Psychological Bulletin
1996, Vol. 119, No. 2, 254–284

0033-2909/96/$3.00

The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory

Avraham N. Kluger
The Hebrew University of Jerusalem

Angelo DeNisi
Rutgers University

Since the beginning of the century, feedback interventions (FIs) produced negative—but largely ignored—effects on performance. A meta-analysis (607 effect sizes; 23,663 observations) suggests that FIs improved performance on average ($d = .41$) but that over ⅓ of the FIs decreased performance. This finding cannot be explained by sampling error, feedback sign, or existing theories. The authors proposed a preliminary FI theory (FIT) and tested it with moderator analyses. The central assumption of FIT is that FIs change the locus of attention among 3 general and hierarchically organized levels of control: task learning, task motivation, and meta-tasks (including self-related) processes. The results suggest that FI effectiveness decreases as attention moves up the hierarchy closer to the self and away from the task. These findings are further moderated by task characteristics that are still poorly understood.

> To relate feedback directly to behavior is very confusing. Results are contradictory and seldom straight-forward. (Ilgen, Fisher, & Taylor, 1979, p. 368)

> The effects of manipulation of KR [knowledge of results] on motor learning . . . reveal . . . some violent contradictions to earlier beliefs about KR, and some glaring absences in our knowledge. (Salmoni, Schmidt, & Walter, 1984, p. 378).

> Feedback does not uniformly improve performance. (Balcazar, Hopkins, & Suarez, 1985, p. 65)

> Few concepts in psychology have been written about more uncritically and incorrectly than that of feedback. . . . Actually, feedback is only information, that is, data, and as such has no necessary consequences at all. (Latham & Locke, 1991, p. 224)

We argue that a considerable body of evidence suggesting that feedback intervention (FI) effects on performance are quite variable has been historically disregarded by most FI researchers. This disregard has led to a widely shared assumption that FIs consistently improve performance. Fortunately, several FI researchers (see epigraphs) have recently recognized that FIs have highly variable effects on performance, such that in some conditions FIs improve performance, in other conditions FIs have no apparent effects on performance, and in yet others FIs debilitate performance (also see U.S. Congress, Office of Technology Assessment, 1987). These conditions, or moderators of the effect of FIs on performance, are poorly understood and go far beyond the view that FIs improve performance unless the feedback is too negative. Yet, many contemporary researchers still assume that FIs consistently improve performance. Therefore, our first goal is to document both the evidence for inconsistent FI effects and the disregard for these data from the onset of FI research. However, inconsistent FI effects could stem from either sampling (and other) artifacts or from real phenomena that require theoretical explanations (Schmidt, 1992). Therefore, our second goal is to quantify the variability of FI effects and to rule out artifact-based explanations to FI effects variability.

We then argue that the major culprit in sustaining the unwarranted assumptions about FIs is a lack of a FI theory. In the absence of a FI theory, FI-related hypotheses were largely derived from the behavioristic law of effect (Thorndike, 1913, 1927). However, these hypotheses were inconsistent with data in many ways (e.g., Annett, 1969). More recent and tenable hypotheses were derived from theories that included feedback as a theoretical component (e.g., goal setting theory, Locke & Latham, 1990; control theory, Carver & Scheier, 1981), but these are limited only to some of the known FI-induced processes (e.g., to motivation or to learning processes, but not to both). As a result, recent FI research is carried out by isolated pockets of researchers who share either a theoretical or a paradigmatic orientation. Therefore, our third major goal is to attempt the integration of the varying theoretical and paradigmatic perspectives on FI, and thus we propose a feedback intervention theory (FIT). In developing FIT, we borrowed heavily

Avraham N. Kluger, School of Business Administration, The Hebrew University of Jerusalem, Israel; Angelo DeNisi, School of Management and Labor Relations, Rutgers University.

This article was partially funded by a grant received from Rutgers University Research Council and by the Recanati Fund of the School of Business Administration, The Hebrew University of Jerusalem.

We wish to thank Mike Campion, Michael Frese, Dan Ilgen, Paul Levy, Ed Locke, Bob Lord, Amnon Rapoport, and Susan Taylor for valuable comments on an earlier version of this article; Dov Eden, Miriam Erez, and Yoav Ganzach for valuable discussions; Jim Gasaway and Andrea Cory for coding the meta-analysis data; Christine Nadolny for library research; Maggie Eisenstaedt for preparing graphs; and Esther Kluger for proofreading. The meta-analysis was presented at the 8th annual convention of the Society for Industrial and Organizational Psychology, San Francisco, CA.

Correspondence concerning this article should be addressed to Avraham N. Kluger, School of Business Administration, The Hebrew University—Mt. Scopus, Jerusalem 91905 Israel. Electronic mail may be sent via Internet to mskluger@olive.mscc.huji.ac.il.

from control theory (Carver & Scheier, 1981), goal setting theory (Locke & Latham, 1990), action theory (e.g., Frese & Zapf, 1994), action identification theory (Vallacher & Wegner, 1987), a variant of learned helplessness theory (e.g., Mikulincer, 1994), and other pertinent research. Our fourth and final major goal is to provide a preliminary test of FIT. The preliminary test of FIT was performed by subjecting some putative moderators, identified by FIT, to moderator analyses of the meta-analytic effects. We conclude with implications of our work for future FI research and practice.

We must begin, however, by defining the scope of FIs in the present article. This article is about FIs defined as *actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance.* This definition is similar to the notion of "knowledge of performance" interventions (Ammons, 1956), "augmented feedback" (Annett, 1969; Salmoni et al., 1984), or "extrinsic feedback" (Annett, 1969; Frese & Zapf, 1994), and as such it has several implications for the boundaries of our investigation. First, FIs include knowledge of results (KR) interventions—the focus of much of the original research in this area—but FIs are broader in scope than KR interventions. A KR intervention may include, for example, a message that "your average typing speed is 100 words per minute." Yet, providing information about *how* one performs a job, such as "you do not use your thumb for typing," is not a KR intervention because it does not contain information about the effectiveness of one's performance (the average typing speed may be 20 or 120 words per min., regardless of whether one uses the thumb). Yet, such an intervention is an FI and therefore is included in our definition. Second, the definition includes FIs for a wide spectrum of tasks, such as test performance, memory tasks, physical tasks, attendance behavior, complying with regulations, and so forth. These tasks are of interest to researchers in educational, social, industrial, organizational, and developmental psychology, as well as other subdisciplines. However, our definition excludes several areas of investigation: (a) natural feedback processes, such as homeostasis, intrinsic feedback, or the negative-feedback-loop of a control system (Carver & Scheier, 1981), that operate *without* an external intervention; (b) task-generated feedback (e.g., a gardener seeing that he or she flooded the plant) that is obtained without an intervention; (c) personal feedback (e.g., "he doesn't like you") that does not relate to task performance; and (d) self-initiated, feedback-seeking behavior (e.g., Ashford & Cummings, 1983). Our definition has some overlap with interventions used by organizational development (OD) change agents (Nadler, 1977); but to the extent that OD efforts involve interpersonal issues, they are outside the focus of our work. In summary, we concentrate on task-performance FIs, including KR interventions, but exclude research dealing with feedback that is not part of an intentional intervention by an external agent. These criteria are designed to focus on the question of what effects can a teacher, manager, boarding school counselor, or computer-program designer expect to obtain from an FI.

