

## CHAPTER 1

### **Control Systems With A Priori Intentions Register Environmental Disturbances A Posteriori**

Wayne A. Hershberger

#### **The Purpose and Precis of this Paper**

Closed-loop control systems act to offset the sensory effects of environmental disturbances, what von Holst called *ex-afference* (von Holst & Mittelstaedt, 1950/1973; von Holst, 1954), even before they are able to assess the magnitude of those effects, with the magnitude of the environmental disturbance being reflected, a *posteriori*, in the system's output or *efference*. For example, the controlled room temperature of my home hovers around 72 degrees Fahrenheit (approximating a reference value set *a priori*), whereas the variable degree to which cold weather disturbs that controlled variable is reflected, a *posteriori*, in my fuel bill. In general, it may be said (using von Holst's terminology) that the negative feedback of closed-loop control systems minimizes *ex-afference*, while maximizing *re-afference*. In contrast, the feedback-negation mechanism in von Holst's *efference-copy* hypothesis was intended to do just the opposite—minimize *re-afference* and maximize *ex-afference*.

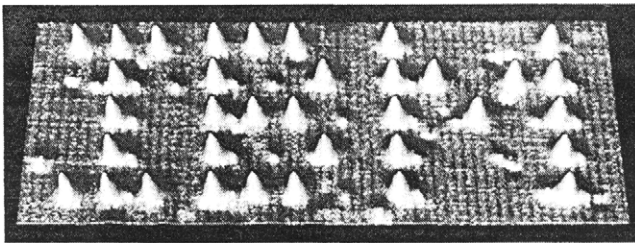
Paradoxically, von Holst's functional schemata (von Holst & Mittelstaedt, 1950/1973; von Holst, 1954) has repeatedly been mistaken for a closed-loop control system in which the *efference copy* serves as a reference value (e.g., see Gallistel, 1980, pp. 166-176; Hinde, 1970, p. 99). This confusion may be particularly widespread among those who are most familiar with the sort of research which von Holst's hypothesis spawned, including the developmental research on sensori-motor coordination (e.g., Held & Hein, 1963) and the research on perceptual adaptation to prismatic displacement (see, Rock, 1966; Welch, 1978). This sort of confusion is also fostered, perhaps, by the fact that an *efference copy* plays an important role (albeit not as a reference signal) in Robinson's (1975) widely accepted, closed-loop model of the oculomotor system. Such confusion can only impede conceptual progress. The purpose of the present chapter is to dispel this confusion by carefully contrasting the *a posteriori* *efference copy* of von Holst and Mittelstaedt's functional schemata with the *a priori* reference signals of closed-loop control systems, such as the living-systems

model currently being championed by Powers (1989) and others (Hershberger, 1989).

### The Re-Afference Principle

A century ago, John Dewey (1896) criticized the reflex arc concept in behavioristic psychology for blinking at the **concurrent** sensory feedback that accompanies an organism's movements, reflexive or otherwise. Dewey illustrated his argument with an optical-ocular example: When looking at a candle flame, any movement of the eye in its orbit is accompanied by a corresponding movement of the flame's image across the retina. If motion of the eye in its orbit is called a motor response (looking) and motion of the retinal image is called a sensory stimulus (seeing), the reflex arc concept, in asserting that looking and seeing are causally related, implies that one of the two comes first, either seeing or looking. But in fact, the corresponding motions are contemporaneous, comprising what Dewey called a sensori-motor coordination.

Dewey would have been fascinated with the picture reproduced in Figure 1 showing 35 Xenon atoms arranged to spell out the IBM logo, because the scanning tunneling microscope (STM) used to "see" the atoms capitalizes upon a sensori-motor coordination of the type Dewey was talking about in his classic critique. (The STM was used to position as well as picture the atoms; Eigler & Schweizer, 1990).



**Figure 1.** Seeing atoms by means of a closed-loop control system, the scanning tunneling microscope (STM). Donald M. Eigler and Erhard K. Schweizer of IBM's Almaden Research Center used an STM to generate, as well as to picture, their employer's logo. (From C. S. Powell, 1990; reproduced with permission)

The STM operates by passing an electrified, ultrafine needle across the surface of the sample being examined. An electrical current "tunneling" across

the gap between the tip of the needle and the surface of the sample is continuously monitored and controlled by simultaneously varying the depth of the needle. A sensori-motor, current-depth, coordination obtains, and because the current is controlled (i.e., kept equal to a reference value set a priori), the depth of the needle reflects the depth of the surface being scanned (a posteriori). Voilà! The STM pictures the surface of the sample by combining a record of all the various positions of the needle's tip.

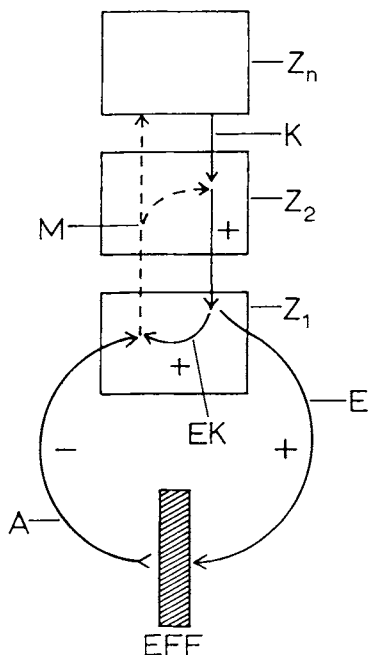
The STM is a closed-loop control system that uses negative feedback to control the value of the sensed current. Although closed-loop-control theory was not well developed until the 1930's (Black, 1934), the general idea had been floating around for centuries (Mayr, 1970), so Dewey, arguably, was suggesting that psychology should adopt a closed-loop-control paradigm in place of the reflex-arc paradigm. Although a number of twentieth century psychologists are currently championing the control theoretic paradigm (e.g., see Bourbon, in press; Hershberger, 1989; Marken, 1990; Powers, 1989), Dewey's critique of the reflex arc concept has had remarkably little effect upon the contemporary Zeitgeist. That is, while the theoretical hegemony of behaviorism has waxed and then waned during the twentieth century, the reflex arc concept has steadily grown to include conditional reflexes (Pavlov, 1927), instrumental responses (Thorndike, 1932), respondent and operant behaviors (Catania & Harnad, 1988; Skinner, 1938), and the like. As a consequence, today, as in Dewey's day, psychologists still blink at the fact of concurrent sensory feedback. Ask a contemporary psychologist about the significance of concurrent sensory feedback and you will likely get a lecture about immediate reinforcement—that is to say, the psychologist will not know what you are talking about. The phenomenon has been overlooked by so many for so long that most contemporary psychologists feel free to dismiss the phenomenon without ever addressing its theoretical implications.

