

Feedback Interventions: Toward the Understanding of a Double-Edged Sword

Avraham N. Kluger and Angelo DeNisi¹

School of Business Administration, The Hebrew University of Jerusalem, Jerusalem, Israel (A.N.K.), and Department of Management, Texas A&M University, College Station, Texas (A.D.)

Feedback intervention (FI), that is, providing people with some information regarding their task performance, is one of the mostly widely applied psychological interventions. Yet there is a growing body of evidence that such interventions yield highly variable effects on performance (Ilgen, Fisher, & Taylor, 1979; Kluger & DeNisi, 1996; Latham & Locke, 1991; Salmon, Schmidt, & Walter, 1984). Indeed, in a meta-analysis, we found that although FIs improve performance on average, they reduce performance in more than one third of the cases (Kluger & DeNisi, 1996; see Fig. 1). The latter fact is contrary to the common belief that FIs most often improve performance. Furthermore, we (Kluger & DeNisi, 1996) found no evidence that information about failure (negative FIs) and information about success (positive FIs) have differential effects, on average, on performance. In summary, the data suggest that, at least under certain circumstances, FIs can impair performance and that the processes through which FIs affect performance require more than simple explanations.

Recommended Reading

- Balzer, W.K., Doherty, M.E., & O'Connor, R., Jr. (1989). (See References)
 Kluger, A.N., & DeNisi, A. (1996). (See References)
 Salmon, A.W., Schmidt, R.A., & Walter, C.B. (1984). (See References)

Although FIs are widely used (e.g., performance appraisals, grades, teaching evaluations), little is known about how they work. As a result, psychologists do not understand when and why FIs might have negative rather than positive (or no) effects on performance. In the present article, we offer an initial explanation of the effects produced by FIs, drawing upon three theoretical constructs that have been developed in connection with control theory: the regulation of feedback-standard discrepancies, locus of attention, and task complexity. These theoretical constructs pertain mostly to the motivational processes induced by FI. The learning processes induced by FI are beyond the scope of this review.² We begin by tracing the development of the assumption that FIs are always highly effective interventions (for a more thorough review, see Kluger & DeNisi, 1996).

BRIEF HISTORICAL REVIEW

Two figures probably contributed the most to the belief that FIs almost always improve performance: Thorndike and Ammons. Thorndike (1913) provided the initial theoretical arguments for the effectiveness of feedback with his law of effect. This theoretical perspective equated a positive FI with reinforcement and a negative FI with punishment (Thorndike,

1927). Both a positive FI and a negative FI should improve performance because one reinforces the correct behavior and the other punishes the incorrect behavior.

Although several reports were empirically consistent with these predictions (e.g., Thorndike, 1927), the law of effect was never sufficiently detailed to account for the inconsistent findings. For example, Thorndike (1913, p. 286) noted that school grades can impede learning, but he suggested that their normative nature (comparison with others) and their low level of specificity attenuate their effectiveness as FIs. The effect of norms cannot be explained by the law of effect, even though the effects of norms are consistent with empirically supported theories linking normative FI with ego involvement ("how well am I doing relative to other people?") versus task involvement ("how can I improve my performance?") (cf. Butler, 1987).

Furthermore, the specificity feature of the law of effect, which suggests that as the FI becomes more specific, its effect on performance becomes more positive, is inconsistent with some data. To salvage the specificity argument, some researchers suggested that moderate levels of specificity have the most positive effects on performance (e.g., Salmon et al., 1984). Yet this revised argument, too, has not received consistent support. In conclusion, the law of effect generated a sizable empirical literature (cf. the review and criticism by Annett, 1969) because it has the advantage of parsimony, but it appears to be too broad to explain the empirical complexities associated with FI.