Historical Perspective

Early FI Research

FI research dates back almost 100 years. Several experiments in the beginning of the century suggested that KR interventions—a form of FIs—increase performance (Arps, 1920; Book & Norvell, 1922; Brand, 1905; Brown, 1932; Crawley, 1926; Elwell & Grindley, 1938–1939; Gates, 1917; Gilliland, 1925; Johanason, 1922; Jones, 1910; Judd, 1905; Manzer, 1935; Smith, 1933; Spencer, 1923; Thorndike, 1927; Waters, 1933; Wright, 1906). Most of these studies, however, suffered from major problems, including inaccurate operationalizations of KR, poor methodology, and lack of attention to inconsistent results.

For example, some researchers labeled their manipulation as KR but manipulated other variables, such as recitation (Gates, 1917), suggestion (Brand, 1905; Jones, 1910), or incentive and punishment (Johanason, 1922). These researchers, who equated suggestions and incentives with KR, in fact, manipulated *preperformance* expectations, such as "you are (un)able" to perform (Brand, 1905; Johanason, 1922; Jones, 1910). Thus, they concluded that KR affected performance, while actually manipulating expectations and never manipulating KR. Parenthetically, other researchers manipulated KR but labeled their manipulations as consciousness of habit (Judd, 1905), reward and punishment (Thorndike, 1927), and incentive (Crawley, 1926; Wright, 1906). Thus, from the beginning, there was confusion about what KR meant.

In addition, many studies had suspect methodologies, including samples of four participants or less (Arps, 1920; Brand, 1905; Crawley, 1926; Judd, 1905; Spencer, 1923; Wright, 1906), and experimenters serving as participants (Judd, 1905; Wright, 1906). Furthermore, most did not have proper experimental controls such as randomization of treatment order in within-subjects designs or randomization of participants in between-subjects designs. Although some authors recognized these problems (e.g., Book & Norvell, 1922, reported inequalities between experimental and control groups), none considered these problems as a threat to their conclusions. In fairness, however, some of these studies were quite rigorous in following the participants, under different treatments and on a daily basis, for a long time (e.g., Jones, 1910, 5 months; Arps, 1920, exceeding 1 year). Nonetheless, the methodological and design problems cast doubts about these researchers' conclusions.

Finally, and perhaps most critically, inconsistencies in the beneficial effect of KR on performance were frequent but often ignored. For example, Judd (1905) found that KR, following practice without KR ("ignorance of results"), actually increased performance errors (at least immediately following the manipulation). Yet, Judd concluded categorically that KR improved performance. Similarly, Arps (1920) found that ignorance of results increased performance after prolonged practice, but Arps reasoned that practice produced mental "imagery," which functioned much the same as KR. Nevertheless, he concluded that KR was generally successful in improving performance, providing an early example of confusing naturally occurring feedback with an FI. Also, although Waters (1933) found that KR increased performance in a time-estimation task but did not affect the distance-estimation task, he was still generally positive about the effects of KR. In other studies, the authors were more critical of the benefits of KR interventions. These authors reported that knowledge of progress in a classroom did not affect motivation (Deputy, 1929; Ross, 1933) and that KR had no significant effects on either learning or retention (Crafts & Gilbert, 1935). Furthermore, Mace (1935) concluded, based on one well-controlled and longitudinal study, that "there are conditions under which knowledge of previous performance may decrease the rate of improvement with practice" (p. 12).

Of course, there were several studies that found positive effects of KR on performance (see Ammons, 1956). The problem is that many other studies did not find unilaterally positive effects and, in many of these cases, the authors seemed to ignore their own inconsistent findings in deriving their conclusions. The problem was compounded when other researchers cited the flawed, uncritical and misleading, or both, studies to support their own conclusions (Book & Norvell, 1922; Brown, 1932; Crafts & Gilbert, 1935; Elwell & Grindley, 1938–1939; Gilliland, 1925; Manzer, 1935; Spencer, 1923). For example, Elwell and Grindley (1938–1939) cited Arps (1920), Crawley (1926), and Johanason (1922) in concluding that KR improves performance but never mentioned the problems or inconsistencies in those studies. Even authors who recognized the problems or inconsistencies in the cited sources (e.g., Brown's, 1932, review of the works of Arps, Wright [1906], and Judd [1905]) still relied on these same sources in concluding that KR improves performance. Thus, when Ammons (1956) reviewed the FI (or KR) literature, an uncritical view of the effectiveness of KR had already begun to develop, despite considerable evidence to the contrary.

Ammons's Review

Ammons (1956) summarized the results of the experiments regarding knowledge of performance (KP) or KR and offered several theoretical statements. More important, Ammons's review had a substantial impact on the FI literature: A manual search of the Social Science Citation Index (SSCI) yielded 100 citations of Ammons's paper since 1965 (we did not have access to citations before 1965).[1] His two most broad statements were that KP increases learning and KP increases motivation. Specifically, Ammons's Generalization 3 states that "Knowledge of Performance affects the rate of learning and level reached by learning," such that, when there is KP, learning is "almost universally" (p. 283) enhanced. However, he did not mention evidence inconsistent with Generalization 3. For example, Ammons duly noted Pressy's (1950) conclusion that the immediate self-scoring device (FI) improves learning (of English vocabulary and psychology material) relative to traditional testing conditions (no FI) but ignored Pressy's report that FI *decreased* learning of Russian vocabulary—a fact that did not receive an adequate explanation from Pressy either. Ammons also cited a study by Book and Norvell (1922) in support of his generalization, although in two of its critical conditions most (over 75%) of the participants in the (no KR) control group also increased their performance.

Much of the evidence in support of Ammons (1956) other major conclusion (Generalization 4)—"Knowledge of Performance affects motivation"—"has been collected informally" and is "inferred" from other findings (p. 285). Specifically, Ammons equated positive attitudes toward receiving FI with actual task motivation, while presenting no evidence for the latter. Furthermore, he stated this generalization even though he recognized, elsewhere in his review, that KP hardly affects performance when an individual is already performing at a high level and that KP may decrease motivation if one is doing poorly. In summary, Ammons's review was an unfortunate milestone in FI research because he reached the two most influential conclusions in the area—FI improves learning as well as motivation—without serious consideration of the contradictory evidence that was available to him (e.g., Crafts & Gilbert, 1935; Ross, 1933) and without reference to some troubling papers that were inconsistent with his major conclusions (Deputy, 1929; Mace, 1935; Spencer, 1923).

After Ammons's (1956) review, empirical inconsistencies continued to accumulate. For example, Locke's early experiments in the goal-setting paradigm did not find any FI effects on performance (e.g., Locke, 1967; Locke & Bryan, 1969). These inconsistencies were appropriately noted by several reviewers (Adams, 1978; Annett, 1969; also see epigraphs).

Furthermore, two meta-analyses, testing theories that contain feedback as a component, found only a weak contribution of feedback to performance. First, Harris and Rosenthal (1985) tested several hypotheses designed to explain the well-documented beneficial effect of expectations of others (agents) on one's performance. When agents (primarily teachers) expect high performance from others (primarily students), they may provide more feedback, more challenging goals, and create a better climate for the students. Their meta-analysis showed that the amount of feedback provided by the agent had only a meager effect on performance (r = .07), whereas other variables, such as the climate that the agent created for the other person (r = .36), had strong effects on performance (Harris & Rosenthal, 1985). Second, a test of the job-characteristics model showed that perceived KR has a weak relationship with performance (r = .09) but a stronger effect on variables such as overall job satisfaction (r = .41; Fried & Ferris, 1987).