There is one notable exception, however: In 1950, the German neuroethologist von Holst published a widely cited theoretical paper (von Holst & Mittelstaedt, 1950/1973) in which he identified the phenomenon and gave it a name (also see von Holst, 1954). Calling sensory feedback "re-afference" (afference and efference are physiological terms for neural input and output, respectively), von Holst acknowledged the factual nature of concurrent sensory feedback and called it *the re-afference principle*. Further, having identified the re-afference principle, he addressed some of its theoretical implications, reporting the results of ingenious experimental tests, and elaborating a theoretical model or functional schemata (see Figure 2), known as the efference-copy hypothesis.

Von Holst's efference-copy hypothesis provides a potential answer to Dewey's critique (i.e., the hypothesis attempts to fill a theoretical void). As such, it is important for two reasons: (a) it provides the *only* answer that has been offered during the twentieth century, apart from control theory, and (b) it provides

the *wrong* answer. It is worthwhile, therefore, to consider von Holst's efference-copy hypothesis in some detail not only to demonstrate the theory's shortcomings but also to explicate the nature of the phenomenon that the theory was designed to address: the re-afference principle.

**Figure 2.** A diagram illustrating Von Holst's theoretical model. Rectangles  $Z_1$  to  $Z_n$  represent different neural centers. Center  $Z_1$  has motor and sensory connections with an effector **EFF**. Any command from  $Z_n$  produces an efferent stream of impulses, **E**, from  $Z_1$  which leaves an efference copy, **EC**, in  $Z_1$ . **E** evokes a corresponding re-afference, **A** (it should be labeled  $A_{re}$ ). When **A** (it should be labeled  $A_{re}$ ) and **EC** combine in  $Z_1$  they are supposed to cancel each other. (From von Holst, E., & Mittelstaedt, H., 1950/1973; reproduced with permission.)



## Efference-Copy Hypothesis

According to von Holst, afference (**A**) comprises two types of receptor input, re-afference ( $A_{re}$ ) and ex-afference ( $A_{ex}$ ), with  $A = A_{re} + A_{ex}$ .

Re-afference is the necessary afferent reflexion caused by every motor impulse; ex-afference is independent of motor impulses....If I shake the branch of a tree, various receptors of my skin and joints produce a re-afference, but if I place my hand on a branch shaken by the wind, the stimuli of the same receptors produce an ex-afference....The same receptor can serve both the re- and the ex-afference. The CNS [central nervous system], must, however,

possess the ability to distinguish one from the other. This distinction is indispensable for every organism, since it must correctly perceive its environment at rest and in movement, and stimuli resulting from its own movements must not be interpreted as movements of the environment. (von Holst, 1954, p. 90)

Von Holst reasoned that if feedback is the problem, feedback negation must be the solution. Therefore, von Holst proposed that the re-afference attending any movement is negated by a putative efference copy (EC) comprising a corollary discharge of the neural efference that initiated the movement in the first place (cf. Sperry, 1950). In von Holst's (1954) words, "the efference leaves an 'image' of itself somewhere in the CNS, to which the re-afference of this movement compares as the negative of a photograph compares to its print; so that, when superimposed, the image disappears" (p. 91). In this manner, the EC is supposed to cancel the  $A_{re}$  portion of  $A$ , yielding pure  $A_{ex}$  as a remainder to mediate veridical perceptions of environmental motions. Algebraically, the efference-copy hypothesis may be stated as follows:

$$A = A_{re} + A_{ex} \quad (1)$$

$$EC = A_{re} \quad (2)$$

$$A_{ex} = A - EC \quad (3)$$

Von Holst's efference-copy hypothesis is his proposed solution to a problem posed by the re-afference principle. Von Holst used the optokinetic reflex of a fly to illustrate both the problem and his proposed solution. Suppose that a fly is standing on a horizontal surface surrounded by a vertical cylinder lined with vertical black and white stripes. Further, suppose that, initially, both the fly and the cylinder are stationary. Because the fly is stationary,  $A_{re}$  is nil, and because the cylinder is stationary,  $A_{ex}$  is nil, and because both of these are nil visual afference is nil ( $A = A_{ex} + A_{re} = 0$ ); consequently, there is no reflexive motion. Now, if  $A_{ex}$  is introduced by rotating the cylinder clockwise at  $1^\circ/s$ , the fly will reflexively rotate in the same direction (cw) at about  $1^\circ/s$ . This well-known optokinetic reflex is said to be elicited by the visual afference,  $A$ , produced by the translation of the optic array across the fly's eyes, in this case, from the left to the right eye at  $1^\circ/s$ . Because this afference ( $A = L \rightarrow R @ 1^\circ/s$ ) was caused by movement of the environment, this  $A$  is all  $A_{ex}$ .

However, suppose that the cylinder is stationary ( $A_{ex} = 0$ ) and the fly spontaneously rotates *counterclockwise* at  $1^\circ/s$ . In this case,  $A$  is all  $A_{re}$ , but otherwise the same as before ( $A = L \rightarrow R @ 1^\circ/s$ ). If the visual afference is the

same in the two cases, and this afference elicits the optokinetic reflex in the former case, why does it not do so in the latter case? Why does the optokinetic reflex not return the fly to its original position whenever a turning motion is initiated? The re-afference principle implies that the fly should be virtually immobilized (rendered virtually catatonic) by its optokinetic reflex. It is not. Why not? Von Holst argued that it is because every re-afference is canceled by an efference copy. That is, if the neural efference responsible for the spontaneous rotation of the fly leaves a copy of itself in the CNS, and this efference copy is equal to the re-afference of that motion, then the contaminating re-afference can be subtracted from afference simply by subtracting the corresponding efference copy from the afference. This is what he hypothesized. *Ideally*, this would yield an afference which is pure ex-afference.

