempts to practice control of the spring involves the stiffness of the spring plus the corrective impulse forces applied by the instructor. The corrective impulse force will be felt by the learner (L-informing function) as a variation in the *compliance* of a virtual spring (i.e., $1/k$, where k is the stiffness coefficient). (The stiffness of the virtual spring is determined by both the metal spring and the muscular spring of the instructor's arm.) When the learner has learned to imitate the instructor perfectly, the spring will feel maximally compliant (so that the L-informing function is nullified). Hence, a measure of the felt corrective feedback (by either learner or instructor) is identical to the measure of felt variations in the virtual spring's stiffness. Minimization of corrective feedback information from instructor to learner maximizes the directive feedforward information from learner to instructor.

From the perspective of the learner, the absolute value of the L-informing function is the independent task variable, and the degree of compliance of the L-controlling function is the dependent task variable. But one must take care in locating these variables in cases of circular causation (feedback) such as this social tracking task. The roles of independent and dependent variables are only relative and interchange when the perspective is shifted from learner to instructor because each acts as an environmental source of mutual and reciprocal constraint on the other.

Figure 4 depicts the stages of learning in the intentional spring task. The intricate interaction of information and control variables can be formally codified and somewhat simplified by the use of the algebra of operators. The operators and their relationships will be ex-

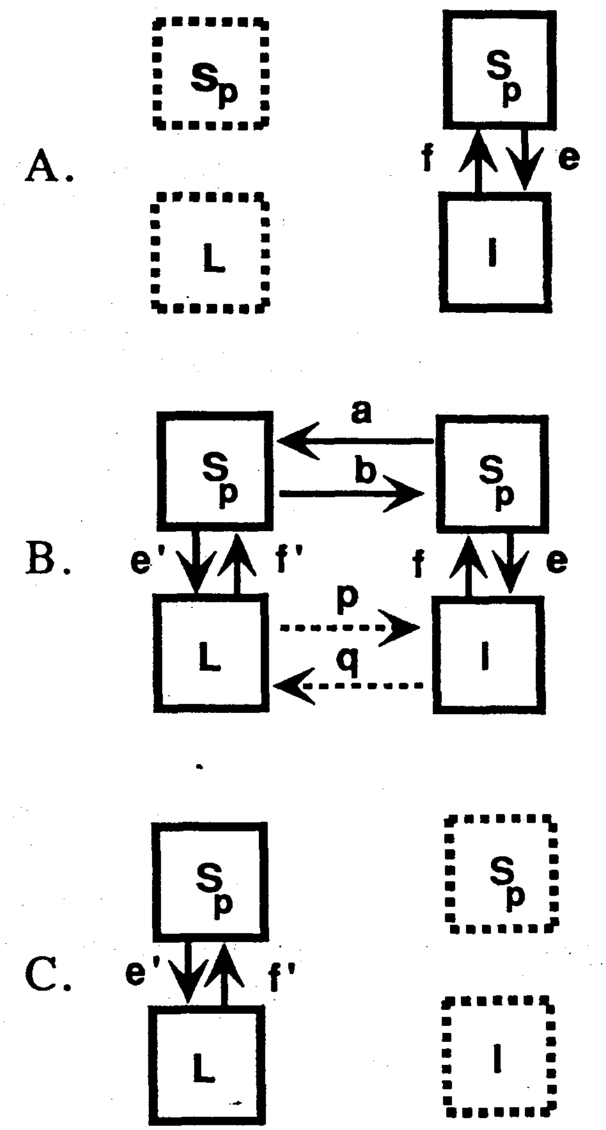

FIGURE 4. The three states of the learning process. (A) shows the initial stage before learner L is coupled through the forcing function of the spring (Sp) to the instructor I; (B) shows a compact schema of the learning process by which the observability and controllability criteria may be satisfied by the passing of a self-adjoint operator from I to L via the informing and controlling functions, respectively; and (C) shows the final stage of learning, in which adaptive autonomy is achieved by L after absorbing I's task intention (boundary conditions) via the passed self-adjoint operator (see text for details).

plicated formally in Section 2 and explicated in even greater detail in Section 3.

Formal Interpretation of the Intentional Spring Task

The following outline, showing the dependency of concepts, will help make explicit the semantic correspondence between the formal and psychological properties of the intentional spring task. These properties apply to the coupled learner-spring-instructor system but not to the decoupled instructor-spring or learner-spring systems.

1. *Boundary conditions* are extrinsically imposed requirements on the initial and final states of the learner-spring-instructor system (e.g., given the task set-up, task goal parameters, and the learner's intention to minimize learning errors and the instructor's intention to maximize instruction effectiveness).

A. *Initial conditions* are requirements imposed on the initial state of a system (e.g., experimental set-up in general; spring's stiffness coefficient; learner's ability and intention to learn; instructor's ability and intention to instruct).

B. *Final conditions* are requirements imposed on the final state of a system (e.g., spring's state of stress and strain at the end of task, instructor's goal to be achieved, learner's goal to be achieved).

2. A *forcing function* has two components: extrinsically imposed hereditary and anticipatory influences on a system. Figure 4B is a compact scheme of operator relationships that in the actual learning situation would be distributed over many iterations of the perceiving-acting cycle.

A. A *controlling function* (on L by I) is an extrinsic source of force that determines hereditary influences on the system. Forces applied to the learner by the instructor via the spring are extrinsic to learner; conversely, forces applied to instructor by learner via the spring are extrinsic to instructor. Both instructor- and learner-applied forces are extrinsic to spring. A controlling function is represented in Figure 4B as $f \rightarrow a \rightarrow e'$.

B. An *informing function* (on L by I) is an extrinsic source of information that determines anticipatory influences on the system. Reactive forces applied to spring by instructor to compensate for learner's errors serve as feedback to learner, where null reactive forces specify no learner error; conversely, active forces applied to spring by learner function as feedback to instructor regarding learner error with respect to task goal. An informing function is represented in Figure 4B as the feedback loop I: $f \rightarrow a \rightarrow e'$ and feedback loop L: $f' \rightarrow b \rightarrow e$.

3. A *Green's functional* is a functional representation of the solution to a boundary-value problem consisting of observability and controllability functionals as double duals. This operator formally expresses the intuitive content of the intentional operator term introduced earlier. The Green's functional applies to the uncoupled instructor-spring or the learner-spring systems, or the coupled instructor-spring-learner system. (See, for example, Equations 1 and 2).

A. The term *observability functional* pertains to the anticipatory influence of information about final states of a system on its current detection and control states. (See Figure 4B: $f \rightarrow a \rightarrow e'$ as feedback after $f' \rightarrow b \rightarrow e$ commutes with p.)

B. The term *controllability functional* pertains to the hereditary (causal) influence of initial states of a system on its current detection and control states (See Figure 4B: $f' \rightarrow b \rightarrow e$ as feedforward after $f \rightarrow a \rightarrow e'$ commutes with q.)

Learning as the Solving of a Boundary-Value Problem

A boundary condition is an extrinsically imposed requirement on the initial and/or final state of a system. A solution to a set of differential equations representing the dynamical system of interest must satisfy these requirements with respect to a specified set of values of the independent variables. A boundary- (initial and/or final) value problem involves finding the solution to the system's equations that meets the specified requirements on the relevant independent variables. Such requirements may depend on either physical, biological, or psychological factors.

In the current example, a learner must solve a boundary- (final) value problem posed by instructor in order to learn the task goal. This is done by absorbing the instructor's intention, which somehow is passed through the controlling function (a component of the forcing function). On the other hand, the instructor must solve a dual boundary (initial) problem regarding what instructions to pass to the learner through the informing function (also a component of the forcing function). Appropriate instructions must reinitialize the learner so that he/she can perform the task successfully and achieve the intended final condition. A natural question to ask is: What are the sufficient conditions that enable the final and initial boundary-value problems to be solved by the learner-spring-instructor system? This condition expresses the intuitive content of the phrase "the intentional dynamics of the task."

The sufficient condition that the learning process should be successful is that the controlling and informing functions allow the learner to satisfy the same controllability and observability criteria as the instructor. This is tantamount to the learner-spring system's achieving as a final condition a Green's functional equivalent to that which the instructor-spring system possessed as an initial condition.

1. The *controllability criterion* is a property of a system that satisfies certain hereditary constraints on certain of its independent variables: Given an appropriate

intitial state and any future state (e.g., the final state), there exists a time-forward path integral that sums the hereditary effects of a controlling function from the initial state to any future state within the time interval.

2. The *observability criterion* is a property of a system that satisfies certain anticipatory constraints on certain of its independent variables: Given an intended final state and any previous state (e.g., the current state), there exists a time-backward path integral that sums the anticipatory effects of an informing function from the intended final state to any previous state within the time interval.

Let Equation 1 represent the Volterra functional satisfying the controllability criterion and Equation 2 the Volterra functional satisfying the observability criterion. The appropriate Green's functional can be represented by an identity satisfied by the hereditary and anticipatory functionals of these dual IDEs; namely,

$$K(t, s) - K(s, t) = K - K^* = 0. \quad (1.3)$$

These are said to be doubly dual, or self-adjoint, kernels of the IDEs depicted in Equations 1 and 2, if $K = K^*$; otherwise, they are singly dual, or adjoint.

The adjointness property will be used extensively to express the duality (i.e., the observed symmetry or antisymmetry) between equations of dual systems. For instance, when the learner becomes as competent with respect to a task skill as the instructor, then the learner–spring equations will be said to have become self-adjoint with respect to the instructor–spring system. Although the form that self-adjointness (duality) takes obviously must differ notationally between ODE and IDE formulations of a system, the property is necessarily inherited when converting from one to the other. Consequently, we shall assume this fact without showing it and refer the interested reader to standard references demonstrating it to be so (e.g., Courant & Hilbert, 1953, pp. 277–280).

Thus, to reiterate our central concern: One needs to address the issue of how an extrinsic intention to carry out a specific task, as imposed on a learner by an instructor, can come to act as an intrinsic (intentional) constraint on the learner's perceiving–acting cycle?

For organisms to become sensitive to the useful dimensions of goal-specific information, there needs to be an "education of attention" as Gibson (1966, 1979) has suggested. Likewise, on the side of skill acquisition and the control of action, there needs to be a corollary "education of intention" to select and execute those voluntary subacts in the manner required to satisfy a more global task goal.

Consequently, any adequate theory of learning must address three types of learning: perceptual learning, action learning, and intentional learning. Intentional learning takes place at a higher level of abstraction than either perceptual or action learning. We believe successful intentional learning serves to bind the other two together into a duality and to modulate their coupling to preserve that duality via the coming into play of Green's operator.

Under the proposed modeling strategy, we attempt to show how an intention imposed by extrinsic rule for the perceptual control of action can become an intrinsic law for doing so. What is usually missing from other accounts is an explicit formulation of how information and energy must be lawfully coupled into a perceiving–acting cycle if the system's rule-governed dynamics is to serve invariantly the stipulated intention: Mathematically, in describing the learner's progress in assimilating the intention of the instructor this requires a move from differential equations with extrinsically imposed boundary conditions to integro-differential equations with boundary conditions. This move from ODEs to IDEs is analogous to moving from an extrinsically imposed rule for changing the elastic coefficient in the linear version of Hooke's law to a Volterra functional that makes the same changes intrinsically in the generalized version of Hooke's law. This move is made formally explicit in Section 2 and Appendix A.