Ammons's contribution to the belief that FIs are almost always effective stems from his authoritative article on the effectiveness of feedback (Ammons, 1956). This highly cited review summarized the results of 50 years of literature regarding knowledge of performance

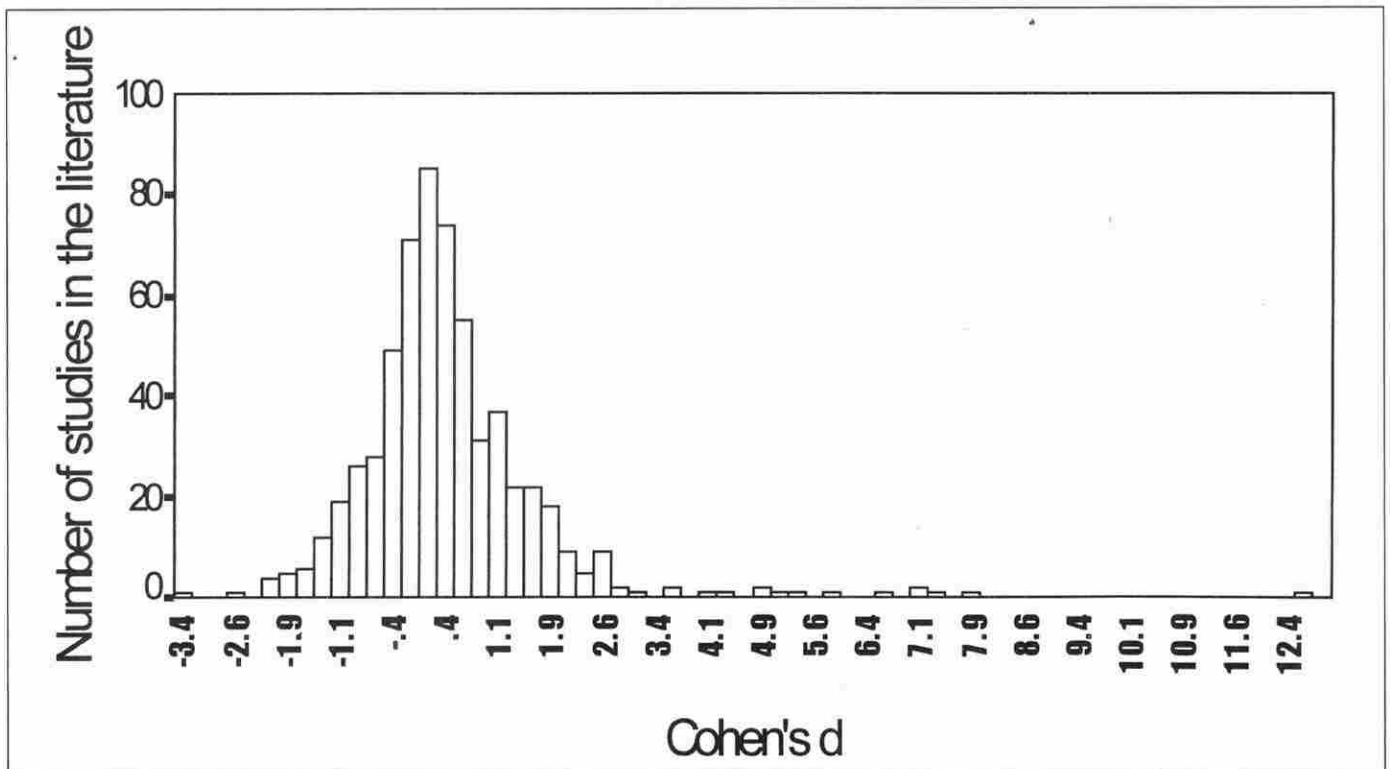


Fig. 1. Distribution of 607 comparisons of performance levels of people who received feedback intervention and people who did not receive feedback intervention. The performance differences are expressed in standard deviation units (d values); positive values indicate that the feedback intervention improved performance, and negative values indicate that the feedback intervention debilitated performance.