### Mittelstaedt's Experiment

An alternate hypothesis is that reflexes are simply disabled during spontaneously emitted movements (operants). To test this alternate hypothesis, von Holst's colleague Mittelstaedt rotated a fly's head  $180^\circ$  about the long axis of its body and fixed it to the thorax upside down so that the two eyes were interchanged. In this case, if  $\mathbf{A}_{ex}$  is introduced by rotating the cylinder clockwise at  $1^\circ/s$ , the afference ( $\mathbf{A} = R \rightarrow L @ 1^\circ/s$ ) is reversed from normal, and the fly reflexively rotates in a counterclockwise direction. More importantly, if the cylinder is stationary ( $\mathbf{A}_{ex} = 0$ ) and the fly begins to rotate counterclockwise at  $1^\circ/s$ , the ccw rotation continues indefinitely as a forced circus movement. This refutes the hypothesis that the reflex is simply disabled during spontaneously emitted movements. Further, according to von Holst, the forced circus movements are driven by uncanceled re-afference, as would be expected from his hypothesis. The explanation is as follows, beginning with a semantic (mathematical) simplification: The *polarity* of the afference,  $L \rightarrow R$  or  $R \rightarrow L$ , is simplified by reducing it to a sign (+ or -), say, with  $L \rightarrow R$  being positive and  $R \rightarrow L$  being negative. Then, if a fly's head is inverted and the fly spontaneously rotates ccw at  $1^\circ/s$ , the re-afference ( $\mathbf{A}_{re} = \mathbf{A} = R \rightarrow L @ 1^\circ/s$ ) would be  $-1^\circ/s$  whereas the efference copy, still corresponding to normal re-afference ( $\mathbf{A}_{re} = \mathbf{A} = L \rightarrow R @ 1^\circ/s$ ), would be  $+1^\circ/s$  (i.e.,  $\mathbf{EC} = +1^\circ/s$ ). Finally, subtracting the latter positive value ( $\mathbf{EC} = +1^\circ/s$ ) from the former negative value ( $\mathbf{A} = -1^\circ/s$ ), as in Equation 3 above, yields an even larger negative value as a remainder ( $\mathbf{A} = -2^\circ/s$ ) to drive the optokinetic reflex in a vicious cycle (ccw in the present example). This vicious cycle is dysfunctional.

When Mittelstaedt returned the fly's head to its normal upright position the fly again behaved normally, moving freely in a stationary cylinder. This is possible, according to the theory, because  $\mathbf{EC}$  and  $\mathbf{A}_{re}$  now share the same polarity (+ or -). For example, using the sign convention adopted above,

whenever a normal fly spontaneously rotates ccw at  $1^\circ/\text{s}$ ,  $\text{EC}$  and  $\mathbf{A}_{\text{re}}$  ( $\text{L} \rightarrow \text{R}$  @  $1^\circ/\text{s}$ ) are both positive ( $+1^\circ/\text{s}$ ), and, therefore, according to Equation 3, the contaminating  $\mathbf{A}_{\text{re}}$  is canceled by the  $\text{EC}$ —providing that  $\text{EC}$  and  $\mathbf{A}_{\text{re}}$  have the same numerical value. If the values of these two variables are not identical, Equation 3 will yield a non-zero residual even when  $\mathbf{A}_{\text{ex}} = 0$ . According to the theory, this residual is always regarded by the fly as ex-afference, whatever its source. Thus, if the cylinder is actually stationary ( $\mathbf{A} = 0$ ), but  $\text{EC}$  and  $\mathbf{A}_{\text{re}}$  are unequal, Equation 3 will yield a non-zero residual that will be regarded by the fly as  $\mathbf{A}_{\text{ex}}$ . That is, the residual (the uncanceled  $\mathbf{A}_{\text{re}}$ ) will give the fly the impression that the cylinder is moving, even though it is not, at an apparent angular velocity equal to the residual.

### Held's Putative Correlation Store

Given that  $\text{EC}$  and  $\mathbf{A}_{\text{re}}$  are functions of the same efference,  $\text{EC}$  will equal  $\mathbf{A}_{\text{re}}$  providing that the two *functions* are the same—which is what von Holst assumed. However, Held (1961) noted that as circumstances change (e.g., muscles become fatigued) the re-afferent consequence of a given neural efference,  $\mathbf{E}$ , will change. That is, the function relating  $\mathbf{A}_{\text{re}}$  to  $\mathbf{E}$ , call it  $\mathbf{f}$ , will change from time to time. If the function relating  $\text{EC}$  to  $\mathbf{E}$ , call it  $\mathbf{f}'$ , does not change simultaneously with every change in  $\mathbf{f}$ , then  $\mathbf{A}_{\text{re}} = \mathbf{f}(\mathbf{E})$  will not equal  $\text{EC} = \mathbf{f}'(\mathbf{E})$ , except by chance. Held reasoned that if von Holst's functional schemata is to work properly (i.e., distinguish ex- from re-afference),  $\mathbf{f}'$  must be an empirical estimate of  $\mathbf{f}$  which is continually being updated. Held postulated a putative mechanism he called the "correlation store" which he supposed could accurately estimate  $\mathbf{f}$ , the "correlation" between  $\mathbf{E}$  and  $\mathbf{A}_{\text{re}}$ , by continuously recalculating the "correlation" between  $\mathbf{E}$  and  $\mathbf{A}$ . Of course, Held's supposition is gratuitous; the correlation between  $\mathbf{A}$  and  $\mathbf{E}$  is clearly *not* the same as the correlation between  $\mathbf{A}_{\text{re}}$  and  $\mathbf{E}$ . Held's putative mechanism (correlation store) can not provide von Holst's functional schemata with the information it needs (accurate estimates of  $\mathbf{f}$ ). *Therefore, von Holst's functional schemata will not work properly to distinguish ex- from re-afference.*

Held's hypothesis (i.e., the putative correlation store) is often cited as a theoretical rationale for experiments concerned with coordinated movements, such as pointing the hand or paw in the same direction as the eyes (e.g., see Held & Hein, 1963; Rock, 1966, Welch, 1978). However, the type of "coordination" that concerns us at present is not movement-movement coordination, but Dewey's sensori-motor coordination—what von Holst called the re-afference principle. The re-afference principle appears to pose a critical perceptual problem for the CNS. According to von Holst, re-afference is potentially worse than useless because, if it can not be distinguished from ex-afference, it can only

generate perceptual illusions. It was von Holst's hypothesis that this critical "discrimination" is effected by means of an efference copy, but this efference-copy hypothesis presupposes a mechanism that is able to keep  $f'$  identical to  $f$ . On this count, Held's hypothesized correlation store is of no help—it does not have access to the necessary data (pure re-afference). To claim that the correlation store does have access to pure re-afference is to imply that other parts of the CNS have access to pure re-afference as well and, thus, the CNS has no need of von Holst's efference copy in the first place—the CNS could always determine the value of  $A_{ex}$  simply by subtracting  $A_{re}$  from  $A$ . In other words, if Held's correlation store could do what it needs to do, it would not need to do what it could do. As a theoretical addition to von Holst's efference-copy hypothesis, Held's correlation store has no utility.