A second question at the heart of our modeling strategy needs to be addressed: If ODEs (e.g., springs and coupled oscillators) are to be used to model dynamical action systems, how do they relate to the ODEs used to model the dynamical perceptual system? With these issues in mind, an overview of the mathematical strategy to be used is provided next.

Overview and Summary of the Mathematical Strategy for Modeling Intentional Learning

Let us attempt to trace the abstract conceptual thread that runs throughout the last two parts of this article. Three related mathematical concepts used extensively throughout this paper are duality, self-adjointness, and Green's function. The last two concepts are intimately related to the first; indeed, they are special cases. The most fundamental duality of interest is that between perceiving and acting, construed formally as the duality between observability and controllability (Kalman et al., 1962). This theorem will be exploited fully to provide a first pass on a mathematical formulation of a duality-based learning theory in Section 3.

Another related expression of duality is that which holds between hereditary and anticipatory influences imposed on a system from the outside. These extrinsic reciprocal influences are expressed in two ways: through forcing functions on (nonhomogeneous) ODEs representing a dynamical system, on the one hand, and through extrinsic requirements, or boundary conditions, placed on certain independent variables, on the other hand. Influences that require extrinsic means for expression when applied to ODE representations of a dynamical system can be expressed intrinsically as IDEs that are formally equivalent to the ODEs. Equations 1 and 2 provide an example of how extrinsic influence functions can be instantiated intrinsically as dual influence functionals.

Psychologically speaking, much can be made of the semantic difference between the formal means that treat hereditary and anticipatory influences extrinsically (ODEs) and those that treat them intrinsically (IDEs). The move from ODEs to IDEs, implicit in the case of Equations 1 and 2, needs to be made explicit as a key part of the proposed mathematical strategy. Recall that our primary objective is to develop a modeling strategy rather than to present a model per se. In accordance with this objective, let us consider how forces acting on a system from the outside can convey information specific to an intention so that the system is instructed about a task goal and progressively learns that intention. These issues are formulated in explicit form using operators in Section 2.

From the above, it is clear that an explicit method is needed to semantically represent the eventual assimilation of an intention into the autonomy of the system that is first imposed on it through instruction from the outside. This involves two stages in modeling, each requiring identifiable mathematical techniques: the absorbing of the control-specific goal information from an extrinsically imposed forcing function and the absorbing of intentional requirements from extrinsically imposed boundary conditions. The Hilbert-Schmidt's technique and the Green's function techniques are used to make explicit the absorption of forcing functions, characteristic of nonhomogeneous ODEs, into the homogeneous ODE representation of the learner–spring system. Green's function technique is used for solving boundary-value problems for ODEs of the Sturm-Liou-

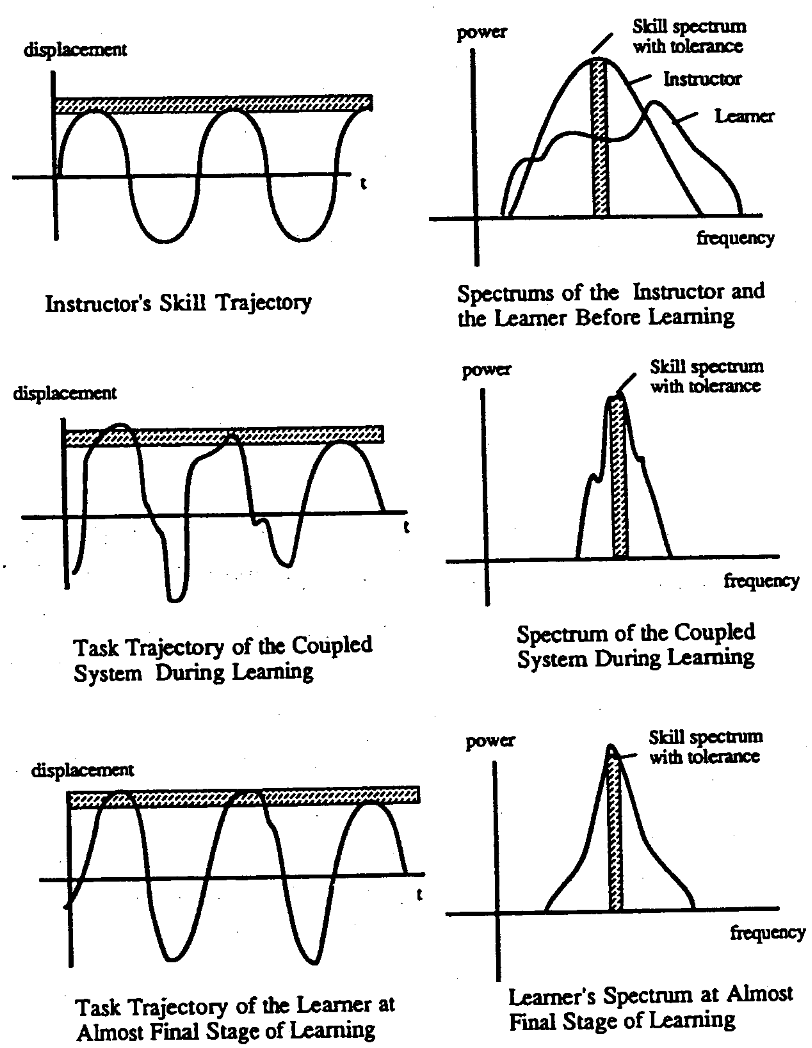

FIGURE 5. Spectrum transfer from the instructor to the learner.

ville type—the type of equations that we postulate for learning. This technique accomplishes the absorption of the intentionally specific boundary conditions into the dual IDE functional representation of hereditary and anticipatory influences.

The most elegant way to describe these absorption techniques is to use adjoint (singly dual) operators and self-adjoint (doubly dual) differential and integral operators. (For a general discussion of operators, see Synge, 1970; for their use in psychology, see, for instance, Solomon, 1988; Solomon, Turvey, & Burton, 1989.) By representing ODEs or IDEs in operator notation, the more complicated techniques of classical calculus and functional analysis can be presented through an algebra of compact linear operators on a Hilbert space (Lanczos, 1961).

Section 2. Absorbing Extrinsic Influences Into the Intrinsic Autonomy of a System

Instructor Skill as a Self-Adjoint Operator

Self-adjointness is a very valuable property of certain linear differential equations (see earlier discussion of Equations 1 and 2). Solutions to the boundary-value problem for a system can often be aided by first finding the adjoint equation of the system and coupling them to form a self-adjoint equation (i.e., Green's function). Formally, it is well known that a linear system with an arbitrary right side (homogeneous or nonhomgeneous) is solvable if and only if its adjoint homogeneous counterpart has no solution that is not identically zero, that is, if it has only the trivial solution. If the system is self-adjoint, then, by definition, its adjoint system is also. For the system $\mathbf{D}u(t) = g(t)$, then, $\mathbf{D}$ denotes the differential operator, whereas the derivatives of a function are denoted in standard operator notation, for example, as $\dot{y} = dy/dt$. This system, as represented, should not have more than one solution given any fixed boundary condition in that case. The system is then said to be tight or complete. The formal property of self-adjointness seems to be a natural condition for most equations used to represent dynamical systems. Indeed, this intuitive insight can be made stronger and more specific (Lanczos, 1961):

> The majority of the differential equations encountered in mathematical physics belong to the self-adjoint variety. The reason is that all the equations of mathematical physics which do not involve any energy losses are deducible from a "principle of least action," that is the principle of making a certain scalar quantity a minimum or maximum. All the linear differential equations which are deducible from minimizing or maximizing a certain quantity, are automatically self-adjoint and vice-versa: *all differential equations which are self-adjoint, are deducible from a minimum-maximum principle* (p. 226).

Accepting this fact, it follows that Nature is a self-adjoint instructor for the animals living in it. In other words, animals have to appreciate the self-adjointness of the environment in which they live and move. As part of nature, animals and humans either evolve or learn to become operators with many self-adjoint operational capabilities. Similarly, if someone has learned a specific skill, it must be a self-adjoint, or a composition of more than one self-adjoint, operational capability. (Please note: In the quote above, the condition cited does not require that all self-adjoint differential equations be deducible from a principle of least action but only that all differential equations that are so deduced will necessarily be self-adjoint.)

For the sake of simplicity, we assume that in our illustrative example of a learning process (the intentional spring task) the specific task to be learned is construed as a self-adjoint operator. In learning the intention of the instructor, the instructor's task is to inform the learner of the self-adjointness required and the learner's task is to absorb that goal-parameter information as boundary conditions on control variables. In this way, figuratively speaking, the instructor must pass an adjoint operator to the learner, who must then (over trials) absorb it (described vis á vis the Hilbert-Schmidt and Sturm-Liouville techniques). Thus, the problem to be solved is: How can the instructor transfer its self-adjoint operational skill to the learner? (The minimal requirement for the learner will be specified in the following section.)

Before solving this problem in a generic fashion, it is necessary to give a formal description of the instructor's skill. According to the well-known Hilbert-Schmidt theorem, any self-adjoint and compact operator $\mathbf{A}$ (compactness is a usual assumption), like any differential operator, can be written in the following form:

$$\mathbf{A}x = \sum_{n=1}^{\infty} \lambda_n (x, v_n) v_n, \tag{2.1}$$

where λ_n and v_n are the eigenvalues and corresponding eigenvectors of $\mathbf{A}$. The eigenvalues are called the *spectrum* of $\mathbf{A}$. The decomposition of $\mathbf{A}$, given above, is called *the spectral decomposition.* One may describe a generic skill as an $\mathbf{A}$ operator with a given spectrum (Figure 5). Intuitively, it follows that learning is a transfer of the instructor's skill spectrum (v_n) and its power (λ_n).

Furthermore, one may assume that in most cases the instructor's spectrum must be wider than the skill spectrum itself, especially if one assumes that the learning is mostly due to direct transfer. Though our intuitive spring example does not belong to the class of complex problems, it is sophisticated enough to illustrate the complexity of the problem of learning. Considering the spectrum to be transferred by the instructor, this task seems simple, because some fundamental frequency (eigenvalue) should be passed to the learner. (In Figure

5, this is represented by the peak value of power.) Note, however, that during the learning process, because the learner and the instructor are coupled and may perturb each other, each will use a wider spectrum than that strictly dictated by the task.

First, the learning process is characterized by pairs of compact linear differential operators on a complete inner product space, or Hilbert space. Initially, the instructor–spring system is represented by a self-adjoint operator **I**, expressing a mastery of the relevant skill in pulling the spring. By contrast, the learner–spring system is represented only by an adjoint operator **L**, expressing a minimal skill capability that also includes an intention to learn (i.e., to follow instructions) (Figure 4A). Linear operators over a Hilbert space constitute a vector space with a inner (scalar) product (denoted by [*, *]). Using this inner product operator, one obtains [**I**, **L**] = c, where (as shown in Section 3) c is an invariant number throughout learning so long as the coupled **I–L** system remains on a solution path to successful learning. As remarked earlier, if learning is the lawful coordination of information detection and action control, as defined relative to an intentionally selected goal, then this inner product invariant can be taken as a requirement of its success.