Nonetheless, scholars continue to ignore findings suggesting that FI effects on performance are highly variable, and Ammons's (1956) review is still cited as evidence for the positive effect of FI on performance (e.g., Ashford & Cummings, 1983). Moreover, scholars citing reviews that concentrated on FI communication processes (e.g., Cusella, 1987; Ilgen et al., 1979) meshed the inaccurate conclusion of Ammons with the emerging literature on FI communication processes and continued to suggest that FIs usually improve performance (e.g., Ashford & Cummings, 1983; Pritchard, Jones, Roth, Stuebing, & Ekeberg, 1988). This view is typified by statements such as "the positive effect of FI on performance has become one of the most accepted principles in psychology" (Pritchard et al., 1988, p. 338). Therefore, when researchers in various psychological fields obtain no FI effects on performance, they search for post

[1] Ammons also wrote a technical report with the same conclusions that predated his 1956 paper, which was available to several of his contemporaries (e.g., Payne & Hauty, 1955), and his work is also cited in numerous books not covered by SSCI (e.g., Irion, 1969). Although this number of citations may not seem very large (Sternberg, 1992; note that the paper with the highest number of citations has been cited 2,000 times), it is nevertheless unusual for a social science paper to receive more than 100 citations. In fact, Garfield (1983) noted that less than one half of 1% of all papers cited between 1975 and 1979 received 51 or more citations, and the typical paper's "half-life," in terms of citations, is only about 4½ years.

hoc explanations for the lack of FI effects in their particular study. For example, Tubbs, Boehne, and Dahl (1993) referred to the lack of FI effects on performance as an "anomaly" (p. 370); Babad (1990) resorted to a liberal interpretation of nonsignificant FI effects as an indicator of a likely population effect.

It is curious that many researchers still do not appreciate the empirical variability of FI effects on performance given that it was noted by several prominent workers in the field (see epigraphs). Perhaps, the variability of FIs effects is not fully appreciated because this variability was noted only in narrative reviews. Therefore, we sought to quantify the variability and to estimate the average FI effects on performance by conducting a meta-analysis.[2] Furthermore, we attempted to determine whether the variance merely reflects sampling-error variance (Hunter & Schmidt, 1990) or some *true* negative effects of FIs on performance.

Meta-Analysis

Articles for consideration in the meta-analysis were identified through a computerized search of the SSCI, PsycInfo, and the National Technical Informational Services. The search, performed in February 1992, requested all citations that included either *feedback* or *KR* as one identifier and *performance* as a second identifier. Also, previous reviews of the FI literature were consulted. Over 2,500 papers and over 500 technical reports satisfied the search criteria.[3]

The studies reported in the approximately 3,000 papers that we considered had to conform to several criteria to be included in the meta-analysis. First, each had to include at least one treatment group that received FI that was not confounded with other manipulations. Therefore, studies that compared the effect of goal setting *and* FI with a control group that received no treatment were excluded. Second, each had to include at least one control or quasi-control group that received no FI. Accordingly, simple before–after designs were excluded because of serious threats to internal validity with performance data (e.g., in novel tasks, a significant performance improvement is very likely to happen solely because of maturation). Third, each study had to measure performance (rather than merely discuss it). Performance measures from various areas of inquiries were used. Performance measures include reading errors, memory retention, test performance, puzzles, performance with computer-aided instruction (CAI), motor performance, reaction time, arithmetic computations, maintenance jobs, and adherence to regulations. For studies that reported more than one measure of performance, all relevant measures were used (including performance ratings). Fourth, only papers that sampled 10 participants or more were included. The exclusion of papers with very small samples was based on the large sampling error that such papers have and the diminished contribution of small samples to various meta-analytic estimates. The cutoff point of 10 participants led to disproportional exclusion of investigations conducted in clinical setting (many of which were based on a single case study).

Finally, we included appropriate studies only if they supplied sufficient information for calculating a d (Cohen's) statistic (Hunter & Schmidt, 1990). Studies that did not report both means and standard deviations (*SDs*), t values, or both exact F values (i.e., not solely $F < 1$) and all relevant means were excluded. For studies that supplied both means and *SD*, d was calculated by subtracting the control group mean from the experimental (FI) group mean and dividing the difference by the pooled *SD* (Hunter & Schmidt, 1990). Therefore, a positive d value reflects a performance gain because of an FI; a negative value reflects a performance decline. Where only t values were provided, they were converted to d according to the formulas provided by Hunter and Schmidt; where exact F values were provided with means but no *SD*, the pooled *SD* was calculated with the square root of the mean-square error (*MSe*; Seifert, 1991).[4]

More important, we identified six papers that yielded 17 effect sizes based on time-dependent statistics (e.g., ARIMA).[5] These were field studies of the behavior modification tradition with reasonable quasi-experimental designs (Chhokar & Wallin, 1984; Goltz, Citera, Jensen, Favero, & Komaki, 1989; Komaki, Collins, & Penn, 1982; Komaki, Heinzmann, & Lawson, 1980; Pritchard et al., 1988; Reber & Wallin, 1984). We created a quasi d statistics for those effects by transforming ARIMA's ts into ds (by treating them as simple t values) but weighing them by the number of participants involved. The effects of this practice are discussed at the end of this section.

We did not correct for unreliability and other artifacts because only two studies contained information about the criterion reliability. Furthermore, even if we knew the artifact distributions, we would not be able to apply them to small effects that are close to zero without risking the creation of a more biased estimate than the estimate based on the raw findings. This problem arises because we assume that FI has both true negative and true positive effects. For example, effects that are truly positive could be observed as negative because of sampling error. Any correction applied to these observation would further bias our estimate.

---

[2] Previous attempts to quantify the effect of FI on performance suffered from several shortcomings: A meta-analysis of several psychologically based intervention programs included a negligible sample of the FI literature in general (Guzzo, Jette, & Katzell, 1985); a quantitative review (Kopelman, 1986) of FI effects on performance included many methodologically suspect studies and excluded well-controlled studies. In addition, it used the percentage of change in the experimental group relative to the control group as an estimate of effect size—a technique that is severely biased by the arbitrary units of the baseline (while ignoring *SDs* within each study).

[3] We estimate that there are over 3,000 dissertations that satisfy the search criteria. If we consider non-English manuscripts, master's theses, and papers that did not contain the keywords either in the title or the abstract, the total number of papers may exceed 10,000. Nevertheless, cost consideration forced us to consider mostly published papers and technical reports in English.

[4] Formula 4 in Seifert (1991) is in error—a multiplier of n, of cell size, is missing in the numerator.

[5] Unfortunately, the technique of meta-analysis cannot be applied, at present time, to such effects because the distribution of d is based on a sampling of people, whereas the statistics of techniques such as ARIMA are based on the distribution of a sampling of observations in the time domain regardless of the size of the people sample involved (i.e., there is no way to compare a sample of 100 points in time with a sample of 100 people). That is, a sample of 100 points in time has the same degrees of freedom if it were based on an observation of 1 person or of 1,000 people.