(KP), also referred to as knowledge of results (KR). (These old terms refer to a form of FI.) Ammons offered two broad statements: KP increases learning, and KP increases motivation. However, his work suffered from three drawbacks. First, he did not explore evidence inconsistent with his generalizations. An example of the partial support for his conclusion regarding learning can be found in his report of Pressey's work on the self-scoring device. The self-scoring device was a mechanical device—used in the precomputer days—that allowed students to see the correctness of their answers to multiple-choice exams. That is, the self-scoring device provided a type of KP (or FI). Ammons (1956) duly noted Pressey's (1950) conclusion that the immediate self-scoring device improves learning in most cases, but ignored Pressey's report that this device decreased learning

in some others (e.g., learning of Russian vocabulary). Second, some of Ammons's conclusions were based on little evidence. For example, the support for his conclusion regarding the effects of KP on motivation was questionable, at best. Specifically, he admitted that the support for the effects of KP on motivation "has been collected informally" and is "inferred" from other findings (p. 285). He suggested that the fact that people often like to receive feedback is evidence for the positive effect of KP on performance. That is, he confused the motivation to hear the feedback with the motivation to improve performance. Finally, Ammons's review of the literature was not comprehensive. He did not even refer to some troubling studies that were inconsistent with his major conclusions (for sources dating back to 1906, see Kluger & DeNisi, 1996).

After Ammons's review, empirical inconsistencies continued to accumulate. But although a few scholars carefully noted these inconsistencies (e.g., Ilgen et al., 1979), the view that has dominated thinking about FIs during the second half of the 20th century is well typified by the following statement: "The positive effect of FB [feedback] on performance has become one of the most accepted principles in psychology" (Pritchard, Jones, Roth, Stuebing, & Ekeberg, 1988, p. 338).

This brief historical review illustrates that the effects of FI on performance have never been consistent or simple. Moreover, it underscores the fact that there is really very little theory concerning how FI might affect performance. As a result, to understand the effects of FI on performance, researchers need to develop theoretical propositions about the

processes that mediate between the FI stimulus and performance. We hope that the theoretical considerations we discuss here will begin to generate research aimed at understanding these processes better.

Our theoretical suggestions are based on control theory (Carver & Scheier, 1981), but also depend heavily on feedback intervention theory (FIT; Kluger & DeNisi, 1996). FIT has three basic arguments that are relevant here: (a) behavior is regulated by comparisons of feedback with goals or standards (and identification of gaps between the two); (b) attention is limited, and only those feedback-standard gaps that receive attention actively participate in behavior regulation; and (c) FIs change the locus of attention and therefore affect behavior.

DISCREPANCIES

Both control theory and FIT claim that behavior is regulated through the control of discrepancies or errors in the system. When a self-regulating system detects discrepancies or errors, the system is motivated to reduce or lower the perceived discrepancies. Even among competing cognitive theories, the detection and evaluation of feedback-standard (or feedback-goal) discrepancies is considered a fundamental source for motivational processes.

However, most cognitive treatments of the process of discrepancy reduction are indifferent to the valence (positive vs. negative) of the discrepancy. That is, these views suggest that effects are symmetrical, and that both a positive discrepancy and a negative discrepancy yield a self-regulatory action that is a function of the absolute magnitude of the discrepancy. Similarly, behaviorism (Thorndike, 1927) has symmetrical predictions, in that rewards and punishment can produce learning equally.

Other theorists have argued, however, that the reaction to positive and negative events is vastly different (cf. Taylor, 1991). That is, they contend that the direction of the feedback-standard discrepancy has major consequences, that reinforcement and punishment have different and asymmetric effects on behavior (Taylor, 1991).