Initialization of Learning as Setting Up a System of Adjoint Equations

In setting up the complete schema for intentional learning, one must specify the conditions and the aim of learning. Initialization of the learning process of the given skilled action requires two conditions: First, the instructor must constrain its self-adjoint **I** operator if it is to be transferred to the learner. At the same time, both the learner and the instructor have to set up their boundary values in an adjoint fashion (Figure 4B). Though, for mathematical convenience, one may choose to separate the differential operators and the boundary conditions, the processes to which they refer in nature do not allow such artificial separation. Indeed one is admonished that ". . . in actual fact the boundary conditions are inseparable ingredients of the adjoint operator without which it loses its significance" (Lanczos, 1961, p. 184).

The learner's operational capability should be modified by instruction. For instance, instructions might be described verbally as the means by which the instructor specifies to the learner the target and manner parameters as well as other aspects of the task to which special attention should be paid. Thus the learner upgrades its operator from **L** to **L′**. As stated earlier, however, the new operators **I′** and **L′** must still conserve the inner product [**I′**, **L′**] = [**I**, **L**] = c, or else the learner is being misinstructed as to the correct nature of the task (see Section 3 for details). This means that as the learner upgrades from **L** to **L′**, the instructor downgrades from **I** to **I′**. But this does not mean that the instructor gradually loses his/her skill, only that it is called on less and less in the training of the learner.

By invoking the Hilbert-Schmidt decomposition theorem, this inverse change in the **I** and **L** operators can be described as a (Hermitian) change of their dual spectum—on the one hand, the need to release less and less information by the instructor and the need to absorb more and more control by the learner, and, on the other hand, the need to release more and more control by the instructor and the need to absorb less and less information by the learner. This is the reason that an inner product invariant exists. (Recall the discussion in Section 1 of the controlling function and the informing function as components of the forcing function that are particularly meaningful in describing the learning process.) Thus the instructor decreases its spectral production (i.e., the power of its spectrum) in the learning process but does not lose the ability to produce it. This necessity for the self-adjointness requirement for **I′** and **L′** follows directly from the arguments already presented.

Learning as an Operator Transfer Process

With respect to the self-adjointness requirement, the same arguments hold during the course of learning just as they do for initializing the learning process. Furthermore, as argued, the inner product of the operators **I′** and **L′** remains invariant during the entire process unless the goal parameters of the task change. (Merely taking a new path to the target does not necessarily count as a violation so long as manner of approaching the target is not violated.) At this point, one may use the mathematical fact that ". . . every differential equation becomes self-adjoint if we complement it by the adjoint equation" (Lanczos, 1961, p. 243). Formally, if **D**u = f has **D′**v = g as its adjoint system, then the coupled system

$$(\mathbf{D}, \mathbf{D}')\begin{pmatrix} u \\ v \end{pmatrix} = (f, g) \tag{2.2}$$

will be self-adjoint.

The Need for a Theory of Evolving Self-Adjointness

Practically speaking, the adjoint instructor–spring–learner (**I′**, **L′**) system, as a coupled system, can be considered self-adjoint. During the dynamic learning process, the participants try to preserve the self-adjointness of their coupled system by gradually modifying the interplay of their operators. Green's function (see Appendix A) seems to be the best candidate for describing the criterion conditions for this complex process. The dynamic transfer of the instructor's operational capability can be described with the help of the Sturm-Liouville theory, which uses Green's function. It seems to us, however, that in requiring self-adjointness of the two coupled systems during the learning process, this theory requires too much. Instead, one needs a theory that initially requires only coupled adjoint systems, with self-adjointness gradually achieved as a final condition of learning over the course of the reciprocal interaction of their spectra.

Consequently, the Sturm-Liouville theory is used only to describe the sufficient condition for successful instantaneous learning and awaits the discovery of a more dynamical description for time-varying systems capable of progressive learning. (Details of the operator transfer process can be found in Appendix A.) As a result of the learning process, in the optimal case, the learner achieves the same operator, $\mathbf{L}' = \mathbf{I}$, that the instructor had at the beginning of the process. In other words, the whole spectrum of $\mathbf{I}$ will be tranferred to $\mathbf{L}'$ (Figure 4C). Let us emphasize that in this generic operator description, the learning process is not analyzed into perceptual learning and action learning. Rather, it is treated as transfer and absorption of self-adjointness through a composite of intentional, perceptual, and action learning.

Intentional learning coordinates the dual processes of perceptual and action learning. To better understand the role of intention as described in operator language, one wants a more complete description of the equations to be coordinated and from which the operators are to be derived. Consequently, in the final section, we present the more complete descriptions and explore some of the most significant consequences of the adjoint systems approach to psychology.

Section 3. Adjointness of Perceptual and Action Learning

Learning as Modeled by a System of Orthogonally Adjoint Equations

Let us begin with a brief summary of the Shaw and Alley paper (1985) that treats learning as a Volterra functional governed by the system of four dual IDEs, depicted in Table 1.

Learning is assumed to take two forms: Action learning tunes the learner toward minimum energy expenditures in orienting its sensory systems and controlling the motor system along an optimal path through task demands to the task goal. Perceptual learning tunes the learner toward maximum information processing needed to orient and steer the motor systems along the optimal path. The optimization of both energy and information processing by the learner in task-specific ways requires modeling by functional equations that are temporally dual (Table 1 viewed row-wise). These equations must also satisfy the controllability (top row) and observability (bottom row) criteria (Kalman et al., 1962).

Learning can be either hereditary or anticipatory. Action learning is an accumulation of experiences over time, which alters the state of readiness (initial condition) of the action/perception systems. Action learning tunes optimal energy expenditures to subsequent recurring task demands. Thus, when successful, action learning satisfies the controllability criterion (Kalman et al., 1962). Although such past-pending control processes

TABLE 1
Table of Duals

Perspective duals	
Organism perspective (Action)	Environment perspective (Perception)
Energy control (+t)	
$y(t) = kx(t) + \int_0^t K_O(t, s)\, x(s)\, ds$	$x(t) = ky(t) + \int_0^t K_E(t, s)\, y(s)\, ds$
Equation 1	Equation 2
Information (observation) (−t)	
$x(t) = ky(t) + \int_0^t K_O^*(s, t)\, y(s)\, ds$	$y(t) = kx(t) + \int_0^t K_E^*(s, t)\, x(s)\, ds$
Equation 3	Equation 4

Note. The symbols in these equations are explained in the text in connection with Equation 3. The transforms whose kernels take the form $\mathbf{K}(t, s)$ designate past-pending integrals that sum energy interactions from some past state (s) to some future state (t). By constrast, those with kernels of the form $\mathbf{K}^*(s, t)$ designate future-tending integrals that sum information transactions from some potential future goal state (t) backward in time to some current state (s).

are goal "blind," they are causally efficacious—being force driven rather than information driven.

By contrast, perceptual learning is an accumulation of expectancies over time, which alters the state of sensitivity of the action/perception systems. Perceptual learning tunes optimal information detection to the relevant task parameters (manner and target parameters) needed to steer a careful course to the task goal. Thus, when successful, perceptual learning satisfies the observability criterion (Kalman et al., 1962). Such future-tending processes have goal "acuity"; they are intentionally efficacious—being information driven rather than force driven.

Compare Figure 6 with Figure 4. Figure 6 shows the interplay of the operators associated with the dual (adjoint) IDEs depicted in Table 1. Here, $\mathbf{K}_O$ can be interpreted as the operator **L** and $\mathbf{K}_E$ as the operator **I**. The reciprocal relationships exhibited by operators **L** and **I** correspond to the observability and controllability functionals that must be satisfied for perceptual and action learning to succeed in satisfying the task intention.

In addition, the learning problem can be viewed from two perspectives: from the internal perspective of the learner as a system of degrees of freedom (state-variables) to be controlled and accommodated to a context of potentially nonlinear, time-varying constraints—the demands of the task environment—or from the external perspective of the task environment as a dual system of such constraints (dual state-variables) that define the task for the learner. Thus, the functional equations for modeling learning are not only temporal duals (Table 1, viewed row-wise), but perspective duals as well (Table 1, viewed column-wise).

These four functional equations define what can be called an *orthogonally adjoint* system. The system is orthogonally adjoint because, in one direction (designated the *temporal duals*), it consists of two pairs of dual equations to model hereditary (reinforcement) influences, on the one hand, and anticipatory (expectancy) influences, on the other. In the other direction (designated the *perspective duals*), two pairs of dual equations are defined over conjugate (information and energy) variables of the actor/perceiver control system, coupled in a mutual but reciprocal relationship with the task environment.

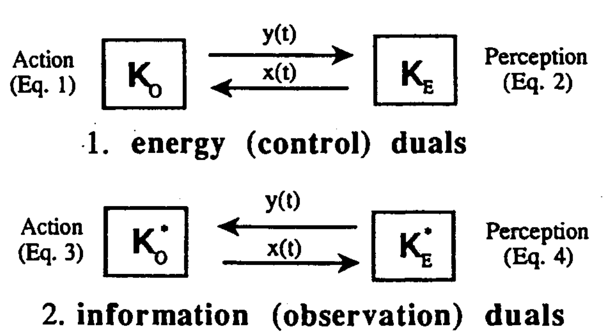

FIGURE 6. Diagram showing the interplay of the adjoint operators involved in the dual pair of dual Volterra learning functionals defined in Table 1.

To anticipate: Two mathematical questions have been raised by us—one general and the other specific. The general question is whether some kind of optimal control system with time lag can satisfy the stipulated system of learning equations. The answer is shown to be affirmative. The specific question is how to model the fact that whenever an action is successfully controlled and the goal obtained, then it must result from the accurate perception of environmental (e.g., task) constraints. This proves to be tantamount to the claim that when learning occurs, then an *inner product operator* exists whereby perceptual information is scaled invariantly over time to the resource requirements of control (e.g., energy). Again the answer is shown to be affirmative.

Let us consider the form that these dual equations for information and control take in adjoint systems theory. The field of adjoint systems is not well known; therefore, to make clear its utility for modeling the problems at hand, we suggest a graded development of this approach. Beginning with the simplest case, the *scalar* one, we then develop the equations for the intermediate *matrix* case, and, finally, for the *time lag*, or hereditary/anticipatory case.

For reasons given previously, the best model for adaptive systems is one which combines reinforcement-like and expectancy-like theories of learning into a single theory by means of the orthogonally adjoint system of equations (Shaw & Alley, 1985). Here, action, as a control variable, is dual to perception, as an observation variable, in the sense understood by modern control theory (Kalman et al., 1962). Interestingly enough, in modern control theory there exist several sets of dual (adjoint) equations that might be used to explain the process of learning as described.