From the papers we reviewed, only 131 (5%) met the criteria for inclusion. We were concerned that, given the small percentage of usable papers, our conclusions might not fairly represent the larger body of relevant literature. Therefore, we analyzed *all* the major reasons to reject a paper from the meta-analysis,[6] even though the decision to exclude a paper came at the first identification of a missing inclusion criterion. This analysis showed the presence of review articles, interventions of natural feedback removal, and papers that merely discuss feedback, which in turn suggests that the included studies represent 10–15% of the empirical FI literature. However, this analysis also showed that approximately 37% of the papers we considered manipulated feedback without a control group and that 16% reported confounded treatments, that is, roughly two thirds of the empirical FI literature cannot shed light on the question of FI effects on performance—a fact that requires attention from future FI researchers.

Of the usable 131 papers (see references with asterisks), 607 effect sizes were extracted. These effects were based on 12,652 participants and 23,663 observations (reflecting multiple observations per participant). The average sample size per effect was 39 participants. The distribution of the effect sizes is presented in Figure 1. The weighted mean (weighted by sample size) of this distribution is 0.41, suggesting that, on average, FI has a moderate positive effect on performance. However, over 38% of the effects were negative (see Figure 1). The weighted variance of this distribution is 0.97, whereas the estimate of the sampling error variance is only 0.09.

A potential problem in meta-analyses is *a violation of the assumption of independence.* Such a violation occurs either when multiple observations are taken from the same study (Rosenthal, 1984) or when several papers are authored by the same person (Wolf, 1986). In the present investigation, there were 91 effects derived from the laboratory experiments reported by Mikulincer (e.g., 1988a, 1988b). This raises the possibility that the average effect size is biased, because his studies manipulated extreme negative FIs and used similar tasks. In fact, the weighted average d in Mikulincer's studies was -0.39; whereas in the remainder of the data, it was 0.47. Yet, after the exclusion of Mikulincer's studies, the weighted variance of d was .92, such that 33% of the remaining effects were still negative.

Finally, the 17 quasi ds derived from time series design were all positive with a mean d of 1.69 and a variance of 0.45. Although this value cannot be interpreted because there is no metric for converting this value into a real d, the fact that none of these effects were negative is unlikely to happen by chance if the expected base rate proportion of negative effects is 38%. This finding suggests that these field studies differ systematically from most studies in the literature and a weakness of the meta-analysis method—issues that are discussed in the conclusion. However, dropping these 17 effects hardly changed the estimate of d (.40) but did increase the observed variance (1.28).

These findings justify a search for moderators and, moreover, suggest that the presence of negative effects of FI on performance are robust and not artifacts. Therefore, we turn to theoretical guidance to identify potential moderators.

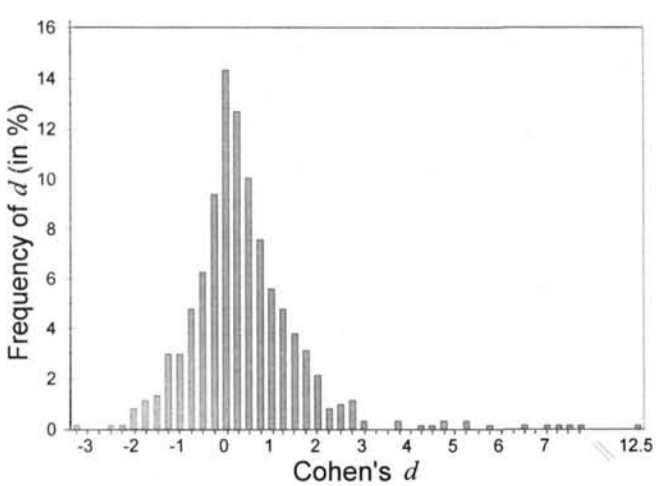

Figure 1. Distribution (histogram) of 607 effects (ds) of feedback intervention on performance.

FI Theories

The Law of Effect

The single most influential theory in this area is Thorndike's (1913) law of effect. Based on the law of effect, a positive FI was equated with reinforcement and a negative FI with punishment. Reinforcement and punishment facilitate learning and hence performance. Both a positive FI and a negative FI should improve performance because one reinforces the correct behavior and the other punishes the incorrect behavior. Whereas 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 noted that grades (i.e., an FI) can impede learning (Thorndike, 1913, p. 286). Thorndike suggested two detrimental properties of such an FI: "Its vice was its relativity [comparison to others] and indefiniteness [low level of specificity]" (p. 288). The property of relativity can not be explained by the law of effect, although it is consistent with empirically supported theories linking normative FIs with extrinsic versus intrinsic motivation (cf. Butler, 1988). The property of specificity can be accommodated by the law of effect

[6] A sample of 100 studies (every 20th) was drawn from the 2,014 studies in this area, whose abstracts are presented on the computerized version of *PsycINFO.* For the majority of the articles, the abstract was sufficient to determine the reasons for exclusion, which were then coded by Avraham Kluger.

In the 100 articles sample, only 3 studies were actually included in the meta-analysis. The reasons for exclusion of the other studies in this sample were lack of a control group (37 articles); lack of any feedback manipulation (27) including review articles, theoretical papers, and mere discussion of feedback as a theoretical explanation for other manipulations; feedback confounded with other treatments (16), primarily with training in behavior modification programs; the dependent variable was not performance (e.g., seeking feedback and satisfaction with feedback; 12); had samples smaller than 10 participants (10), primarily in reports of treating mentally retarded people; feedback was manipulated by removing some source of natural feedback or altering physiological sensations (9), primarily through obscuring vision and change of pitch; and insufficient information to calculate a d statistic (2).

but is inconsistent with data. To salvage the specificity explanation, some researchers suggested that specificity may have an inverted U-shaped relationship with learning (Irion, 1969; Salmoni et al., 1984; indirectly suggested by Hogarth, Gibbs, McKinzie, & Marquis, 1991). Unfortunately, the support for this position has not been consistent, and what support for the inverted-U hypothesis that has been obtained is not consistent with a specificity or precision explanation (e.g., Salmoni et al., 1984). Furthermore, empirical results are inconsistent with the law of effect in many other ways as well (for a review, see Adams, 1978; Annett, 1969; Bilodeau, 1969; Irion, 1969). Yet, despite all the logical (e.g., Powers, 1973) and empirical shortcomings of the law of effect, it had a substantial influence on FI researchers. Therefore, the law of effect was blamed by some for hindering FI research (e.g., Adams's review, 1978; Locke & Latham, 1990). In summary, the law of effect generated sizable empirical literature (cf. the review and criticism by Annett, 1969) because it has the advantage of parsimony, but it is too broad to explain the empirical complexities associated with FI.

Other theories that were applied to FIs have more empirical success but are more limited in their scope. These theories are reviewed below, and their components are incorporated into the proposed FIT. In proposing FIT, we first sought to offer an overarching approach to integrate various perspectives on FI effects and seemingly unrelated areas of investigation. Second, we sought to use FIT to generate hypotheses to be tested on the meta-analytic effects. We, therefore, turn to review the FI-related theories and their relations to the proposed theory.