Yet, despite these disagreements, the theories that emphasize symmetry actually recognize asymmetry, and vice versa (for more details, see Kluger, in press). Thus, both theoretical approaches may be correct. People may possess parallel systems that in concert support survival; one operates with symmetric rules and the other with asymmetric rules. These systems may contain both affective and cognitive subsystems. Indeed, the more positive is the direction of the feedback-standard discrepancy (overshooting vs. undershooting the standard), the higher is the resultant pleasantness (the most salient dimension of affect; Kluger, Lewinsohn, & Aiello, 1994) and the amount of nontask (other) thoughts (Kluger, in press). These effects are asymmetrical about the standard. In contrast, the larger the absolute size of the feedback-standard discrepancy (regardless of direction), the higher the resultant arousal (the second dimension of affect; Kluger et al., 1994) and the amount of task-related thoughts (Kluger, in press). These effects are symmetrical about the standard.

Thus, we can offer an initial explanation for the perplexing finding that the valence of feedback does not have a simple moderating effect on FI effectiveness. We suggest that it does not have a simple effect because it activates two response systems, one that responds to valence symmetrically and one that responds asymmetrically. The coexistence of two types of responses to FIs hints that these pro-

cesses may have different effects on performance (e.g., pleasantness may enhance creativity, but arousal may debilitate it). Understanding the role of these systems in mediating the effects of FIs on performance awaits more theoretical development and empirical investigation.

LOCUS OF ATTENTION

The second relevant theoretical construct is locus of attention. We assume that FIs are interventions with high potential to change locus of attention and that knowing where attention is directed provides a better position to predict FIs' effects on performance. That is, after receiving feedback, an individual is very likely to be thinking about something different from what he or she was thinking about before receiving the intervention. Attention can be directed to the self, to the task at hand, or even to the details of the task at hand. We predicted that when FIs cause attention to be directed to the self, the risk that FIs will debilitate, rather than enhance, performance increases (Kluger & DeNisi, 1996).

Our reasoning was that attention to the self can attenuate the effects of FIs because it depletes cognitive resources necessary for task performance (Kanfer & Ackerman, 1989) and produces affective reactions that may interfere with task performance. Therefore, we hypothesized that FIs that contain cues that direct attention to the self, or that are given in a self-threatening environment, will produce weak or even negative effects on performance. Indeed, both FIs that contain praise and FIs that contain destructive criticism (which are likely to direct attention to the self) yield lower performance effects than FIs that do not contain cues to the self (Kluger &

DeNisi, 1996). Moreover, we have shown that other cues that are likely to direct attention to the self are correlated with attenuated effects of FIs on performance (e.g., verbal FIs vs. computerized FIs).

Attention is sometimes directed to the self via normative cues, that is, cues that make the individual compare himself or herself to other people (as percentile scores do). Evidence showing that FIs that direct attention to the self via normative cues are largely ineffective is found in education research. For example, traditional teacher evaluation forms appear to have little effectiveness for developing college professors (Marsh & Roche, 1997). When these forms are used, professors usually receive information on their standing relative to their peers. In contrast, developmental efforts that rate the individual on various performance dimensions that highlight his or her weaknesses and strengths appear to be more effective (Marsh & Roche, 1997). Similarly, in one study, grades increased ego involvement, but did not affect performance relative to a control condition with no FI, whereas task-focused FI (specific comments) increased task involvement and consequently performance (Butler, 1987).

Although attention to the self can debilitate the effectiveness of FIs, this need not always be the case. There are two factors—one cognitive and one motivational—that may determine whether an FI that directs attention to the self has an enhanced effect: the type of task being performed (discussed in the next section) and the types of self-goals activated by the FI. Self-goals may be better understood in light of Higgins's classification of the self into actual, ideal, and ought self (e.g., Higgins, 1997). The actual self is what a person believes he or she is, the ideal self is what a person believes he or she wishes to be, and the ought self is what a person

believes he or she should be. According to Higgins, people can pay attention, among other things, to discrepancies between the actual and the ideal self or to discrepancies between the actual and the ought self. Attention to discrepancies from the ideal self focuses people on promotion goals (possible gains), and leads people to try to attain their ideals (which in our view can never be reached if defined broadly as wishes and dreams). In contrast, attention to discrepancies from the ought self focuses people on prevention goals (possible losses), and leads people to try to meet socially prescribed task standards. Consequently, the reaction to FIs that direct attention to the self will depend largely on the aspect of the self that becomes salient.