The problem posed by the re-afference principle is not unique to insects (e.g., see Sperry, 1951). Smith and Molitor (1969) and Yarbus (1962) found that reversing the polarity of visual afference is just as debilitating in humans as it is in insects: When their experimental participants wore contact lenses fitted with a prism (Smith) or a mirror (Yarbus) that reversed left and right, the participants' intentional eye movements elicited interminable forced nystagmus.

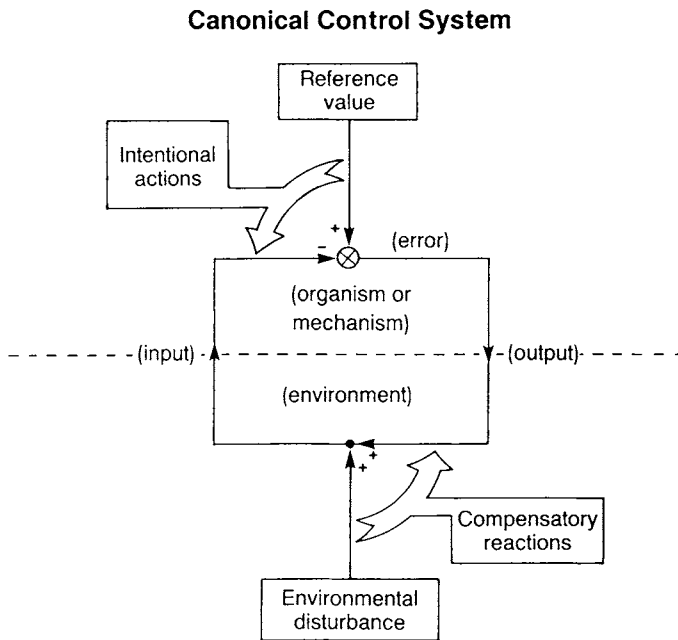
### **Closed-Loop-Control Hypothesis**

The alternative to von Holst's efference-copy hypothesis is the closed-loop-control hypothesis, currently championed by Powers (1973, 1989) and by others (e.g., Bourbon, in press; Cziko, 1992; Hershberger, 1986, 1987a, 1987b, 1988, 1989; Marken, 1990, 1992; McClelland, 1994). Figure 3 is a canonical representation of closed-loop control. Any system that controls the value of a single variable by means of negative feedback may, in principle, be reduced to the canonical loop illustrated in Figure 3. That is, a schematic diagram of a very complex closed-loop control system may be reduced to the simple canonical form depicted in Figure 3. Note that Figure 3 looks a bit like Figure 1, von Holst's functional schemata. The major feature of both figures is a closed loop. This loop represents the re-afference principle, which both theories readily acknowledge. However, the similarity ends there.

A control system monitors the value of the variable being controlled, compares that value with a reference value, and uses the discrepancy to drive output in a direction that reduces the error. (Several of these functions may be carried out by a single component, so when analyzing a system one should not always expect to find as many components as functions.) Because the output is error driven and error reducing, the feedback is said to be negative; the error signal negates itself. Further, because it is an error signal that is being negated, negative feedback is good. Conversely, positive feedback (the proverbial vicious cycle) is bad. (When Mittelstaedt rotated a fly's head 180° he reversed the



polarity of visual feedback, replacing good negative feedback with bad positive feedback. The notion in psychology that vicious cycles of self-sustaining activity may sometimes prove useful is called the “circular responses hypothesis”; see Dennis, 1954.)



**Figure 3.** A canonical control loop mapped onto the interface (dashed line) between an organism (or mechanism) and its environment. (From Hershberger, 1989; reproduced with permission.)

In order to control an environmental variable, a control system requires exteroceptors that can monitor the controlled variable and effectors which can influence the controlled variable. That is, the control loop must extend into the control mechanism’s (or organism’s) environment. Accordingly, the canonical loop in Figure 3 is mapped onto an interface between a control mechanism (or organism) and its environment. Everything above the dotted line is part of the mechanism (or organism), and everything below the dotted line is part of the environment. Although the mechanism (or organism) has one input (the controlled variable), the control loop has two inputs, the reference value and the environmental disturbance. The reference value represents the mechanism’s (or organism’s) intended input. The environmental disturbance comprises all the

environmental factors which potentially disturb the mechanism's (or organism's) input.

### **An Example of Closed-Loop Control**

The cruise control on modern automobiles provides a familiar example of closed-loop control. The automobile's speedometer is a receptor sensing road speed, and the automobile's engine is an effector influencing road speed. Road speed is also influenced (disturbed) by the slope of the roadway, with an inclining roadway tending to decrease road speed and a declining roadway tending to increase road speed. One "sets" the reference value of the cruise control (the intended road speed) by bringing the speedometer reading to that value and pressing a button. Thereafter, the cruise-control takes over—controlling road speed. The cruise-control system compares the sensed road speed with the intended value and automatically increases the flow of fuel to the engine whenever sensed speed is too low, and decreases the flow of fuel whenever sensed speed is too high. The output of the engine (torque) is, thus, error driven and error reducing. Note that this negative feedback loop controls the speed of the automobile but not the torque of the engine. Having set the reference value (intended cruising speed) one can predict the speedometer reading but not the torque of the engine. The torque of the engine is altered by the slope of the terrain, with engine torque increasing during ascents and decreasing during descents. A coupling of two types of behavior (controlled road speed and elicited engine torque) is apparent, represented in Figure 3 by the two large blocked arrows labeled "Intentional actions" and "Compensatory reactions." In this example the intentional action is the road speed and the compensatory reaction is the engine torque: The road speed is determined by the reference value (an intention) while variations in engine torque are determined (elicited) by variations in the terrain.