FI-Related Theories and FIT

Although there is no theory dedicated to FIs, several theories and research paradigms contain the concept of feedback as a central component: control theory, also known as cybernetics, (Annett, 1969; Podsakoff & Farh, 1989), goal setting theory (Locke & Latham, 1990), multiple-cue probability learning paradigm (MCPL; Balzer, Doherty, & O'Connor, 1989), social cognition theory (Bandura, 1991), and a variant of learned helplessness theory (Mikulincer, 1994). These theories have even been used to test FI effects with some limited success. Yet, the domain of the *applications* of these theories to FIs is either motivation or learning. We emphasize that for some theories the application domain is limited, but the theories themselves are not.

For example, goal setting theory postulates strategy development, and control theory postulates problem solving as a component in the theory, although in the context of FIs both were primarily used to theorize about FI effects on motivation. Similarly, the MCPL paradigm was used extensively to investigate FI effects on learning but not on motivation. Yet, non-FI motivational effects on learning in this paradigm were studied, such as incentives (Ganzach, 1994) and time pressure (Rothstein, 1986). Furthermore, a research program of learned helplessness provided extensive information about FI effects on performance (for a summary, see Mikulincer, 1994) but was largely ignored by FI researchers. Therefore, we propose a hybrid theory—FIT—that is geared toward integrating the existing theories, accounting for known processes that are not addressed by existing theories, and attempting to explain the observed inconsistencies in the effects of FI on performance.

FIT has five basic arguments: (a) Behavior is regulated by comparisons of feedback to goals or standards, (b) goals or standards are organized hierarchically, (c) attention is limited and therefore only *feedback-standard gaps* that receive attention actively participate in behavior regulation, (d) attention is normally directed to a moderate level of the hierarchy, and (e) FIs change the locus of attention and therefore affect behavior. These arguments are interdependent, and each consecutive argument is built on the preceding argument. The first argument is already a common component of many FI-related theories. The second through the fourth arguments are largely borrowed from control theory. The last argument, that FIs change the locus of attention, is unique to FIT and is crucial for understanding the FI-performance link. In departure from the existing FI literature, we claim that the important question is, What does an FI do to one's attention?, and not whether it affects task learning and task motivation. That is, FI effects on task learning and task motivation are only some of the possible reactions to an FI, but to better understand these, we need to investigate the total reaction to the FI. We first treat each assumption separately. Second, we integrate them and discuss how FIs may induce task motivation, task learning, and meta-task processes. Third, only then do we offer general testable research propositions.

FIT: Assumptions

Feedback-Standard Comparisons

The first argument is that a basic mechanism in behavior regulation is the evaluation of and reaction to a feedback-standard comparison. This argument is found in several prominent theories such as goal setting theory (Latham & Locke, 1991) and control theory (Carver & Scheier, 1981). Both theories share several key assumptions about the role of feedback in self regulation (Wood & Locke, 1990), despite criticism leveled at control theory by goal setting theorists (e.g., Latham & Locke, 1991). First, both theories view behavior as goal directed. To achieve goals or standards, people use feedback (whether provided by an intervention or not) to evaluate their performance relative to their goals. The result of a comparison of FI to a goal or a standard creates a *feedback sign* (positive or negative evaluation of one's performance relative to the goal)—an argument accepted by researchers across a variety of theoretical orientations (e.g., Bandura, 1991, social cognition theory; Mikulincer, 1994, learned helplessness theory; Podsakoff & Farh, 1989, control theory; Locke & Latham, 1990, goal setting theory).

Although many theorists agree about the evaluation of the discrepancy, they differ with regard to the reaction to it. According to control theory, when a discrepancy is noted, people are motivated to reduce it. Because people strive to reduce the discrepancy, the single system (consisting of a goal, feedback, comparison of the two, and an action to reduce sensed discrepancy) is referred to as a *negative-feedback-loop,* also known as test-operate-test-exit. According to goal setting theory, however, people are motivated to achieve the goal, rather than eliminate the discrepancy.

Regardless of the theoretical differences, both theories recognize that people have several behavioral options in reacting to a feedback-standard discrepancy. In control theory language, the discrepancy can be eliminated by changing behavior to change the future feedback, by changing the standard so it matches the present feedback, by rejecting the feedback, or by escaping the situation (physically or mentally) that signals discrepancy. In goal-setting language, one can strive to attain the goal, change the goal, reject the feedback, or abandon commitment to the goal. Typically though, people choose to eliminate the feedback-standard discrepancy by attempting to attain the standard. When people try to attain the standard, the FI signals that performance falls short of the standard effort is typically increased; when the FI signals that performance exceeds the standard, effort is typically reduced (or maintained). Indeed, participants receiving a negative FI are likely to exert more effort than those who receive a positive FI (S. Anderson & Rodin, 1989; Campion & Lord, 1982; Kernan & Lord, 1991, at least initially for a high-valence goal; Podsakoff & Farh, 1989).

The Four Strategies of Eliminating Feedback-Standard Gap

However, the multitude of coping mechanisms is a theoretical challenge because one needs to predict a priori which of these coping mechanisms will be activated. This challenge was partly addressed by the seminal work of goal setting theory researchers. They have demonstrated convincingly that it is possible to increase the likelihood that the goal will be attained, rather than acting on any other behavioral option, by a goal-setting intervention (Locke & Latham, 1990). In such interventions, initial performance is typically below the standard, that is, the feedback sign is negative. When the feedback sign is negative, people choose to increase their effort, rather than lower the standard, when the goal is clear, when high commitment is secured for it, and when belief in eventual success is high (e.g., high self-efficacy; Bandura & Cervone, 1983). Furthermore, adding an FI to a goal-setting intervention is likely to further block any other coping mode except for changing behavior. Therefore when the feedback sign is negative—as is typically the case in hard goal-setting interventions—adding a clear FI improves motivation to attain the standard (Erez, 1977).

The second strategy in which the discrepancy can be eliminated is to abandon the standard. Abandonment of the standard seems to happen when the discrepancy is perceived to have a low likelihood of being eliminated through actions (e.g., Bandura, 1991; Mikulincer, 1988b). In fact, Mikulincer amassed an impressive body of evidence showing that an *extremely and repeated* negative FI, not just a negative FI, results in a "classical learned helplessness" response (Mikulincer, 1988a, 1988b, 1989a, 1989b, 1989c, 1990; Mikulincer, Kedem, & Zilkha-Segal, 1989; Mikulincer & Nizan, 1988; Mikulincer, Yinon, & Kabili, 1991).

The third strategy for discrepancy reduction is changing the standard, rather than abandoning it altogether. People may lower the standard when they receive negative feedback and when they cannot or do not want to abandon the standard. Alternatively, people may raise their standard when they receive positive feedback (Lewin, Dembo, Festinger, & Sears, 1944) and therefore produce improvements in future performance.

Finally, the fourth strategy for eliminating the perceived feedback-standard gap is to reject the feedback message (e.g., Ilgen et al., 1979). It seems that a negative feedback sign is more likely than a positive feedback sign to lead to feedback rejection (e.g., Ilgen et al., 1979). Feedback rejection may have real-world effects (Pearce & Porter, 1986). Specifically, a satisfactory, as opposed to a high, performance appraisal was perceived (at least among managers) as a negative FI, which was also perceived as unfair and lowered organizational commitment. Yet, the rejection of negative feedback may be culturally dependent. In collectivistic cultures, negative feedback may be accepted more readily than positive feedback (Markus & Kitayama, 1991). Indeed, several studies in Japan and other Asian societies suggest that the effect of feedback sign on its rejection depends on the divergent attributions that may be evoked in different cultures (Markus & Kitayama, 1991).