Specifically, when attention is focused on the ought self, people will compare the feedback to their ought standard, that is, their perceived obligation. The result of the comparison will push performance toward the standard: Feedback about superior performance will be followed by performance decline, and feedback about inferior performance will be followed by performance improvement. In other words, people receiving positive feedback reduce performance, and people receiving negative feedback improve performance. Such effects are well known, and have been observed both in laboratory experiments (Podsakoff & Farh, 1989) and in field studies of reactions to performance appraisals (e.g., Reilly, Smither, & Vasilopoulos, 1996).

In contrast to FIs that direct attention to discrepancies from the ought standard, FIs that direct attention to discrepancies from the ideal standard are likely to yield feedback with negative valence only. We hypothesize such an effect because we assume that ideals, as opposed to obligations, can

never be achieved. (Our definition of ideals refers to wishes and dreams and is more inclusive than Higgins's definition of ideals as aspirations that can be achieved.) In fact, we obtained data consistent with our argument that ideals cannot be achieved. Specifically, in a small-scale experiment ($N = 21$), we asked students to mark an endpoint on a line reflecting the size of either their actual-ideal discrepancy or their actual-ought discrepancy. The length of the actual-ideal discrepancy line was 42 mm (35%) longer than the actual-ought discrepancy line, $t(19) = 1.91$ $p < .05$, one-tailed. If an FI directs people's attention to a comparison of the actual and ideal selves, they are likely to assess the gap as negative, and they will be motivated to increase activities that will narrow the perceived actual-ideal discrepancies (albeit only as long as the gap is perceived to be bridgeable). We suspect that FIs focusing on the ideal self are common in various developmental training programs (e.g., the Center for Creative Leadership; Guthrie & Kelly-Radford, in press).

An interesting applied question is how organizations can make their employees frame their jobs as part of their ideals and not as part of their obligations. A partial answer to this question may lie in the leadership literature that contrasts transformational (charismatic) leaders with transactional (more typical) leaders (e.g., Howell & Avolio, 1993). That is, charismatic leaders are thought to direct the attention of their followers to their ideals, whereas traditional leaders are thought to direct attention to the obligation of the follower to a contractual transaction with the leader. Therefore, future research should test whether various FI cues are interpreted when attention is directed to various self-loci and whether they, consequently, pro-

duce different motivational and performance outcomes.

TASK PROPERTIES

The final theoretical construct that should be taken into account in trying to understand how FIs affect performance is task properties. Analyses that we conducted (Kluger & DeNisi, 1996) indicated that the effectiveness of an FI depends on the type of task, yet we do not have a theory that successfully differentiates among task types. For example, we do not know what crucial features result in the different effects of FIs regarding tennis playing and FIs regarding managing a group of employees. Resorting to a simple classification, we can, however, consider task mastery (subjective difficulty) and task complexity (objective difficulty; e.g., remembering 5 cues vs. 15 cues). From the perspective of control theory, FIs that direct attention to the self on complex tasks deplete the resources needed for task performance and direct some of these resources to self-related goals (e.g., self-enhancement). In contrast, FIs that direct attention to the self on simple tasks may augment performance in a manner similar to social facilitation effects. (Social facilitation effects are the effects of the presence of other people on performance: Performance of subjectively simple tasks is facilitated, and performance of subjectively complex tasks is hindered.) Indeed, our analyses suggest that the effects of FIs grow more positive either as the task becomes more subjectively familiar or as it becomes more objectively simple. Ironically, then, people who probably need feedback the most benefit the least from typical FIs. These findings are consistent with findings regarding other motivational interventions whose performance benefits are at-

tenuated or even reversed as task complexity increases (cf. Kanfer & Ackerman, 1989).