### **Two Types of Behavior: Intentional Actions and Compensatory Reactions**

The two blocked arrows in Figure 3 represent lineal cause and effect relationships that emerge from the underlying circular feedback process. The two blocked arrows represent emergent properties of the whole loop. They are *not* integral parts of the loop itself. Note that the two blocked arrows point in a counterclockwise direction, which is opposite to the clockwise direction of the feedback loop itself. *Closed-loop control systems do not control their input by controlling their output. Nor do environmental disturbances elicit compensatory output by being sensed.*

The two types of behavior, the intentional actions and the compensatory reactions, are synergistically coupled. That is, although intentional actions and

compensatory reactions are mutually exclusive *types* of action they are not mutually exclusive actions. On the contrary, they are always found to go hand in hand in any system that controls its own input. For example, the flight path of an airplane is the pilot's (or autopilot's) doing only to the degree that the pilot's (or autopilot's) reactions automatically offset any would-be aerodynamic disturbances to the intended flight path. Otherwise, he, she (or it) is merely along for the ride.

The intentional actions and automatic reactions represented by the blocked arrows in Figure 3 are both entirely dependent upon, but emergent from, the underlying negative feedback loop. Because negative feedback is the re-afference principle put to good use, control theorists view the re-afference principle in an altogether different light than von Holst's. Control theorists view re-afference as essential to closed-loop control and, thus, as helpful. Von Holst viewed re-afference as a contamination of afference and, thus, as harmful. Accordingly, von Holst's efference-copy hypothesis deals with the re-afference principle in a manner that is altogether different from that of the control-theoretic model. In fact, the two models are functionally antithetical: Whereas von Holst's ideal functional schemata is supposed to rid *afference of all re-afference* by means of feedback negation, an ideal closed-loop control system *rids afference of all ex-afference* by means of negative feedback (e.g., in the cruise-control example, the speed of the vehicle does not vary with the terrain). The antithesis is twofold: (a) feedback negation versus negative feedback, and (b) ridding afference of re-afference versus ridding afference of ex-afference.

### Von Holst's Mistake

Although von Holst was aware of negative feedback and its importance in control theory (what he referred to as technological cybernetics) he

emphasized that this "negative feedback" is not a necessary component of the reafference principle and that it should not be confused with the latter! The decisive point in the principle is the mechanism distinguishing reafference and exafference [i.e., the putative efference copy]. This distinction plays no part in cybernetic technology. (von Holst & Mittelstaedt, 1950/1973, footnote 7)

Von Holst was half right (which is to say that he was wrong). Control systems do not need to distinguish ex- from re-afference in order to control the value of afference, but to the degree that the afference is controlled, it is free of ex-afference. So, in a closed-loop control system there is no need for an efference-copy mechanism to distinguish ex- and re-afference in the first place. When Reichardt and Poggio (1976) quantitatively analyzed the visual-orientation

behavior of flies during flight, they found no evidence of von Holst & Mittelstaedt's putative mechanism for distinguishing "between self movement and object movement" (p. 358). Further, even von Holst's colleague, Mittelstaedt, has had second thoughts (MacKay & Mittelstaedt, 1974).

### **Efference as a Reflection of Ex-Afference**

However, this is not to say that there is no purpose to which sensed efference in the form of an efference copy might be put. On the contrary. When closed-loop control systems control their input (afference), environmental disturbances (would-be ex-afference) are mirrored in the system's output (efference), because in order to control the value of the input, every environmental disturbance (every would-be ex-afference) would have to be offset by compensatory output (efference). Consequently, in a closed-loop control system, an efference copy would mirror would-be ex-afference.

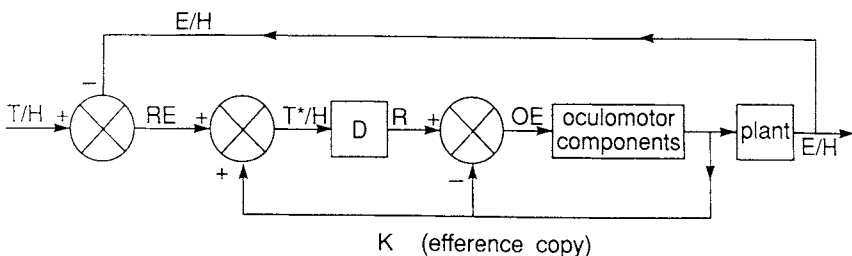
The scanning tunneling microscope, mentioned above, capitalizes on this fact. The STM controls the electric current "tunneling" across the gap between the tip of a scanning needle and a scanned surface by continuously adjusting the depth of the needle. Taken collectively, the myriad positions of the needle's tip represent the shape of the surface being scanned; thus, a corollary discharge (i.e., efference copy) of the signal that determines the depth of the needle, taken together with the signals that generate the two-dimensional scan, provide all the information necessary to picture the three-dimensional shape of the scanned surface (would-be ex-afference). Rather than corresponding to re-afference, the efference copy corresponds to ex-afference, exactly opposite to what von Holst supposed.

A comparable example from psychology is "seeing" the visual direction of a visible target by looking at it. For instance, when one is watching an aircraft fly overhead (i.e., controlling the retinal locus and motion of its image—keeping the image nearly stationary on the fovea) neither the visual direction of the aircraft, nor the aircraft's motion is represented on the retina. Rather, they are represented in extraretinal, oculomotor signals corresponding to the various orientations and movements of the eyes (e.g., see Hershberger, 1987b). The position of the image on the retina is relatively fixed, and simply reflects the intent to watch the object. That is, the reference values for the retinal slip and the retinal eccentricity of the retinal image are both set to zero, a priori. The motion of the aircraft (i.e., its changing visual direction) is registered in the nervous system a posteriori in terms of the oculomotor efference required to keep the aircraft's image on the fovea. Accordingly, the extra-retinal signal representing sensed direction of gaze is generally acknowledged to be an efference copy comprising corollary discharges from premotor integrator neurons in the brainstem (Robinson, 1975).

## Robinson's (1975) Model of the Oculomotor System

Robinson's closed-loop model of the oculomotor system, depicted in Figure 4, merits special attention because it is a closed-loop control system that happens to utilize an efference copy as the controlled variable.