In summary, all four options for feedback-standard discrepancy reduction received some empirical support. Therefore, FIT incorporates this self-regulation mechanism as one of its building blocks. That is, behavior is regulated with a comparison of feedback to a standard, where a detected discrepancy leads to one of four coping options. Two options involve the *change* of one of the components creating the discrepancy (i.e., change the behavior or change the standard), and two other options involve the *elimination* of one of the components responsible for the discrepancy (i.e., abandon the standard or reject the feedback). However, although the assumption of a feedback-discrepancy reduction mechanism received some support, the challenge of predicting a priori which reduction mode will be selected is only partially met. To predict the effects of FI on task motivation, one needs to know a priori the strength of both the goal and the feedback and the likelihood that either of them can be changed.

The Limitation of the Feedback-Standard Comparisons Argument

In addition to the difficulties in predicting a priori the mode of feedback-standard gap resolution, the feedback-standard comparison argument is insufficient on three additional grounds: (a) It requires an elaboration to account for the role of multiple standards, (b) it cannot account for various findings regarding detrimental FI effects on learning, and (c) it does not incorporate recent findings regarding the effects of FI-induced affect on performance. We therefore turn to review each of these challenges for FIT and then present the additional arguments of FIT that address these challenges.

Multiple standards. The derivation of feedback sign may be a complicated process because feedback is frequently compared with more than one internal standard (e.g., Lewin et al., 1944; Locke & Latham, 1990). Feedback may be compared with a norm (performance relative to others); a prior expectation (Ilgen, 1971; Ilgen & Hamstra, 1972; Kluger, Lewinsohn, & Aiello, 1994); past performance levels, which are used to assess progress in approaching a goal (Carver & Scheier, 1990; Hsee & Abelson, 1991); performance of other groups; and an ideal goal (Lewin et al., 1944). Furthermore, there is ample empirical evidence that multiple standards influence the affective reaction to FIs (Bandura, 1991; Ilgen, 1971; Ilgen & Hamstra, 1972; Kluger et al., 1994; Locke & Latham, 1990). These findings suggest that various feedback-standard discrepancies are weighted and summed into an overall affective evaluation of the FI. Yet, the issue of multiple standards may have two complications. First, multiple standards are not always present. In fact, when an FI is provided on a novel task, the recipient has no clear standard against which to compare the feedback, and therefore it is not surprising that in such cases FIs have no apparent effect on mobilization of effort, if goals are not provided (Bandura & Cervone, 1983; Erez, 1977); in such a case, feedback may be used to calibrate a new task-related standard. The second complication in deriving feedback sign may occur when the FI message conflicts with other naturally occurring sources of feedback in the environment (Albright, Levy, & Williams, 1992). Therefore, FIT must account for the selection of standards for deriving feedback sign and for integration of each of the perceived FI-standard discrepancies into a general perception of the FI.

Detrimental FI effects on learning. Research based on the MCPL paradigm suggests that *outcome* FI (mere KR) impedes learning of complex tasks and subsequently, task performance (Azuma & Cronbach, 1966; Hammond & Summers, 1972; Hammond, Summers, & Deane, 1973; Schmitt, Coyle, & Saari, 1977; for a recent review, see Balzer et al., 1989), leading some researchers to question the basic assumption that people can learn complex rules from FIs or any feedback at all (Brehmer, 1980). Research on computer-aided instruction has demonstrated that adding an FI to CAI programs impairs learning relative to FI-free CAI programs (Carroll & Kay, 1988; Lepper & Gurtner, 1989) or at best has no effects on CAI learning (Wise, Plake, Pozehl, Barnes, & Lukin, 1989). Consistent with these findings, Jacoby, Mazursky, Troutman, and Kuss (1984) reported that seeking outcome feedback was negatively correlated with performance.

The effects of FI-induced affect. FIs strongly influence both pleasantness (e.g., Isen, 1987) and arousal (Kluger et al., 1994). Pleasantness (also known as evaluation, satisfaction, valence, or pleasure) and arousal (also known as activity or activation) are considered to be the two dominant dimensions underlying the affective experience (Mano, 1991; Russell, 1980; but cf. Watson & Tellegen, 1985; for recent criticism of a dimensional approach to the study of affect, see Ortony, Clore, & Collins, 1988; Lazarus, 1991), and they are found across all studies and cultures (Russell, 1991). Typically, FIs with a positive sign elicit positive moods (pleasantness), and FIs with a negative sign elicit negative moods (unpleasantness). Moreover, the size of the feedback-standard gap has been shown to elevate arousal (Kluger et al., 1994), such that FIs with both an extremely positive sign and an extremely negative sign elevate arousal in comparison with FIs with both a moderate sign and no FIs.

Mood researchers have suggested that both pleasantness and arousal affect performance. Specifically, pleasantness has both inhibitory and facilitating effects on cognition and performance (Forgas, Bower, & Moylan, 1990; Isen, 1987; Mano, 1992). For example, *un*pleasant moods have been shown to facilitate elaboration of attitudinal messages (Bless, Bohner, Schwarz, & Strack, 1990) but also to inhibit cognitive flexibility and consequently decrease performance quality on a creative task (Murray, Sujan, Hirt, & Sujan, 1990). Moreover, this research suggests that once pleasantness is induced, regardless of the manner in which it was induced, it has predictable effects on cognitive processing, which extends to unrelated tasks (e.g., Forgas, 1993). Therefore, once an FI has induced pleasantness, the pleasantness has substantial effects on cognitive processing.

FI-induced arousal may also affect performance. The major theory linking arousal and performance, which has received reasonable support, is Easterbrook's (1959) cue-utilization hypothesis (Christianson, 1992). (For information regarding the [in]famous Yerkes–Dodson hypothesis, see K. J. Anderson, 1990; Christianson, 1992; and Neiss, 1990.) According to the cue-utilization hypothesis, arousal increases attention to focal cues and reduces attention to peripheral cues. Hence, complex tasks that require frequent attentional shifts between focal and peripheral cues may be debilitated by a high level of arousal. Indeed, Lewinsohn and Mano (1993) have shown that arousal restricts the cognitive complexity of judgment strategies. Moreover, Eysenck's (1992) review suggests that performance of simple vigilance tasks is enhanced with high arousal and that performance of creativity tasks (requiring multiple attentional shifts) is debilitated, whereas performance of some cognitive tasks (IQ tests) is unaffected. In summary, FIs induce strong affective reactions, which in turn were shown to have automatic and pervasive effects on performance even on tasks other than the one that induced the affect.

In summary, the assumption that behavior is regulated through feedback-standard comparisons and discrepancy reduction is too simple. In its simple form, it cannot accommodate the simultaneous activation of several standards, the detrimental FI effects on learning, and the performance effects of FI-induced affect. By adding additional assumptions presented below, the preliminary FIT sets the foundation to accommodate these challenges.

Hierarchy

Negative-feedback-loops are organized hierarchically.[7] Negative-feedback-loops at the top of the hierarchy contain goals of the self, whereas those at the bottom of the hierarchy contain physical action goals (e.g., open the door). Loops that are high in the hierarchy can supervise the performance of lower level loops, such that the output of higher level loops may be the change of goals for lower level loops. The assumption of hierarchy has received some direct and indirect support and has a theoretical advantage of parsimony.