There are two new avenues to explore regarding how task properties moderate the effects of FIs on performance. First, a motivational intervention can have opposing effects on various components of task performance. Kairi (1996) measured both reaction time (time from stimulus onset to release of a finger from a waiting key) and movement time (time from finger release to hitting the target) in an "odd man out" task. Participants were asked, in each trial, to choose from among three lights the one that was the greatest distance from the other two. Kairi manipulated social facilitation by having an experimenter sit next to the participants in the experimental group and by letting participants in the control group perform alone in a room. The presence of the experimenter improved movement time, but slowed (insignificantly) reaction time. Perhaps the difficulty in finding performance effects of FIs is due, in part, to their opposing effects on different components of overall task performance.

The second avenue to be explored is based on the distinction between two cognitive systems. Many scholars recognize that some cognitions are governed by a rational or rule-based system, and others are governed by an association-based or experiential-based system (Slooman, 1996). The rational system may be more susceptible to resource depletion, and hence tasks that are largely dependent on this system may be more susceptible to negative effects on performance. This is another possibility that awaits empirical research.

PRACTICAL IMPLICATIONS

Our review suggests that FIs can be double-edged swords. Practitioners

may ask what they can do to minimize the documented risks associated with FIs. One clear answer lies in using FIs only in combination with goal-setting intervention. Providing FIs that relate to previously established goals is likely to direct attention to the task at hand and not to the self. Indeed, both our meta-analysis and other reviews (Kluger & DeNisi, 1996) suggest that a goal-setting intervention augments FIs' effects on performance. Moreover, we have found that employees who wish to have more feedback than they are receiving often suffer from the absence of clear goals. Similarly, current models of effective training evaluation emphasize that building measures for evaluation requires a process of need analysis and goal setting. It seems that providing FIs without clear goals increases the risk that the recipient's goals will not be those intended by the FI provider. But, perhaps more critically, we also suggest that the practitioner interested in developing and implementing FIs take the time to test the effectiveness of these interventions rather than simply assuming that they will work.

Acknowledgments—The work reported here was supported by a grant from the Israel Foundations Trustees and by a grant from the Recanati Fund of the School of Business Administration at the Hebrew University.

Notes

1. Address correspondence to Avraham N. Kluger, School of Business Administration, The Hebrew University—Mt. Scopus, Jerusalem 91905, Israel; e-mail: mskluger@pluto.mssc.huji.ac.il.

2. Detrimental FI effects on learning processes have been noted with respect to learning of judgment tasks (for a review, see Balzer, Doherty, & O'Connor, 1989), motor skills (Salmoni et al., 1984), and other tasks (Kluger & DeNisi, 1996).