A caveat is in order: Whether the controllability criterion is to be defined over least energy, least action, least work, least impulse, or the extremum of some other kinetic quantity is an empirical question. There is no a priori universal metric over which a trajectory might be variationally defined. The same must be said for observability and the corresponding kinematic quantities. Whether least time, least distance, least directional changes, or whatever extrema are to figure in the inner product is an empirical question to be determined by individual task requirements. What is to be recognized, in the spirit of general analysis, is the necessary existence of a minimax duality between some observable and some controllable quantities, with their inner product remaining invariant whenever a perceptually controlled intentional act is successful (Strang, 1986).

These are the fundamental properties that hold analogically over all tasks of this kind.

The next section summarizes the relevant equations from modern control theory that might be used in learning theory to model the duality of perception, formally construed as observation, and action, formally construed as control (Shaw & Alley, 1985). These equations can be expanded by considering a special class of control problems that can be described by adjoint equations. Such equations provide the tools needed to express the mutual and reciprocal relationship between the actor/perceiver and the task environment construed formally as dynamical systems. (For additional discussion, see Shaw & Mingolla, 1982; Shaw & Todd, 1980; Shaw & Turvey, 1981; Shaw et al., 1982).

Finally, equations are proposed for a time delay system that exhibits hereditary effects, and whose adjoint system exhibits anticipatory effects—on the analogy of reinforcement and expectancy theories, respectively. Thus, the time delay equations define adjoint systems with memory and anticipation. In addition, these systems exhibit the sequential (possibly nonlinear) effects of trials and other variables, such as learning to learn, that may significantly affect learning and transfer effects.

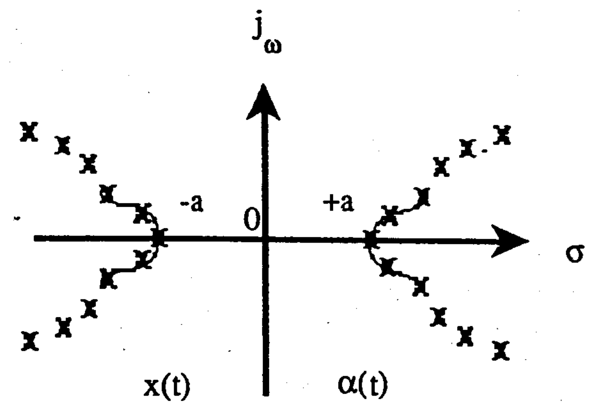

FIGURE 7. (Left) the pole-zero pattern of the original system $x(t)$ and its adjoint $\alpha(t)$ for a scalar, first-order ODE. (Right) original system $x(t)$ integrates forward in time, from initial time t_0 to final time t_f, whereas the adjoint system $\alpha(t)$ integrates backward in time, from final time t_f to initial time t_0.

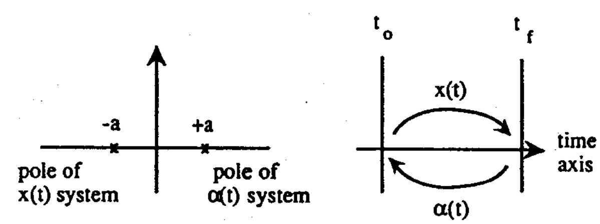

FIGURE 8. The pole-zero pattern for the time lag system and its adjoint system. The pole pattern for the original system $x(t)$ is on the left, whereas that for the adjoint system $\alpha(t)$ is on the right.

The Adjoint System: Scalar Case

Let us consider a simple scalar differential equation process and investigate its adjoint system. The scalar form of the adjoint system is the simplest case defined by temporally dual equations. Recall that this fundamental duality defines the relationship between perceptual information and action control. Here $a(t)$ is generally described as a continuous function of time t.

$$\dot{x}(t) = dx(t)/dt = a(t)x(t). \tag{3.1}$$

Equation 1 is integrated forward in time from $t = t_0$, the initial time, until $t = t_f$, the final time. Associated with the function $x(t)$ is the adjoint system $\alpha(t)$, which satisfies

$$\dot{\alpha}(t) = -a(t)\alpha(t), \tag{3.2}$$

where $\alpha(t)$ is specified at the terminal time. What is noteworthy is that $\alpha(t)$ is integrated backward in time from $t = t_f$ to $t = t_0$, and also differs from the right-hand side of Equation 1 by a change in sign. If constant $a > 0$, then Figure 7 illustrates the pole-zero diagram of the original system. Their adjointness, or duality, is manifested in their appearing as mirror images about the complex (j_ω) axis. All adjoint systems, including the vector/matrix cases and the time lag (hereditary) cases, exhibit this involutive character when plotted against their original system. Figure 8 illustrates how each equation is integrated forward or backward in time, respectively.

It is also possible for $a(t)$ in Equation 1 to vary with time: The adjoint system can be defined accordingly. An important quantity in analyzing the solutions to these equations is the fundamental, or state transition matrix, which is defined for the system Equation 1 as follows:

$$\dot{\phi}(t, t_0) = a(t)\phi(t, t_0), \quad \phi(t_0, t_0) = I \tag{3.3}$$

where I indicates 1 for the scalar case and a unit diagonal matrix for the matrix case. This reduces to

$$\phi(t, t_0) = e^{a(t - t_0)}, \tag{3.4}$$

if a is constant. The solution $x(t)$ of Equation 1 can be written as

$$x(t) = \phi(t, t_0)\, x(t_0). \tag{3.5}$$

For the adjoint system, the state transition matrix satisfies

$$\dot{\phi}_a(t, t_f) = a(t)\phi_a(t, t_f), \tag{3.6}$$

$$\phi_a(t_f, t_f) = I, \tag{3.7}$$

which, for a is constant, reduces to

$$\phi_a(t, t_f) = e^{a(t_f - t)}. \tag{3.8}$$

From this, it follows that

$$\phi_a(t_0, t) = \phi(t, t_0) \tag{3.9}$$

or

$$\phi_a^{-1} = \phi. \tag{3.10}$$

Thus, the adjoint system acts like an inverse system but proceeds backward in time. The most interesting relationship between the original system and the adjoint system is revealed by the inner product operator. For two scalars, $x(t)$ and $\alpha(t)$, the inner product operator is defined as the scalar product of the components. This turns out to be equal to a constant over the specified interval (e.g., the individual's learning interval); that is,

$$<\alpha(t), x(t)> = \alpha(t)\, x(t) = \text{constant}. \tag{3.11}$$

If one uses the simple example illustrated thus far, then for a = constant,

$$<x(t), \alpha(t)>' = x(t_0)e^{a(t - t_0)}\, \alpha(t_f)e^{a(t_f - t)}$$

$$= x(t_0)\alpha(t_f)\, e^{a(t_f - t_0)} = \text{constant}. \tag{3.12}$$

To better illustrate what happens if $a(t)$ depends on time, consider a time varying example (Bellman, 1973) that more explicitly demonstrates these calculations.

Example 1: Given the system

$$\dot{x}(t) = -t\, x(t), \text{ with } x(t_0) \text{ specified}, \tag{3.13}$$

then the solution is

$$x(t) = x(t_0)\, e^{(t_0^2 - t^2)/2}, \tag{3.14}$$

$$\phi(t, t_0) = e^{(t_0^2 - t^2)/2}, \quad \text{for } t \geq t_0; \tag{3.15}$$

whereas

$$\phi(t, t_0) = 0 \quad \text{for } t < t_0. \tag{3.16}$$

The adjoint system satisfies

$$\dot{\alpha}(t) = t\, \alpha(t), \text{ with } \alpha(t_f) \text{ specified}. \tag{3.17}$$

The solution is

$$\alpha(t) = \alpha(t_f)e^{(t^2 - t_f^2)/2} \tag{3.18}$$

and

$$\phi_a(t, t_f) = e^{(t^2 - t_f^2)/2}. \tag{3.19}$$

Note: $\phi_a(t, t_f)\phi(t, t_f) = e^{(t_0^2 - t_f^2)/2} = \text{constant}. \quad (3.20)$

also

$$<x(t), \alpha(t)> = x(t_0)e^{(t_0^2 - t^2)/2}\alpha(t_f)e^{(t^2 - t_f^2)/2} \tag{3.21}$$

or

$$<x(t), \alpha(t)> = x(t_0)\, \alpha(t_f)\, e^{(t_0^2 - t_f^2)/2} = \text{constant}. \tag{3.22}$$

The Adjoint Sytem: Matrix Case

The results presented here for the scalar case easily generalize to the matrix case. One can now define the dual properties of controllability and observability by which the reciprocal role of action and perception in the learning process can be modeled. In addition, one can also define the inner product operator, the means by which perceptual information can be scaled to the control of action. Following Kalman et al. (1962), the original sytem equations can be represented as an n component column vector $x(t)$ as follows:

$$\dot{\bar{x}}(t) = A(t)\bar{x}(t) + B(t)\bar{u}(t) \text{ with } \bar{x}(t_0) \text{ specified}, \tag{3.23}$$

where $A(t)$ is an $n \times n$ matrix, $B(t)$ is an $n \times p$ matrix, and $\bar{u}(t)$ is a $p \times 1$ control vector.

The solution to this system is given by

$$\bar{x}(t) = \phi(t, t_0)\, \bar{x}(t_0) + \int_{t_0}^{t} \phi(t, s)\, B(s)\, \bar{u}(s)\, ds, \tag{3.24}$$

where $\phi(t, t_0)$ and $\phi(t, s)$ are the state-transition matrices of the free system

$$\dot{\bar{x}}(t) = A(t)\, \bar{x}(t_0) \tag{3.25}$$

from $x(t_0)$ to $x(t)$ and from $x(s)$ to $x(t)$. Associated with the system depicted by Equation 23 is an observation vector $\bar{y}(t)$, an m component vector, which satisfies

$$\bar{y}(t) = H(t)\, \bar{x}(t), \tag{3.26}$$

where $H(t)$ is an $m \times m$ matrix relating the observation in the vector $\bar{y}(t)$ from $\bar{x}(t)$. If the system is real, the adjoint system associated with Equations 23 and 26 is given by

$$\dot{\bar{\alpha}}(t) = A^T(t)\, \bar{\alpha}(t) + H^T(t)\, \bar{v}(t), \tag{3.27}$$

$$\bar{z}(t) = B^T(t)\, \bar{\alpha}(t). \tag{3.28}$$

$\bar{\alpha}(t)$ is specified, and Equations 27 and 28 are integrated backward in time. The superscript T indicates matrix transpose. With this matrix notation, it is now possible to define some properties of the original system Equations 23 and 26, such as, *complete controllability and complete observability.* It is also possible to generalize the definition of the inner product operator as presented in the next section.

Properties of the Original System (Complete Controllability and Complete Observability) and the Inner Product Operator

The system, represented by Equations 23 and 26, is completely controllable if there exists some input $\bar{u}(t)$ that takes the system from any initial state $\bar{x}(t_0)$ to any other state $\bar{x}(t_1)$ in a finite length of time $t_1 \geq t_0$. This property holds if the following matrix is nonsingular for some $t_f > t_0$:

$$W(t_0, t_f) = \int_{t_0}^{t_f} \phi(t_f, t)\, B(t)\, \phi^T(t_f, t)\, dt. \tag{3.29}$$

The measure of complete controllability is related to the minimum amount of control energy $\bar{u}(t)$ necessary to transfer $\bar{x}(t_0)$ to $\bar{x}(t_f)$ in $t_f - t_0$ seconds.