Direct empirical support for the notion of hierarchy is provided by action identification theory research (Vallacher & Wegner, 1987). According to action identification theory, people can think about their actions in various levels of meaning. For example, the same act can be construed as "reading words" and "investing in my scientific career." People tend to think about a task in increasingly higher levels as they learn the task. As people learn a task, and hence automate its performance, attention is directed at higher levels of action regulation where the meaning of the action is more self related.[8]

[7] Annett (1969) proposed a hierarchical cybernetic model for FI but did not consider the nontask content of the hierarchy or the attention issue, which we see as crucial for any FI theory.

[8] Yet, the ultimate locus of attention is not the core of the self but somewhere below the apex (for reasons discussed in *Normal Locus of Attention*).

Indirect support for the notion of hierarchy is provided by the explanatory power of the concept of hierarchy. The concept of hierarchy can explain differences in action time of various activities (e.g., Lord & Levy, 1994)—phenomena such as depression (Hyland, 1987), personality differences, and leadership styles (Cropanzano, James, & Citera, 1993)—and conditions in which people have incorrect knowledge of their own actions (Frese & Stewart, 1984). Furthermore, this concept provides control theory and FIT with parsimony that allows the integration of various FI-related theories into one paradigm: aspiration level theory, attribution theory, social cognition theory (the notion of self-efficacy), and anxiety theory (see Carver & Scheier, 1981). For example, low self-efficacy is an output of a loop high in the hierarchy that lowers the standard or goal for a lower level standard. That is, low self-efficacy is only one example of the influence of higher order loops on the setting of a lower level standard. Similarly, anxiety is a perception of discrepancy from the ideal self that can be created by an output of an unresolved lower level loop that directs attention to high-level goals. The discrepancy at the self level can be resolved at the task level by escaping the task that causes continuous negative feedback outputs.

The concept of hierarchy can also address criticism that is leveled against simple cybernetic models. According to this criticism, cybernetic models cannot explain the observation that sometimes people increase their goals, that is, widen the feedback-standard gap (e.g., Bandura, 1991). However, this criticism can be accommodated by the assumption that higher level loops always have some unattained goals, as observed, for example, by Jewish writings of the first millenium: "Nobody departs from the world with half his desire gratified" (Midrash Rabbah: Ecclesiastes, 1939, p. 39). If some higher level loop has an unattained goal, then the output of such a high-level goal can be the proactive goal settings for lower level loops (try a new endeavor) or the raising of an existing standard. This suggests that people increase the feedback-standard gap—an apparent contradiction to the operating principle of a negative-feedback-loop. However, this apparent contradiction is solved, when the newly created feedback-standard gap is eventually resolved, thus signaling a reduction of the discrepancy to the higher level loop.

More important, not all loops are linked hierarchically. On the contrary, in any level of the hierarchy, there may be chains of loops that are linked sequentially and organized as scripts or programs (e.g., Gollwitzer, Heckhausen, & Ratajczak, 1990), and occasionally a lower level loop can take control over the system (see the concept of a weak hierarchy in Frese & Zapf, 1994). Therefore, we view programs (e.g., Gollwitzer et al., 1990) as sets of loops that are loosely embedded in the hierarchy.

Attention

At any level of the hierarchy, there may be discrepancies in many negative-feedback-loops, but only those loops receiving attention are acted on. The assumption of limited attention capacity is not problematic and is accepted in a wide spectrum of theories (for a review of the concept, see Carver & Scheier, 1981). However, we do not view the locus of attention as an "all or nothing" phenomenon. Rather, it is a probabilistic process, where most attention is likely to be at one foci, but it can be present simultaneously, or with quick alternations, at different levels of hierarchy and across several standards within the hierarchy.

Normal Locus of Attention

Attention is normally directed to a moderate level of the hierarchy, that is, not to the ultimate goals of the self or to the detailed components of an ongoing activity. The assumption that attention is typically below the self level is based on evidence that people do not like under many circumstances to direct attention to the self (Wicklund, 1975) and the observation that many behaviors are automated and do not require attention to their details for successful operation (Carver & Scheier, 1981). However, the exact level of normal attention typically varies in the moderate levels of the hierarchy as a function of task familiarity as suggested by action identification theory (Vallacher & Wegner, 1987).

FI Effect on Locus of Attention

FIT adds one additional and crucial assumption: FIs command, and often receive, considerable attention. FIs are unlikely to be ignored because any FI has potentially serious implications for the self. Therefore, the question of the perception of FIs is not similar to the classical question of the perception of attitudinal messages. Because FIs receive considerable attention, FIs have the capacity to alter the locus of attention. The question of FIs perception is about the *what* (will receive attention) and not about the *if* (it will be perceived at all). Attention to the FI message may affect the process at different levels of the hierarchy.

FIT: Integrating the Assumptions

To simplify the presentation, we divided the hierarchy into three levels of linked processes involved in the regulation of task performance: *meta-task processes* involving the self, *task-motivation processes* involving the focal task, and *task-learning processes* involving the task details of the focal task. Meta-task processes are at the top of the hierarchy, and task-learning processes are at the bottom. The three-level hierarchy of processes is only an abstraction. The hierarchy is probably more complex and contains more sublevels (Carver & Scheier, 1981; Lord & Levy, 1994). However, this abstraction facilitates the exposition of the major processes proposed in FIT.

We refer to processes that occur *above* the focal task level as meta-task processes to indicate that these processes have the potential to control the focal task processes. Meta-task processes include processes that link the focal task with higher order goals, such as the evaluation of the implication of task performance for the self. Meta-task processes include processes that have considerable effects on performance, such as attention to the self, affect, and possibly framing effects. In our definition of meta-task processes, we include nonfocal task processes and nonfocal task-learning processes, such as a motivation to retaliate against the feedback messenger (M. S. Taylor, Fisher, & Ilgen, 1984) and learning that the feedback sender is untrustworthy. Such processes may not be in themselves meta-task processes; but in the context of FI, they are likely to be activated by meta-task processes to serve higher order goals. The postulation of the meta-task processes is a departure from previous treatments of FI (e.g., Payne & Hauty, 1955) because FIT differentiates between FI-induced motivation and learning processes that are task related and those that are not task related.

In accordance with the assumption that attention is typically at a moderate level of the hierarchy, we start by describing the reaction to an FI at the moderate level where task progress may be monitored and where an FI can affect task motivation. Next, we discuss FI effects on task learning, and finally, we consider FI effects on meta-task processes.

FI Effects on Task-Motivation Processes

FIT incorporates here the mechanisms of feedback sign assessment and the decision on how to react to the perceived feedback sign, as suggested by prior research which is reviewed above. To recapitulate, in the simple case, an FI is compared with a task standard, and effort is increased if the feedback sign is negative and decreased or maintained if the sign is positive. Yet, if a task-motivational process did not culminate in elimination of feedback-standard discrepancy, FIT suggests that attention may be diverted either to lower level processes—task-learning processes—or to higher level processes—meta-task processes. Furthermore, a positive feedback sign may signal to the self that the focal task presents an opportunity for self-enhancement and hence leads to raising the standard for performance and consequently improving future performance (Lewin et al., 1944). These possible processes are depicted in Figure 2 where the processes of a simple negative-feedback-loop are shown by the bold arrows.