References

- Ammons, R.B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *Journal of General Psychology*, 54, 279-299.
- Annett, J. (1969). *Feedback and human behaviour*. Harmondsworth, England: Penguin.
- Balzer, W.K., Doherty, M.E., & O'Connor, R., Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, 106, 410-433.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79, 474-482.
- Carver, C.S., & Scheier, M.F. (1981). *Attention and self regulation: A control theory of human behavior*. New York: Springer-Verlag.
- Guthrie, V.A., & Kelly-Radford, L. (in press). Developing leaders with a feedback intensive program. In C.D. McCauley, R. Moxley, & E. Van Velsor (Eds.), *The Center for Creative Leadership handbook of leadership development*. San Francisco: Jossey-Bass.
- Higgins, E.T. (1997). Beyond pleasure and pain. *American Psychologist*, 52, 1280-1300.
- Howell, J.M., & Avolio, B.J. (1993). Transformational leadership, transactional leadership, locus of control, and support for innovation: Key predictors of consolidated-business-unit performance. *Journal of Applied Psychology*, 78, 891-902.
- Ilgen, D.R., Fisher, C.D., & Taylor, M.S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64, 349-361.
- Kairi, B. (1996). *The moderating effect of social facilitation on the relationship of response time and intelligence*. Unpublished master's thesis, The Hebrew University of Jerusalem, Jerusalem, Israel.
- Kanfer, R., & Ackerman, P.L. (1989). Motivation and cognitive abilities: An integration/apptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74, 657-690.
- Kluger, A.N. (in press). Feedback-expectation discrepancy, arousal and locus of cognition. In M. Erez, U. Kleinbeck, & H. Thierry (Eds.), *Work motivation in the context of a globalizing economy*. Mahwah, NJ: Erlbaum.
- Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: Historical review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254-284.
- Kluger, A.N., Lewinsohn, S., & Aiello, J. (1994). The influence of feedback on mood: Linear effects on pleasantness and curvilinear effects on arousal. *Organizational Behavior and Human Decision Processes*, 60, 276-299.
- Latham, G.P., & Locke, E.A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, 50, 212-247.
- Marsh, H.W., & Roche, L.A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, 52, 1187-1197.
- Podsakoff, P.M., & Farh, J.H. (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes*, 44, 45-67.
- Pressey, S.L. (1950). Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. *Journal of Psychology*, 29, 417-447.
- Pritchard, R.D., Jones, S.D., Roth, P.L., Stuebing, K.K., & Ekeberg, S.E. (1988). Effects of group feedback, goal setting, and incentives on organizational productivity [Monograph]. *Journal of Applied Psychology*, 73, 337-358.
- Reilly, R.R., Smither, J.W., & Vasilopoulos, N.L. (1996). A longitudinal study of upward feedback. *Personnel Psychology*, 49, 599-612.
- Salmon, A.W., Schmidt, R.A., & Walter, C.B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, 95, 355-386.
- Slovan, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22.
- Taylor, S.E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, 110, 67-85.
- Thorndike, E.L. (1913). *Educational psychology: Vol. I. The original nature of man*. New York: Teachers College, Columbia University.
- Thorndike, E.L. (1927). The law of effect. *American Journal of Psychology*, 39, 212-222.

The Localization of a Simple Type of Learning and Memory: The Cerebellum and Classical Eyeblink Conditioning

Joseph E. Steinmetz¹

Department of Psychology, Indiana University, Bloomington, Indiana

One of the most intriguing problems in psychology and neuroscience that has been widely studied over the past century is how the vertebrate brain encodes learning and memory. During this time, a number of researchers using a variety of methods have systematically explored locations in the brain where learning and memory may be encoded. These studies have shown that the brain is composed of a variety of learning and memory systems that are involved in encoding the rich variety of classes of learning and memory that vertebrates are capable of exhibiting.

Classical eyeblink conditioning in rabbits is one form of simple associative learning that has been widely studied, and this paradigm has become the model behavioral system of choice for studying many aspects of the neural correlates of simple learning and memory. This simple yet very elegant set of procedures was initially described and characterized by Gormezano and his colleagues (Gormezano, Kehoe, & Marshall-Goodell, 1983). In this paradigm, a tone or light is the conditioned stimulus (CS). The unconditioned stimulus (US) is an airpuff or electric shock near the eyes. Ini-

tially, the US causes a vigorous reflexive eyeblink called the unconditioned response (UR). With continued paired presentation of the CS before the US, however, the CS comes to elicit an eyeblink response (the conditioned response, or CR). For eyeblink conditioning to occur, the time between the presentation of the CS and US can

Recommended Reading

- Anderson, B.J., & Steinmetz, J.E. (1994). Cerebellar and brainstem circuits involved in classical eyeblink conditioning. *Reviews in the Neurosciences*, 5, 251-273.
- Lavond, D.G., Kim, J.J., & Thompson, R.F. (1993). Mammalian brain substrates of aversive classical conditioning. *Annual Review of Psychology*, 44, 317-342.
- Steinmetz, J.E. (1996). (See References)
- Thompson, R.F., & Krupa, D.J. (1994). Organization of memory traces in the mammalian brain. *Annual Review of Neuroscience*, 17, 519-549.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.