Of interest to determining the optimality of the control is the degree to which the amount of work done approaches the minimum. For this, one needs an equation defining minimum energy:

$$\text{Min E} = \bar{x}^T(t_f)\, W^{-1}(t_0, t_f)\, \bar{x}(t_f). \tag{3.30}$$

Small values of $W(t_0, t_f)$ imply little controllability, because large amounts of energy are required to transfer $\bar{x}(t_0)$ to $\bar{x}(t_f)$, and vice versa.

Perceptual information guides action: Hence, a duality must exist between the energy required for control and the information that provides the measure of control. Such a measure is guaranteed by the duality of complete controllability to complete observability. This condition is defined next.

A system is said to be *completely observable* if it is possible to determine the exact value of $\bar{x}(t_0)$ given the values of $\bar{y}(t)$ in a finite interval (t_0, t_f), where $t_0 < t_f$. The original system represented by Equations 23 and 26 is completely observable if the following matrix is positive definite for some $t_f > t_0$:

$$M(t_f, t_0) = \int_{t_0}^{t_f} \phi(t, t_0)\, H(t)^T\, H(t)\, \phi(t, t_0)\, dt. \tag{3.31}$$

This motivates the next important lemma.

Lemma 1: A system is completely controllable if and only if its dual (adjoint) is completely observable.

The proof of this lemma follows directly from the dual relationships

$$\phi(t_f, t) = \phi_a^T(t, t_f) \text{ and} \tag{3.32}$$

$$B(t) = H_a^T(t), \tag{3.33}$$

and by substituting into Equations 29 and 31 the relationships specified by Equations 32 and 33.

Analogous to the case of minimum energy, one can ask: What happens to information when the system successfully achieves control of action with respect to some goal? Given the duality of complete observability with complete controllability, then whenever energy is minimized information must be maximized—a minimax duality (Strang, 1986). Thus, the measure of complete observability is related to the maximum amount of perceptual information as follows:

$$\text{Max info} = \bar{y}^T(t_f)\, M^{-1}(t_f, t_0)\, \bar{y}(t_f). \tag{3.34}$$

The last item of interest for the matrix case involves the inner product of the original system with its dual, for it provides a measure of the amount of control exercised as compared to the amount of information detected over the task interval. The definition of the inner product operator for the matrix case can also be given.

Definition: Inner Product Operator

Let $|\bar{x}|^2 = [\bar{x}^T \bar{x}]$ and $<\bar{x}, \bar{x}> = [\bar{x}^T \bar{x}]$,

where $\bar{x}$ is an $n \times 1$ column vector.

Using the above definition, a lemma can be constructed to show in the matrix case, as in the scalar case, that the inner product between the original system and its adjoint is a dynamical invariant. This invariant quantity expresses the fundamental law of intentional dynamics —the conservation of intention that elects and sustains the coordination of action with perception (i.e., the preceiving–acting cycle) (Shaw et al., 1990).

Lemma 2: If $\bar{u}(t) = 0$ in Equation 27, then $<\bar{x}(t), \bar{\alpha}(t)> = \bar{x}^T \bar{\alpha}$ = a constant.

Alternatively, this may occur if $u(t)$ is in feedback form,

$$\bar{u}(t) = -K(t)\, \bar{x}(t), \tag{3.35}$$

and Equation 35 is substituted into Equation 23.

These results may be extended to systems with hereditary influences, sometimes called systems with retardation, or, more commonly, time lag. Such systems are minimal for modeling adaptive changes in control caused by learning through reward. These results can also be extended to an adjoint system that exhibits dual anticipatory influences characteristic of learning as a function of change in expectancies.

The Adjoint System With Time Lag

The results of adjoint systems theory extend naturally to systems with time lag (Bellman, 1973; Eller, Aggarwal, & Banks, 1969). In this case, the plant (vector equation) satisfies

$$\dot{\bar{x}}(t) = A_1(t)\, \bar{x}(t) + A_2(t - \tau)\, \bar{x}(t - \tau) + B(t)\, \bar{u}(t) \tag{3.36}$$

$$\bar{y}(t) = H(t)\, \bar{x}(t). \tag{3.37}$$

An important change is introduced now into the matrix case as described by Equations 36 and 37. Whereas the simpler matrix case of control systems requires independent evaluation of the initial conditions of its differential equations, the time lag case does not. Instead, systems with time lag require an initial function to express any influence that builds up over time as a result of learning, transfer of learning, learning to learn, or fatigue. The initial function, a hereditary functional, is an operator that, in a sense, automatically updates the initial conditions of the differential equations of the matrix case system. It is this capability that renders the time lag systems truly adaptive so that the extrinsically controlled "tweaking" of individual parameters over trials or tasks is unnecessary.

The hereditary functional $\bar{y}(t)$ for the time lag system is a given continuous function on the interval $[t_0 - \tau, t_0]$. It follows that for the solution $\bar{x}(t)$,

$$\bar{x}(t) = \bar{\phi}(t) \text{ for } t \in [t_0 - \tau, t_0]. \tag{3.38}$$

The more precise notation for the solution of the system Equations 36 and 37 is

$$\bar{x}(t, t_0, \bar{\Phi}(t), \bar{u}(t)) = \bar{x})t, t_0, \bar{\Phi}(t), 0) + \int_0^t \phi(t, s)\, B(s)\, \bar{u}(s)\, ds \quad (3.39)$$

where $\phi(t, s)$ is the associated transition matrix, for which

$$d\phi(t, s)/ds = -\phi(t, s)\, A_1(s) - \phi(t, s + \tau)A_2(s + \tau),$$

$$t_0 < s < t_f - \tau,\ \phi(t, s) = 0 \text{ for } s > t, \text{ and } \phi(t, t) = I. \quad (3.40)$$

Also associated with system Equations 36 and 37 is the adjoint system $\bar{\alpha}(t)$, which satisfies

$$\dot{\bar{\alpha}}(t) = A_1^T(t)\, \bar{\alpha}(t) + A_2^T(t + \tau) + H^T(t)\, \bar{v}(t), \quad (3.41)$$

$$\bar{z}(t) = B^T(t)\, \bar{\alpha}(t). \quad (3.42)$$

The solution of this dual equation has a form similar to $\bar{x}(t, t_0, \bar{\Phi}(t), \bar{u}(t))$ in Equation 39 above (Repperger, 1974; Repperger & Koivo, 1974; Weiss, 1970).

Definition: Complete Controllability

The sufficient condition for the system represented by Equations 36 and 37 to be completely controllable is that the following matrix be of full rank:

$$G_1(t_0, t_f) = \int_{t_0}^{t_f} \phi(t_f, s)\, B(s)\, B^T(s)\, \phi^T(t_f, s)\, ds, \quad \text{for } t_f > t_0. \quad (3.43)$$

Definition: Complete Observability

The system is likewise completely observable if the following matrix is of full rank:

$$G_2(t_0, t_f) = \int_{t_0}^{t_f} \phi(s, t_f)\, H(s)\, H^T(s)\, \phi^T(s, t_f)\, ds, \quad \text{for } t_f > t_0. \quad (3.44)$$

The inner product definition now extends to the following:

$$<\bar{x}(t), \bar{\alpha}(t)> = \bar{x}^T(t)\, \bar{\alpha}(t) + \int_{-\tau}^{0} \bar{x}^T(t + s)\, \bar{\alpha}(t + s)\, ds. \quad (3.45)$$

It is useful to study the time lag system and its adjoint system in the complex plane. Figure 8 illustrates the pole-zero diagram for the original system $\bar{x}(t)$ and its adjoint system $\bar{\alpha}(t)$ for the scalar time lag example. System $\bar{x}(t)$ can be classified as an infinite dimensional system, with its poles having a particular pattern. The mirror image about the j_ω-axis of the pattern is the diagram for $\bar{\alpha}(t)$, the adjoint system. Because the poles of $\bar{x}(t)$ and $\bar{\alpha}(t)$ are both ordered, then the time delay system $\bar{x}(t)$ is considered a hybrid between a finite dimensional system (a finite number of poles), as shown in Figure 7, and a true infinite dimensional system (such as a partial differential equation system), as shown in Figure 9. For a true infinite dimensional system, the pole-zero pattern would be scattered throughout the complex plane with many different curves extending to infinity in various directions.

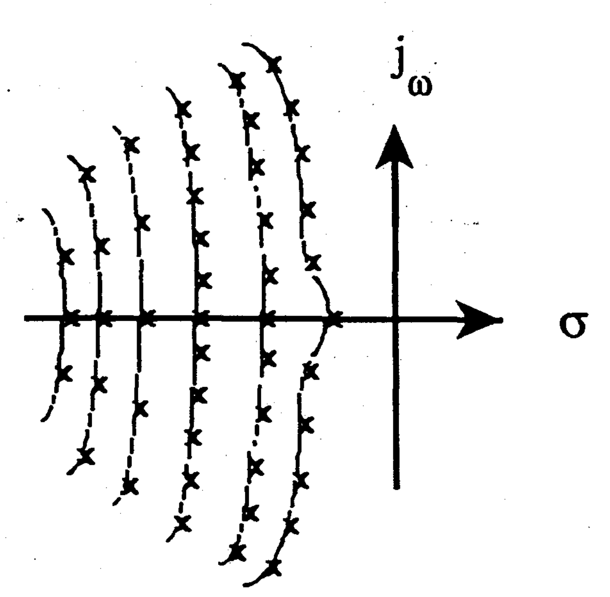

FIGURE 9. The pole-zero pattern for a true infinite dimensional system—a partial differential equation.

Final Formulation of the Adjoint Learning Model: A Summary

From the numerous equations in the earlier sections, it is useful to select a set of equations to fit those postulated in Section 1. A time-forward system is needed to model action and a time-backward system to model perception. The inner product operator demonstrates the existence of an invariance that should hold if and only if the learning process is successful. Furthermore, because learning appears to evolve in a continuously cumulative fashion, and one wishes to have a formulation that relates over time, the time lag system with an initial function representation seems quite plausible. Let us see how this time lag variable works.

As argued above, the most general form for the learning model would be a time lag system directed forward in time to represent the action (control) variable. Recall Equations 38 and 36 in which the initial function $\phi(t)$ represents hereditary changes due to experience. The dual, or adjoint variable required to measure perceptual learning could be modeled by $\bar{\alpha}(t)$, which is directed backwards in time. Also recall Equations 41 and 42 for $\bar{\alpha}(t)$. The state transition matrices $\phi(t, s)$ satisfy Equations 39–41. Thus, by the earlier definitions, the system is completely controllable if $G_1(t_0, t_f)$ of Equation 43 is of full rank, and completely observable if $G_2(t_0, t_f)$ of Equation 44 is also of full rank. Because the adjoint system proceeds backward in time, this seems to be a viable model.