Robinson's closed-loop model of the oculomotor system utilizes two separate indices of the variable being controlled (as do all closed-loop control systems): a feedback signal, representing the current value of the controlled variable, and a reference signal, specifying the intended value of the feedback signal. The feedback signal in Robinson's model is an efference copy,  $K$ , comprising corollary discharges from premotor integrator neurons in the brainstem; because the eyeball is not subjected to environmental disturbances (i.e., no variable loads), the eyeball's orientation is determined by this efference. The reference signal in Robinson's model is a centralized command signal that specifies the intended value of the feedback signal; that is, it is a reference signal ( $R$ ) representing intended eye position. The centralized reference signal ( $R$ ) is a *controlling* command signal. The efference copy ( $K$ ) is a *controlled* command signal. Because both are command signals and one is a reference signal, it is tempting to suppose that both are reference signals, but this is not the case. There



**Figure 4.** A simplified version of Robinson's (1975) model of the saccadic oculomotor system.  $T/H$ , target position relative to head;  $E/H$ , orientation of the eyes relative to the head;  $RE$ , retinal error (retinal eccentricity of the target's image);  $D$ , delay;  $T^*/H$ , neural estimate of target position relative to the head;  $K$  (efference copy), sensed orientation of the eyes in the head;  $R$  (reference signal), intended eye orientation;  $OE$  (oculomotor error), the difference between the intended and sensed orientation of the eyes in the head ( $R - EC$ ).

is no controlled variable in Robinson's oculomotor model that is driven into correspondence with the model's efference copy ( $K$ ), so  $K$  is not a reference signal. It is also tempting (even more so) to suppose that  $R$  and  $K$  are both efference copies, but that is not the case either. The reference signal,  $R$ , is a central command which is not sent to the extraocular muscles. Therefore,  $R$  is not efference, let alone an a posteriori copy of efference. However, inasmuch as any reference signal in any control system is an a priori representation of the value of the variable being controlled, it could be said to be a copy, a priori, of the value of the controlled variable. Thus, the value of  $R$  in Robinson's model may be said to be a copy (of one sort) of a copy (of another sort). That is,  $R$  is a copy, a priori, of  $K$ , while  $K$  is a copy, a posteriori, of efference.

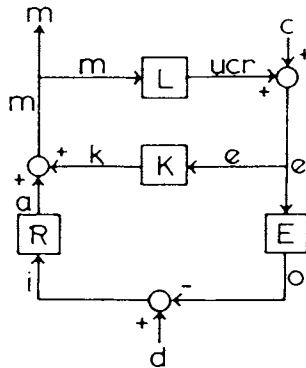
### Reference Signals Are A Priori Copies

Because the reference signal of a closed-loop control system is an a priori copy of the controlled variable, and because the controlled variable is typically monitored, or sensed, by a receptor, the reference signal is typically an a priori copy of sensed input (afference), and, thus, could be called an *afference copy*, as I have done from time to time (Hershberger, 1978, 1987b; Hershberger & Jordan, 1992, 1996). However, the expression *afference copy* is inappropriate when the controlled variable is a corollary discharge of efference, as is the case in Robinson's model of the oculomotor system. *Therefore, whenever a reference signal is called a copy, it should be called an a priori copy to distinguish it from von Holst's a posteriori, efference copy.*

### The A Priori Copy in von Holst's Functional Schemata

By analyzing von Holst's functional schemata (von Holst & Mittelstaedt, 1950/1973) using functional calculus (Di Stefano, Stubberud, & Williams, 1967), it is possible to determine what system variable, if any, is serving as an a priori copy (reference signal) in von Holst's model.

The schematic diagram shown in Figure 5 is equivalent to one that Mittelstaedt (1958) used to illustrate their (von Holst & Mittelstaedt, 1950/1973) original hypothesis, but in control theoretic terms. The lines in Figure 5 represent variables and the blocks represent transfer functions relating one variable to another. The transfer functions and variables are labeled with upper-case and lower-case letters respectively. The letters are mnemonic; the key is shown. The arrows indicate the direction of influence of one variable upon another. The branching points indicate that one variable influences two others. The open circles represent summing points where one variable is either added to or subtracted from the other as indicated by a plus or minus sign.

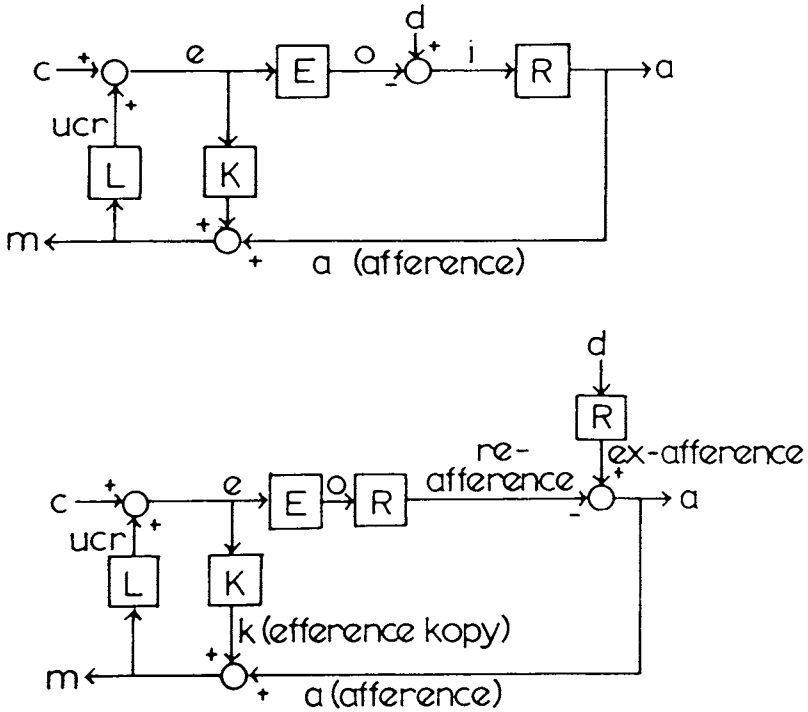


VARIABLES	TRANSFER FUNCTIONS
c: central command	E: Effector I/O
e: neural efference	R: Receptor I/O
o: effector output	K: Copy function
d: environmental disturbances	L: Link (reflexive)
i: receptor input	
a: neural afference	
k: efference copy	
m: meldung	
ucr: unconditional reflex	

**Figure 5.** A detailed version of von Holst's theoretical model (von Holst & Mittelstaedt, 1950/1973) adapted from a control-theoretic rendition by Mittelstaedt (1958).

The diagram in Figure 5 is functionally equivalent to the one in Figure 2 except for three features: (a) environmental disturbances are represented, (b) a receptor is included, and (c) the point at which the polarity of feedback is inverted is shifted from the sensori-motor area to the environment. The latter difference is trivial, and the former two features correct errors of omission in Figure 2. Using functional calculus, von Holst and Mittelstaedt's original functional schemata, represented in Figure 5, may be reduced step by step to its canonical form. This reduction is illustrated in Figures 6 through 9.