The invariant rule determined by the inner product suggests a measure of the system's innate capacity for learning to learn or, conversely, for fatigue:

$$Q = \langle\alpha(t), x(t)\rangle = \text{constant}; \quad (3.46)$$

$$Q = x^T(t)a(t) + \int_t^0 x^T(t+s)\,\alpha(t+s)\,ds. \quad (3.47)$$

Notice that large values of Q indicate high values of coordination between action and perception, which implies a large capacity for learning to learn. By contrast, when Q is small in value, this indicates a low capacity for learning to learn or may be related to subject fatigue.

Summary and Conclusions

Learning is better characterized as a (possibly nonlinear) hereditary/anticipatory (Volterra) functional than as more simplified models. We have offered reasons that perceiving and acting must relate as dual functionals that reciprocate in a cyclic manner that satisfies specific adjointness properties and become self-adjoint if intentional learning is successful. Arguments were presented that a semantic correspondence might be set up between certain mathematical properties in the proposed modeling strategy and the most significant characteristics of any intentional learning process (e.g., cumulativity, hereditary, and anticipatory influences (Section 1).

An important caveat should be recognized however: No model for the mechanism of intentional learning (e.g., filter model, neural net) has been offered, rather only a general strategy for formally representing the most lawful aspects of any intentional learning process (Section 2)—the sufficient condition under which such learning might be successful. In addition, it has been shown (Section 3) that the proposed mathematical strategy of using adjoint functionals has substantial justification in many fundamental theorems of control theory (e.g., Kalman's duality theorem and inner product invariant).

We close by summarizing some of these most significant aspects of the intentional dynamics theory of learning.

1. The effects of learning are cumulative, in that previous experiences facilitate later performance. The differential equation system denoted by Equations 36–38 and 41–42 models this property by the forward and backward integration of $x(t)$ and $\alpha(t)$. These values determine anticipatory and hereditary influences on learning that accumulate forward and backward in time, respectively.

2. Learning, in the most general case, can be characterized as intentional modulation of the learner's perceiving–acting cycle, as predicted from the principle of duality (or mutuality, or reciprocity)—the basic principle of ecological psychology. The set of Equations 38, 36, and 41–42 has this effect where $x(t)$ supports the past-pending (action) perspective and $\alpha(t)$ supports the future-tending (perception) pespective.

3. From the symmetry of the system to equations defined over temporal (information/energy) duals, Equations 38, 36, and 41–42 imply the perspective duals, defined by reciprocal constraints of the environment (instructor) and organism (learner). From points (3) and (4), one sees that these equations provide a mathematically coherent account of an adaptive ecosystem. They provide both the motive and means for incorporating learning theory into ecological psychology, as originally proposed by Shaw and Alley (1985). (See also Newell, 1991.)

4. Intentional learning consists of three component types of learning: action learning, characterized here by $x(t)$; perceptual learning, characterized here as $\alpha(t)$; and intentional learning, characterized here as the (nonlocal) constraint that sets up the inner product $\langle x, \alpha\rangle = a$. Perceptual learning and action learning give rise to two sets of data points that must change in a coordinated reciprocal way as the time in the learning situation increases, if learning the task intention is to be successful. This results in their inner product remaining invariant over the learning, that is $\langle x, \alpha\rangle = a$ constant, whenever the boundary conditions defining the task intention are satisfied.

5. The boundary conditions for the task intention (specified by goal parameters) must be passed from the instructor to the learner, construed as a self-adjoint operator, via the shared forcing function. The forcing function, a vehicle for instruction, includes both a controlling function and an informing function (Figure 4B). When these component functions are well coordinated by intention over time, as evidenced by their invariant inner product, then the learner comes to satisfy the controllability and observability criteria characteristic of successful intentional learning. This point is sufficiently important to be underscored.

If the learner's intention wavers so that his/her attention to the relevant information falters, then this psychological fault corresponds to the mathematical fact that the inner product does not remain invariant, and the attempt to learn to do the task as instructed fails. On the other hand, the successful education of intention means that intention does not waver and the inner product invariance is assured. Hence, large changes that occur in $\alpha = \Delta\alpha$ produce large changes in $x = \Delta x$. That is, large changes in the sensitivity of detection processes induce large changes in the precision of action-control processes. Likewise, small changes in α induce small changes in x.

6. Learning to learn and fatigue effects also have natural interpretations within the adjoint systems context. If $Q_1 = \langle x, \alpha\rangle$ is one experimental condition and $Q_2 = \langle x, \alpha\rangle$ another, and if $Q_1 \gg Q_2$, then a greater

disposition for learning is represented by the first experimental condition. Alternatively, Q_2 may indicate a condition of greater fatigue for the subject than Q_1.

7. The integro-differential equations suggested by Shaw and Alley (1985) for learning, as depicted in Table 1, are seen now to apply to the response variable $y(t)$ in terms of the trial-to-trial variables $x(t)$ and take the general form:

$$y(t) = kx(t) + \int_a^b K(t, s)\, x(s)\, ds. \qquad (3.48)$$

This is satisfied by Equations 38, 36, and 41–42, as can be easily shown.

8. The ecological approach to psychology argues that there should be a duality between action and perception. The equations for $x(t)$ and $\alpha(t)$ given here, indeed, are duals in a strict mathematical sense. If $x(t)$ is completely controllable, then $\alpha(t)$ is completely observable. If $x(t)$ is completely observable, then $\alpha(t)$ is completely controllable. Thus, orthogonally adjoint systems provide a rigorous mathematical formulation of this most fundamental assumption of the ecological approach and the fundamental law of intentional dynamics.

Figure 10(a) depicts the asymptotic learning curve derived from measurement of change in the action

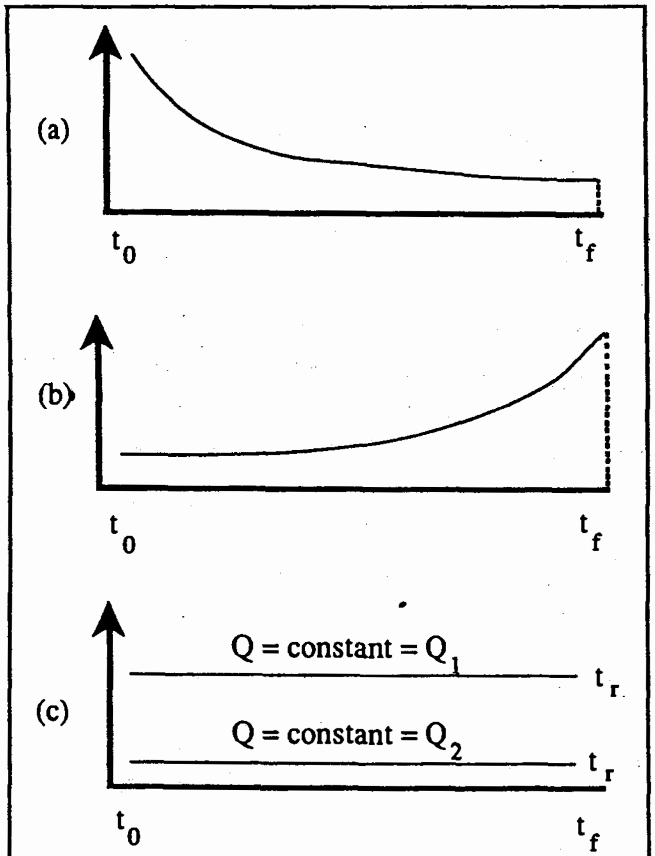

FIGURE 10. Diagram of the inner product invariant that exists between action and perception over successful learning (see text for details).

(energy) variable $x(t)$ over trials $= \tau$. Figure 10(b) depicts the corresponding dual learning curve derived from measurement of the adjoint perception (information) variable (t) over trials $= t_f - \tau$. Figure 10(c) shows the curves resulting from the inner product operator $= \langle x(t), \alpha(t) \rangle = \dot{Q}$, a constant over the successful learning session. Relatively high Q values, (Q_1), represent learning to learn, whereas relatively low values, (Q_2), represent fatigue or low capacity for learning. (Note: This is not the Q value of Kugler & Turvey, 1987, but may be related. A mathematical derivation would be required to draw this conclusion, however.)

9. Because action response measurements tend to stabilize with time (e.g., error scores on the same task with repeated trials), then, conversely, one would expect the processing of perceptual information, the dual of the action variable, to exhibit unstable response characteristics. Figures 10(a–c) illustrate the form of the relationships these variables take. One can notice from Figure 8 that if $\dot{x}(t)$ is stable (left-half plane pole), then $\alpha(t)$ should show unstable characteristics (right-half plane pole). Thus, if $x(t)$ is asymptotically stable in time, $\alpha(t)$ should show an unstable solution. This may help to explain goal gradient effects, that is, why uncertainty decreases and energy expenditure increases as the organism approaches its goal. Figure 10(c) illustrates how this effect of the proportionality of perception and action is an invariant quantity determined by the operation of the inner product operator Q on $x(t)$ and $\alpha(t)$. Notice that Q, the measure of the disposition to learn, is illustrated for both high and low values of Q.

These are some of the main properties revealed by a first pass on the adjoint approach to intentional learning. Future efforts should be directed toward seeking empirical evidence for the parametric interpretation of the system of dual functionals that underlies this approach for specific learning tasks. One such effort has already begun. This is described next and used as a vehicle for an interesting generalization.

Minimal Covariant Coupling as Lawful Basis of the Kinematically Specified Dynamics (KSD) Principle

A paradigm of extreme interest to our adjoint systems approach has been recently developed by Richard Schmidt in Michael Turvey's Action Laboratory at The University of Connecticut. This paradigm exploits the fact that the controlling (directive feedforward information) and informing (corrective feedback information) functions have a dual phasic relationship across I and L systems via their shared forcing function. Except, here there is only an information coupling of the two systems so that no active or reactive forces exist. Here, strictly speaking, there is no forcing function in the usual sense!

In this paradigm, social dyads are coupled by visual rather than haptic information. The task of the learner is to be visually instructed to swing his/her self-controlled pendulum (a leg) in phase or out of phase by seeing how the instructor swings his/her manually controlled pendulum (a leg). Our discussion also fits self-instructional paradigms, such as the Kugler and Turvey (1987) bimanual swinging pendula task or the Kelso, Scholz, and Schöner (1986) paradigm that uses bimanual oscillation of two index fingers. Thus, instructor–learner interactions are strictly kinematic; there are no kinetics in the forcing function.

Nevertheless, for the learner to succeed in the task of learning to swing his/her pendulum in phase or out of phase with the instructor's pendulum, there must be a self-adjoint (Green's) operator defined on the controlling and informing functions that are somehow supported by the kinematically defined forcing function; that is, "light must get to the appropriate muscles" (see Schmidt, 1990, for details). Kugler and Turvey (1987) have spoken of such informational couplings as "soft" (kinematic) couplings, as opposed to "hard" (kinetic) couplings whenever there is no mass term in Newton's law. Let us suggest one way this might work.

If one takes literally the absence of a mass term in Newton's law, then one has $F_L = a_I$ (rather than $F = ma$), where F_L specifies the learner's controlling forces needed to do the task and a_I specifies the instructor-produced instructional kinematics. Clearly, this formulation leaves the forcing function undefined. A better formulation incorporates the dual influence (Green's) functional required for the task intention by defining it over the following kinematically specified forcing function

$$F_L/km_L = a_I, \text{ where } km_L = m_I,$$

and k is a scaling factor that relates the masses in the L-pendulum system to the masses in the I-pendulum system.

Hence, the concept of this forcing function is more aptly described as a *minimal covariant* coupling, which is satisfied whenever the I-tool system commutes with the L-tool system (say, in the sense of their Lie bracket product being zero or their covariant derivative being symmetrical in the manner described by Shaw et al., 1990). The notion of the forcing function as a covariant coupling constrasts sharply with the notion that it is a very low energy coupling, say, in the sense of entrainment. This covariant view provides an explicit formulation to Runeson's so-called KSD principle, which asserts that dynamics can be kinematically specified (Runeson & Frykholm, 1981).

In other contexts, we have called this a *reciprocity mapping* (Shaw et al., 1990) and k the common *ecometric* bases to information detection and energy control measures (Shaw & Kinsella-Shaw, 1988). If intention, as an abstract dynamical invariant, is to be made explicit, the action over the coupling must then be a generalized Hamiltonian (in the sense of Shaw et al., 1990) rather than an ordinary Hamiltonian. Viewed in this way, the Schmidt (1990), Kelso et al. (1986), and Kugler and Turvey (1989) paradigms exemplify intentional pendulum tasks analogous to our intentional spring task. Both tasks come under this ecometric rubric, with the learner performing the task successfully if and only if task intention is a conserved quantity under a (self-)instructing function in lieu of the usual forcing function. The pertinent question, of course, is what tasks do not fall under the purview of the proposed theory of intentional dynamics.

ACKNOWLEDGMENT

This work was supported in part by a summer faculty research award to the senior author from the Air Force Office of Scientific Research and by NSF grant number NSF IRI-8902755. Thanks are due to Rik Warren and the staff of the Human Engineering and Human Factors Division of the Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, for their help and support on various aspects of this project. In addition, the contribution of Thomas Carolan of the Institute for Simulation and Training, University of Central Florida, is gratefully acknowledged. We thank also Claudia Carello, Oded Flascher, Tony Morris, and Greg Burton at the Center for the Ecological Study of Perception and Action for help in the production of the manuscript.

APPENDIX A

Absorbing the Forcing Function

The learner system is allowed to wrest control from the instructor by incorporating the forcing function into its own boundary conditions and, hence, intentionalizing them in the process. The system takes the following form:

$$\mathbf{L}[y] = P(t)\ddot{y} + q(t)y + r(t)y = g(t). \tag{A.1}$$

Solution y for this system depends on the forcing term g(*t*). The linear operator **L** on the left-hand side of Equation A.1 can be used to represent the ODE on the right-hand side. Using the operator notation, one can translate the ODE problem into another problem defined in Hilbert space. Finding a solution for our differential equations becomes a problem of finding functions in the Hilbert space of double differentiable functions that are mapped by **L** into g(*t*).

The mathematical techniques used to transfer this goal-specific control from the instructor system to the learner system involves the well-known fact that nonhomogeneous SODEs with homogeneous boundary conditions can always be put into a homogeneous form with inhomogeneous boundary conditions. This is accomplished by reformulating the boundary conditions that the equation must satisfy so as to "absorb" the controlling influence of the operator's forcing function. Formally, this requires the knowledge of a particular solution to Equation A.1. The well-known mathematical theorem says that, with the help of a particular solution, the task to find any other solution is reducible to the task of finding solutions of the homogeneous equation

$$\mathbf{L}[y] = p(t)\,\ddot{y} + q(t)\,\dot{y} + r(t)\,y = 0. \tag{A.2}$$

The general solution of Equation A.1 can be expressed with the linear combination of any two independent solutions of Equation A.2 and a particular solution of Equation A.1.

$$y(t) = y_p(t) + c_1y_1(t) + c_2y_2(t). \tag{A.3}$$

One may translate this formal decomposition of solutions into the language of ecological psychology. Finding a particular solution means proving the viability of the required action, that is, showing that the initial inner product between information detection (observability) and action control (controllability) that held for the instructor-spring system as a self-adjoint system remains invariant when passed to the learner. Learning means that the learner–spring system has successfully searched the functional space of solutions to the homogeneous equation.

Getting rid of the forcing function as an extrinsic influence accomplishes the first step in intentionalizing a dynamical system. But how are the other outside influences eliminated, the boundary conditions, so that the system might become intentionally autonomous? This problem is addressed next.

Absorbing Boundary Conditions: The Sturm-Liouville Theory

Two important points need to be noted about natural boundary conditions and the relationship they have to differential equations used to capture the laws of nature: First, a differential equation alone, without boundary conditions, cannot provide a unique solution. But boundary conditions are not part of the law that might be expressed in differential terms. Rather, they are logically complementary to the law and, therefore, must be added as extrinsically imposed constraints on the system. ". . . from the physical standpoint a 'boundary condition' is always a simplified description of an unknown mechanism which acts upon our system from the outside. A completely isolated system will not be subjected to any boundary conditions" (Lanczos, 1961, p. 504).

Second, learning entails the assimilation of intention from one system by another system. This assimilation process can be interpreted as implying that the boundary conditions become absorbed into that system's dynamical equation. But this cannot happen in the case of differential equations. Integro-differential equations, however, have no boundary conditions outside the limits placed on its kernel transform. The absorption process can be accomplished as follows:

Equation A.2 is said to be *exact* if it can be written in the form

$$d[p(t)\dot{y}]/dt + f(t)y = 0, \qquad \text{(A.4)}$$

where $f(t)$ is expressed in terms of $p(t)$, $q(t)$, and $r(t)$. By equating the coefficients of the above equations, and then eliminating $f(t)$, one finds that a sufficient and necessary condition for exactness is

$$\ddot{p}(t) - \dot{q}(t) + r(t) = 0. \qquad \text{(A.5)}$$

If a linear homogeneous SODE is not exact, then it can be made so by multiplying by an appropriate integrating factor $s(t)$. For this, one needs $s(t)$ such that

$$s(t)\,p(t)\,\ddot{y} + s(t)\,q(t)\,\dot{y} + s(t)\,r(t)\,y = 0, \qquad \text{(A.6)}$$

which can be written as

$$d[s(t)\,p(t)\,\dot{y}]/dt + f(t)\,y = 0. \qquad \text{(A.7)}$$

Again, by equating coefficients in these two equations and then eliminating $f(t)$, one discovers that the function s must satisfy

$$p\ddot{s} + (2\dot{p} - q)\dot{s} + (\ddot{p} - \dot{q} + r)\,s = 0. \qquad \text{(A.8)}$$

This is the adjoint equation to the original Equation A.2.

For a large class of physically meaningful equations, it can be shown that the original equation has an adjoint whose ad-

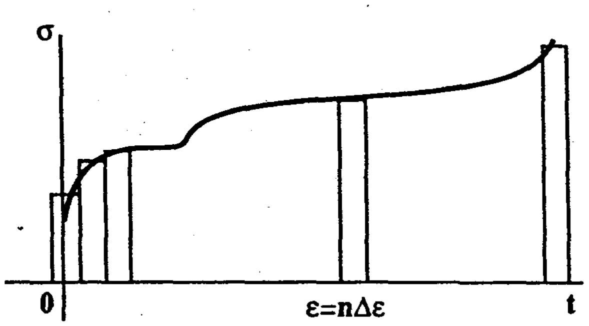

FIGURE 11. Boltzmann's superposition principle: Stress as a superposition of impulse strains.

joint is the original equation. This is called a *self-adjoint equation*. The original equation stated above is a case in point. It can be shown that the condition for Equation A.2 to be self-adjoint is that

$$\dot{p}(t) = q(t). \qquad \text{(A.9)}$$

A very useful fact is that any SODE can be made self-adjoint by multiplying from the left by

$$\exp \int \frac{q(t) - \dot{p}(t)}{p(t)}. \qquad \text{(A.10)}$$

Because the concept of Hilbert space and the L linear operator on it has been introduced, one can talk about eigenvalues and eigenvectors (eigenfunctions) of this operator. Given this theoretical framework, it is natural to seek a solution of $L[y] = g$ by looking for L^{-1}—the inverse operator of L. In many important cases, it can be found. The inverse operator is an integral operator in the form

$$L^{-1}[g(t)] = \int_a^b G(t, v)\, g(v)\, dv. \qquad \text{(A.11)}$$

The G kernal is called the Green's function of the operator L.

First, the existence of the Green's function needs to be guaranteed. The general condition for its existence is that the homogeneous equation with the given boundary conditions must have no solutions that are nondegenerative. Generally, finding the Green's function is not a simple problem. For a Sturm-Liouville system, however,

$$d[p(t)\dot{y}] + f(t)\, y = g, \qquad \text{(A.12)}$$

A Green's function can always be found by means of the following theorem:

Theorem 1. Suppose that the system given by Equation A.12 has no null eigenvalues (i.e., has only a nondegenerative eigenfunction) on the interval $[a, b]$, and assume the imposed boundary conditions are

$$c\, u(a) + d\, u'(a) = 0, \qquad \text{(A.13)}$$

$$e\, u(b) + d\, u'(b) = 0, \qquad \text{(A.14)}$$

where $p(t) > 0$, $dp(t)/dt$ exists and is continuous in $[a, b]$, and $f(t)$ is continuous in $[a, b]$ where $\{f(t) \in C[a, b]\}$, c, d, e, and f are non zero. Then, for any $f(t) \in C[a, b]$, the system has a unique solution,

$$u(t') = \int_a^b G(t', t)\, g(t)\, dt, \qquad \text{(A.15)}$$

where $G(t', t)$ is the desired Green's function given by

$$G(t', t) = \frac{u_2(t')\,u_1(t)}{p(t)\,W(t)},\ \text{for } a < t < t',$$

$$G(t', t) = \frac{u_1(t')\,u_2(t)}{p(t)\,W(t)},\ \text{for } t' < t < b, \qquad \text{(A.16)}$$

where the $u_i(t')$ functions are nonzero solutions to the homogeneous system given by Equation A.12, with the boundary conditions specified by Equations A.13 and A.14, and where $W(t)$ is the Wronskian. Thus, notice that the given boundary conditions, as extrinsically imposed requirements on the ODE representation of the system represented by Equation A.12, are now absorbed into the IDE representation given by equation A.15.

APPENDIX B

Expressing Hereditary Influences by Means of IDEs

Hooke's law ($F = -kx$) provides a simple linear model for elastic materials. Practically speaking, however, there is no pure elastic material. Traditionally, viscous materials are contrasted with elastic materials. Experimental data show that materials are neither viscous nor elastic, but rather viscoelastic. They exhibit a variety of hereditary effects at different time scales. The more elastic the material, the longer the time scale required to produce a hereditary effect. On the other hand, the more viscous the material, the shorter the time scale to produce the effect.