The diagrams in Figure 6 distinguish ex- and re-afference. The upper diagram in Figure 6 involves no reduction; lines are lengthened, shortened, and bent in order to highlight the variable labeled afference. In the lower drawing, the receptor transfer function,  $R$ , is moved in front of the summing point. This is done simply by using the distributive rule of multiplication over subtraction. That is, since  $a = R \times i$ , and  $i = d - o$ , it follows that  $a = R(d - o)$ ; or, using the distributive rule,  $a = (R \times d) - (R \times o)$ . The product of  $R$  times  $d$  is the intervening variable von Holst labeled ex-afference. The product of  $R$  times  $o$  is



**Figure 6.** A reduction of von Holst’s model illustrating the intervening variables, ex-afference and re-afference.

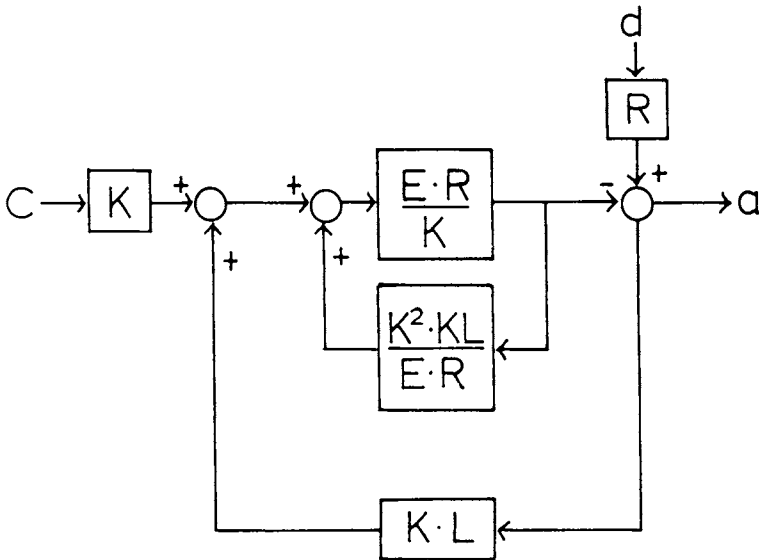
the intervening variable von Holst labeled re-afference. AffERENCE is equal to the sum of the two, with re-affERENCE having a negative sign.

It is apparent in Figure 6 that the feedback loops have no influence whatsoever upon the relationship obtaining between re-affERENCE and the efference copy. If the product  $(E \times R)$  equals  $K$ , then the values of the re-affERENCE and the efference copy will correspond exactly for any value of efference. However, if  $(E \times R)$  does not equal  $K$ , then the system can not and does not drive re-affERENCE into correspondence with the efference copy, neither does it drive the value of affERENCE into correspondence with the efference copy, as becomes apparent with a further reduction of the schematic diagram.

The diagram in Figure 7 is the result of moving and combining transfer functions in the manner just illustrated. Two loops are now apparent: an inner, positive-feedback loop, and an outer negative-feedback loop. By replacing the inner loop with its closed-loop transfer function, one gets the feedback system in the canonical form shown in Figure 8.

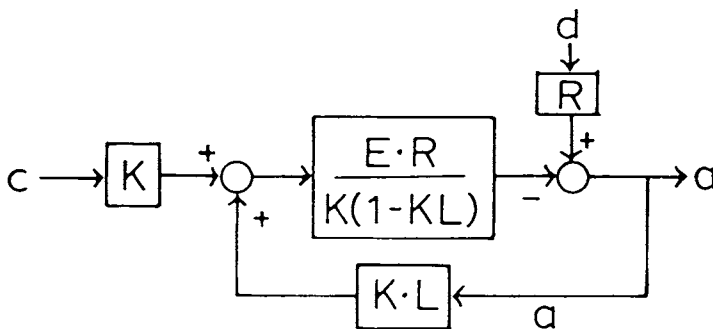
Finally, by removing the block from the feedback path of the canonical loop, one gets the unity loop shown in Figure 9. Because of the minus sign, the



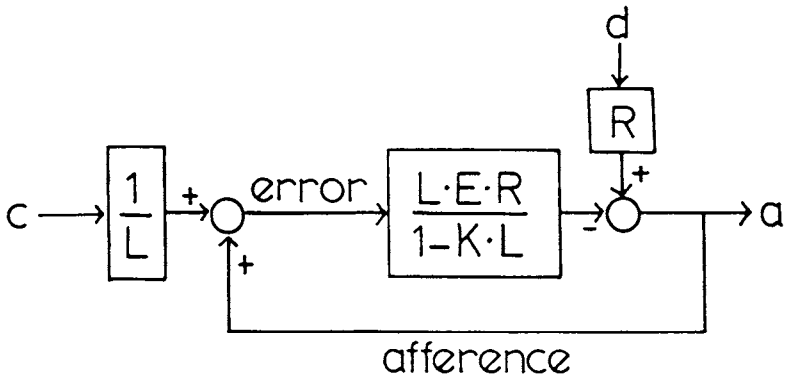


**Figure 7.** A further reduction of von Holst's model showing a positive feedback loop.

loop involves negative feedback. Assuming that the system is stable, one can see that the value of the afference at steady state (i.e., when error is zero) must be  $-c/L$ . Thus, assuming that the system is stable, the system variable which serves as an a priori copy of the controlled afference is  $-c/L$ . That is, the functional schemata's reference signal, set point, Sollwert, should-be value, or the like is expressed in terms of the central command,  $c$ , rather than the efference,  $e$ , or the efference copy,  $k$ .



**Figure 8.** A diagram of von Holst's model reduced to its canonical form.



**Figure 9.** A diagram of von Holst's model reduced to a unity loop. Assuming that the system is stable, the controlled efference must be equal to  $-c/L$  at steady state, when error is zero. Thus, the model's reference value is expressed in terms of the central command,  $c$ , rather than the efference,  $e$ , or the efference copy,  $k$ .