Originally, Hooke's law was formulated to describe the force versus elongation relationship for a simple spring. In reality, however, the k value is apt to change, even though in the original equation, $F = -kx$, k is treated as a constant. The description given below can be found in any basic book of continuum mechanics. The Volterra model is an integro-differential form that specifically for viscoelastic materials takes the generic form

$$\sigma(x, t) = -k(x, t)\,\epsilon(x, t) + \int_{-\infty}^{t} -k(x, t, \tau)\,\epsilon(x, \tau)\,d\tau,$$

where σ is the stress, k is the viscoelasticity (so-called *relaxation*) parameter, and ϵ is the amount of strain. The independent variables are: the spatial variable x, which is mostly constant, refers to the assumed homogeneity of the material; t and τ both denote time variables. This equation can be derived simply by using Boltzmann's superposition principle. That is, the integral denotes the superposition of small stretch impulses. To simplify the explanation, one assumes homogeneity. In this way, one eliminates the x variable. One also assumes that ϵ is differentiable with respect to time, which yields

$$\sigma(t) = K(t, -\infty)\,\epsilon(-\infty) + \int_{-\infty}^{t} K(t, \tau)\dot{\epsilon}(\tau)\,d\tau,$$

where $\frac{d}{dt}K(t, \tau) = -k(t, \tau)$, and $K(t, -\infty) = -k(t)$.

In many cases, $\epsilon(-\infty)$ vanishes, which leads to further simplification

$$\sigma(t) = K(t, 0)\,\epsilon(0) + \int_0^t K(t, \tau)\,\dot{\epsilon}(\tau)\,d\tau$$

or to the equivalent form

$$\sigma(t) = K(t, 0)\,\epsilon(0) + \int_0^t \frac{d}{d\tau}K(t, \tau)\,\epsilon(\tau)\,d\tau,$$

where $K(t, 0)\,\epsilon(0)$ represents the initial condition before the hereditary change starts. These equations are simply the condensed form of Figure 11, illustrating Boltzmann's superposition principle.

REFERENCES

Beek, P. (1988). Exploring the dynamics of juggling. In A. M. Colley & J. R. Beech (Eds.), *Cognition and action in skilled behavior* (pp. 229–246). Amsterdam: North-Holland.

Beek, P. J., & Beek, W. J. (1988). Tools for constructing dynamical models of rhythmic movement. *Human Movement Science, 7,* 301-342.

Bellman, R. E. (1973). *Methods of nonlinear analysis.* New York: Academic Press.

Bernstein, N. (1967). *The coordination and regulation of movements.* Oxford: Pergamon Press.

Bizzi, E. (1980). Central and peripheral mechanisms in motor control. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 131–144). Amsterdam: North-Holland.

Boyce, W., & DiPrima, R. (1986). *Elementary differential equations and boundary value problems.* New York: Wiley.

Bransford, J. D., Stein, B. S., Shelton, T. S., & Owings, R. A. (1981). Cognition and adaptation: The importance of learning to learn. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment.* Hillsdale, NJ: Erlbaum.

Cooke, J. D. (1980). The organization of simple, skilled movements. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 199–209). Amsterdam: North-Holland.

Courant, R., & Hilbert, D. (1937). *Method of mathematical physics, 1.* New York: Interscience Publishers.

Einstein, A., & Infeld, L. (1938). *The evolution of physics.* New York: Simon and Schuster.

Eller, D. H., Aggarwal, J. K., & Banks, H. T. (1969). Optimal control of linear time delay systems. *IEEE Transactions on Automatic Control, AC-14,* 678–687.

Feldman, A. G. (1966a). Functional tuning of the nervous system with control of movement or maintenance of a steady posture. II. Controllable parameters of the muscle. *Biophysics, 11,* 565-578.

Feldman, A. G. (1966b). Functional tuning of the nervous system with control of movement or maintenance of a steady posture. III. Mechanographic analysis of execution by man of the simplest motor task. *Biophysics, 11,* 766–775.

Feldman, A. G. (1980). Superposition of motor programs: I. Rhythmic forearm movements in man. *Neuroscience, 5,* 81-90.

Feldman, A. G. (1986). Once more on the equilibrium-point hypothesis (λ model) for motor control. *Journal of Motor Behavior, 18,* 17-54.

Fodor, J. A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. In S. Pinker & J. Mehler (Eds.), *Connections and symbols*l (pp. 3-71). Cambridge, MA: MIT Press.

Fre'chet, M. (1925). L'analyse ge'ne'rale et les ensembles abstraits [General analysis and abstract sets]. *Revue de metaphysique et morale.* Vol. 32

Fung, C. (1981). *Biomechanics: Mechanical properties of living tissue.* New York: Springer-Verlag.

Golden, J. M., & Graham, G. A. C. (1988). *Boundary value problems in linear viscoelasticity.* New York: Springer-Verlag.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Hinton, G. (1984). Some computational solutions to Bernstein's problems. In H. T. A. Whiting (Ed.), *Human motor* *actions: Bernstein revisited* (pp. 413–438). Amsterdam: North-Holland.

Kalman, R. E., Englar, T. S., & Bucy, R. S. (1962). Fundamental study of adaptive control systems. ASD-TR-61-27, 1.

Kelso, J. A. S., & Kay, B. A. (1987). Information and control: A macroscopic analysis of perception-action coupling. In H. Heuer & A. F. Sanders (Eds.), *Tutorials in perception and action* (pp. 3-32). Amsterdam: North-Holland.

Kelso, J. A. S., Sholz, J. P., & Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters, 118,* 279–284.

Kramer, E. E. (1970). *The nature and growth of modern mathematics.* New York: Hawthorn.

Kugler, P. N., & Shaw, R. E. (1990). Symmetry and symmetry-breaking in thermodynamic and epistemic engines. A coupling of first and second laws. In H. Haken & M. Stadter (Eds.), *Synergetics of cognition* (pp. 296–330). Proceedings of the International Symposium at Sloss Elmau, Bavaria, June 4–8, 1989.

Kugler, P. N., Shaw, R. E., Vicente, K. J., & Kinsella-Shaw, J. M. (1990). Inquiry into intentional systems. I: Issues in ecological physics. *Psychological Research, 52,* 98–121.

Kugler, P. N., & Turvey, M. T. (1987). *Information, natural law, and the self-assembly of rhythmic movement: Theoretical and experimental investigations.* Hillsdale, NJ: Erlbaum.

Lanczos, C. (1961). *Linear differential operators.* London: D. Van Nostrand.

Lindsay, R. B., & Margenau, H. (1957). *Foundations of physics.* New York: Dover.

Newell, K. M. (1991). Motor skill acquisition. *Annual Review of Psychology, 42,* 213–237.

Picard, E. (1907). La mecanique classique et ses approximations successives [Classical mechanics and its successive approximations]. *Revista di scienza, 1,* 4–15.

Repperger, D. W. (1974). An algorithm for the computation of the minimum energy solution of systems with retardation. *The International Journal of Control, 19,* 1047-11055.

Repperger, D. W., & Koivo, A. J. (1974). On stable forward filtering and fixed-lag smoothing in a class of systems with time delay. *IEEE Transactions on Automatic Control, AC-19,* 266–268.

Repperger, D., & Shaw, R. E. (1990). An adjoint systems approach to learning and transfer of training: 2. Mathematical formulation. *PAW Review: A Technical Report of the Perceiving Acting Workshop in Ecological Psychology.*

Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 733–740.

Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 227–247.

Shaw, R. E. (1987). Behavior with purpose. *Contemporary Psychology, 3,* 243–245.

Shaw, R. E., & Alley, T. R. (1985). How to draw learning curves: Their use and justification. In T. D. Johnston & A. T. Pietrewicz (Eds.), *Issues in ecological study of learning* (pp. 275–443). Hillsdale, NJ: Erlbaum.

Shaw, R. E., & Carolan, T. (1988). Adjoint optimal control. *Behavioral and Brain Sciences, 11,* 146–147.

Shaw, R. E., & Kinsella-Shaw, J. M. (1988). Ecological mechanics: A physical geometry for intentional constraints. *Human Movement Science, 7,* 155–200.

Shaw, R. E., Kugler, P. N., & Kinsella-Shaw, J. M. (1990). Reciprocities of intentional systems. In R. Warren & A. Wertheim (Eds.), *Control of self-motion* (pp. 579–619). Hillsdale, NJ: Erlbaum.

Shaw, R. E., & Mingolla, E. (1982). Ecologizing world graphs. *Behavioral and Brain Sciences, 5,* 648–650.

Shaw, R. E., & Repperger, D. (1989a). An adjoint systems approach to learning and transfer of training: 1. Introduction. *PAW Review: A Technical Report of the Perceiving-Acting Workshop in Ecological Psychology.*

Shaw, R. E., & Repperger, D. (1989b). An adjoint systems approach to learning and transfer of training: 3. Formulation and support for the learning model. *PAW Review: A Technical Report of the Perceiving-Acting Workshop in Ecological Psychology.*

Shaw, R. E., & Todd, J. (1980). Abstract machine theory and direct perception. *Behavioral and Brain Sciences, 3,* 400–401.

Shaw, R. E., & Turvey, M. T. (1981). Coalitions as models of ecosystems: A realist perspective on perceptual organization. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 343–415). Hillsdale, NJ: Erlbaum.

Shaw, R. E., Turvey, M. T., & Mace, W. M. (1982). Ecological psychology: The consequences of a commitment to realism. In W. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes II* (pp. 159–226). Hillsdale, NJ: Erlbaum.

Smith, T. J., & Smith, K. U. (1987). Feedback control systems of human behavior. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 251-309). New York: Wiley.

Solomon, H. Y. (1988). Movement-produced invariants in haptic explorations: An example of a self-organizing, information-driven, intentional system. *Human Movement Science, 7,* 201–223.

Solomon, H. Y., Turvey, M. T., & Burton, G. (1989). Gravitational and muscular variables in perceiving extent by wielding. *Ecological Psychology, 1,* 265–300.

Strang, G. (1986). *Introduction to applied mathematics.* Wellesley, MA; Wellesley Cambridge Press.

Synge, J. L. (1970). *Talking about reality.* Amsterdam: North-Holland.

Tricomi, F. G. (1957). *Integral equations.* New York: Dover.

Turvey, M. T., & Shaw, R. E. (1979). The primacy of perceiving: An ecological reformulation of perception for understanding memory. In L.-G. Nilsson (Ed.), *Perspectives on memory research* (pp. 167–222). Hillsdale, NJ: Erlbaum.

Weiss, L. (1970). An algebraic criterion for controllability of linear systems with time delay. *IEEE Transactions on Automatic Control, AC-15,* 443–444.

Whiting, H. T. A. (Ed., 1984). *Human motor actions: Bernstein Revisited.* Amsterdam: North-Holland.

Submitted January 15, 1991
Revised May 1, 1991

Reprinted from the **Journal of Motor Behavior**, published by HELDREF PUBLICATIONS, 1319 Eighteenth Street, NW, Washington, DC 20036-1802.