## Conclusions

The re-efference principle is something psychology must address. There are two opposing views, one leading to von Holst's efference-copy hypothesis, the other leading to the closed-loop-control hypothesis. The efference-copy hypothesis is fatally flawed. The closed-loop-control hypothesis is viable. A reference signal, a significant feature of any closed-loop-control system, can even be found in von Holst's functional schemata. However, this a priori copy is not to be confused with the sort of a posteriori copy von Holst called an efference copy. If von Holst's efference copy is to have any theoretical utility it is simply as a form of sensed efference or innervation sensation (for a history of this hoary idea, see Scheerer, 1987). When closed-loop control systems control their input, environmental disturbances (would-be ex-efference) are mirrored in the system's output (efference), because in order to control the value of the input, every environmental disturbance (every would-be ex-efference) would have to be offset by compensatory output (efference). Consequently, rather than corresponding to re-efference, von Holst's efference copy corresponds to ex-efference, exactly opposite to what he supposed.

## References

- Black, H. S. (1934). Stabilized feed-back amplifiers. *Electrical Engineering*, January, 114-120.

- Bourbon, W. T. (in press). Perceptual control theory. In H. L. Roitblat & J-A Meyer (Eds.), *Comparative approaches to cognitive science*. Cambridge, MA: MIT Press.
- Catania, A. C., & Harnad, S. (Eds.). (1988). *The selection of behavior*. New York: Cambridge University Press.
- Cziko, G. A. (1992). Purposeful behavior as the control of perception: Implications for educational research. *Educational Researcher*, 21, 10-18, 27.
- Dennis, W. (1954). A note on the circular response hypothesis. *Psychological Review*, 61, 334-338.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, 3, 359-370.
- Di Stefano, J. J., Stubberud, A. R., & Williams, I. J. (1967). *Feedback and control systems*. New York: McGraw-Hill.
- Eigler, D. M., & Schweizer, E. K. (1990). Positioning single atoms with a scanning tunneling microscope. *Nature*, 344, 524-526.
- Gallistel, C. R. (1980). *The organization of action*. New York: Erlbaum.
- Held, R. (1961). Exposure history as a factor in maintaining stability of perception and co-ordination. *Journal of Nervous and Mental Disease*, 132, 26-32.
- Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, 56, 872-876.
- Hershberger, W. A. (1978). Afference copy: The closed-loop analogue of von Holst's efference copy. *Cybernetics Forum*, 8, 97-102.
- Hershberger, W. A. (1986). An approach through the looking glass. *Animal Learning & Behavior*, 14, 443-451.
- Hershberger, W. A. (1987a). Of course there can be an empirical science of volitional action. *American Psychologist*, 42, 1032-1033.
- Hershberger, W. A. (1987b). Saccadic eye movements and the perception of visual direction. *Perception & Psychophysics*, 41, 35-44.
- Hershberger, W. A. (1988). Some overt behavior is neither elicited nor emitted. In A. C. Catania & S. Harnad (Eds.), *The selection of behavior* (pp. 107-109). New York: Cambridge University Press.
- Hershberger, W. A. (Ed.). (1989). *Volitional action: Conation and control*. Amsterdam: Elsevier/North-Holland
- Hershberger, W. A., & Jordan, J. S. (1992). Visual direction constancy: Perceiving the visual direction of perisaccadic flashes. In E. Chekaluk & K. R. Llewellyn (Eds.), *The role of eye movements in perceptual processes* (pp. 1-43). Amsterdam: Elsevier/North-Holland.
- Hershberger, W. A., & Jordan, J. S. (1996). The phantom array. *Behavioral & Brain Sciences*, 19(3), 552-553.

- Hinde, R. A. (1970). *Animal behavior* (2nd. ed.). New York: McGraw-Hill.
- MacKay, D. M., & Mittelstaedt, H. (1974). Visual stability and motor control (reafference revisited). In W. D. Keidel (Ed.), *Cybernetics and bionics* (pp. 71-80). Munich: Oldenbourg.
- Marken, R. S. (Ed.) (1990). Purposeful behavior: The control theory approach [Special issue]. *American Behavioral Scientist*, 34(1)
- Marken, R. S. (1992). *Mind readings: Experimental studies of purpose*. Chapel Hill NC: New View.
- Mayr, O. (1970). *The origins of feedback control*. Cambridge: MIT Press.
- McClelland, K. (1994). Perceptual control and social power. *Sociological Perspectives*, 37, 461-496.
- Mittelstaedt, H. (1958). The analysis of behavior in terms of control systems. In B. Schaffner (Ed.), *Group processes, Transaction of the Fifth Conference*. New York: Josiah Macy, Jr. Foundation.
- Pavlov, I. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London: Oxford University Press.
- Powell, C. S. (1990). Science writ small. *Scientific American*, 262, 26.
- Powers, W. T. (1973). *Behavior: The control of perception*. New York: Aldine de Gruyter.
- Powers, W. T. (1989). *Living control systems*. Gravel Switch, KY: CSG Books.
- Reichardt, W., Poggio, T. (1976). Visual control of orientation behavior in the fly: Part I. A quantitative analysis. *Quarterly Review of Biophysics*, 9, 311-375.
- Robinson, D. A. (1975). Oculomotor control signals. In G. Lennerstrand & P. Bach-y-Rita (Eds.), *Basic Mechanisms of ocular motility and their clinical implications* (pp. 337-374). New York: Pergamon Press.
- Rock, I. (1966). *The nature of perceptual adaptation*. New York: Basic Books.
- Scheerer, E. (1987). Muscle sense and innervation feelings: A chapter in the history of perception and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on Perception and Action* (pp. 171-194). Hillsdale, N.J.: Lawrence Erlbaum.
- Skinner, B. F. (1938), *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual neural inversion. *Journal of Comparative & Physiological Psychology*, 43, 482-489.
- Smith, K. U., & Molitor, K. (1969). Adaptation to reversal of retinal feedback of eye movements. *Journal of Motor Behavior*, 1, 69-87.
- Sperry, R. W. (1951). Mechanisms of neural maturation. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley.
- Thorndike, E. L. (1932). *The fundamentals of learning*. New York: Teachers College Press.

- von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behavior*, 2, 89-94.
- von Holst, E., & Mittelstaedt, H. (1973). The reafference principle. In R. D. Martin (Ed. and Trans.), *The behavioral physiology of animals and man: Selected papers of E. von Holst* (Vol. 1). Coral Gables: University of Miami Press. (Original work published 1950)
- Welch, R. B. (1978). *Perceptual modification*. New York: Academic Press.
- Yarbus, A. L. (1962). *Eye movements and vision*. (B. Haigh, Trans.) New York: Plenum Press.