Figure 2, as well as Figures 3 and 4, presents possible FI effects on performance. However, it does not speak about the probability that each of the paths will be activated. Therefore, one cannot count the number of paths that lead to positive performance effects, contrast them with the number of paths leading to negative performance effects, and predict the likely effects of FI on performance. This is feasible only when the probability of activating each path is known from additional research. Moreover, the processes at all levels may coexist and even modulate each other. Therefore, these figures present the processes in isolation and as deterministic processes for presentation purposes only. In fact, we view these processes as interdependent and probabilistic.

The task-motivational processes interface with both the task-learning processes and the meta-task processes, and therefore the effects of FI are probably never as simple as the motivational processes would suggest. The processes that interface the task-motivational processes are discussed in the next sections below.

FI Effects on Task-Learning Processes

Learning processes may be activated by the motivational process when the feedback sign is negative, additional effort is deemed insufficient, and the preferred strategy for eliminating the perceived feedback-standard discrepancy is to change behavior (high goal commitment). When people are confronted with subjective failure that they want to overcome, they first try to work harder (Wood & Locke, 1990). Working harder is the output of the motivational process, and it is accomplished by activating programs or scripts for action that are available from past experience. These programs are lower level negative-feedback-loops supervised by the motivational processes. They are activated by default because they require only the allocation of little additional cognitive resources. If working harder fails, people may try to work smarter by generating a hypothesis regarding means for improved performance. This view is consistent with a model proposed by Wood and Locke (1990) to account for the effects of goal setting on complex tasks. According to their model, a motivated performer first activates an universal strategy that can work on most tasks. The universal strategy is to expend more effort, persist, and focus attention on the tasks. If the universal strategy fails, people may search for a task-specific plan; if the latter fails or is not available, people may try to develop a new strategy. The universal strategy and the task-specific strategy correspond to the task-motivational level and the task-learning level in FIT, respectively. This view of the learning end is also consistent with Salmoni et al.'s (1984) suggestion that FI increases cognitive elaboration (more thinking about the task) which results in deeper processing, better retention, and hence a possible learning effect.

Learning processes may also be activated directly by FI cues. FI cues that refer to components of the task (e.g., you are not using your thumb for typing) are likely to direct attention to learning processes and generate working hypotheses, or at least cause their reevaluation.

According to FIT, hypotheses generated by task-learning processes are also standards (considered a feed-forward mechanism in mathematical literature on adaptive systems; Casti, 1989). These hypotheses-standards are used to evaluate the success of new behaviors. This evaluation process is similar to resolving a discrepancy in the motivational processes, that is, the actor is engaged in the loop monitoring the hypothesis either until the results of the behavior match the hypothesis (the hypothesis is confirmed, and a new thing is learned) or until one gives up on that hypothesis and returns to the motivational level. At that level, a new decision may be made either to generate a new hypothesis or to quit trying.

FI-induced attention to learning processes does not guarantee an improvement in performance. On the contrary, when the task is well practiced, attention to task details is likely to interrupt the execution of automatic scripts (well-tested hypotheses) and impair performance (Vallacher & Wegner, 1987). Furthermore, MCPL research suggests that *outcome* FIs (mere KR) impede learning of complex tasks and subsequently task performance (e.g., Balzer et al., 1989; see *Detrimental FI effects on learning*). In fact, outcome FIs cause participants to experiment with successful task strategies, resulting in poorer task performance relative to no-FI controls (Hammond & Summers, 1972). This MCPL finding is consistent with FIT, that is, FIs cause a motivated recipient to test new hypotheses regarding more efficient ways to perform a task. However, the mere *motivation to learn* may backfire because the more varied and elaborate attempts at the task (i.e., decreased cognitive consistency) are often futile. Indeed, Earley, Connolly, and Ekegren (1989) showed in MCPL experiments that an increase in motivation leads to an increase in dysfunctional strategy search. Yet, *process* FIs (i.e., information about judges' policy rather than judges' accuracy) are not detrimental to performance consistency, but their effects on learning are not clear (Adelman, 1981, Experiment 2; Lindell, 1976; Schmitt et al., 1977; Steinman, 1974). This suggests that FI effects on strategy may not always affect overall performance because the alternative strategy is equally effective. Indeed, in some

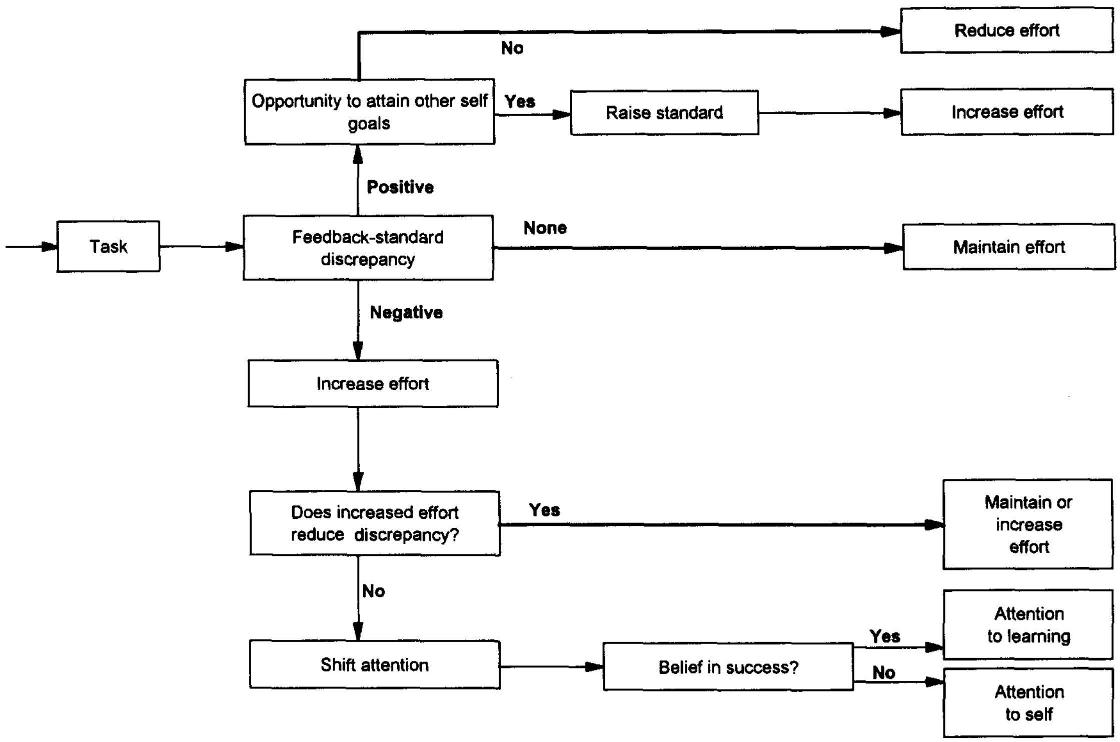

Figure 2. The effects of feedback intervention (FI)-induced attention on task-motivation processes and their consequences for performance. (Simple cybernetic processes are marked with wide arrows; putative FI-performance effects are illustrated by the boxes at the right-hand side of the figure.)

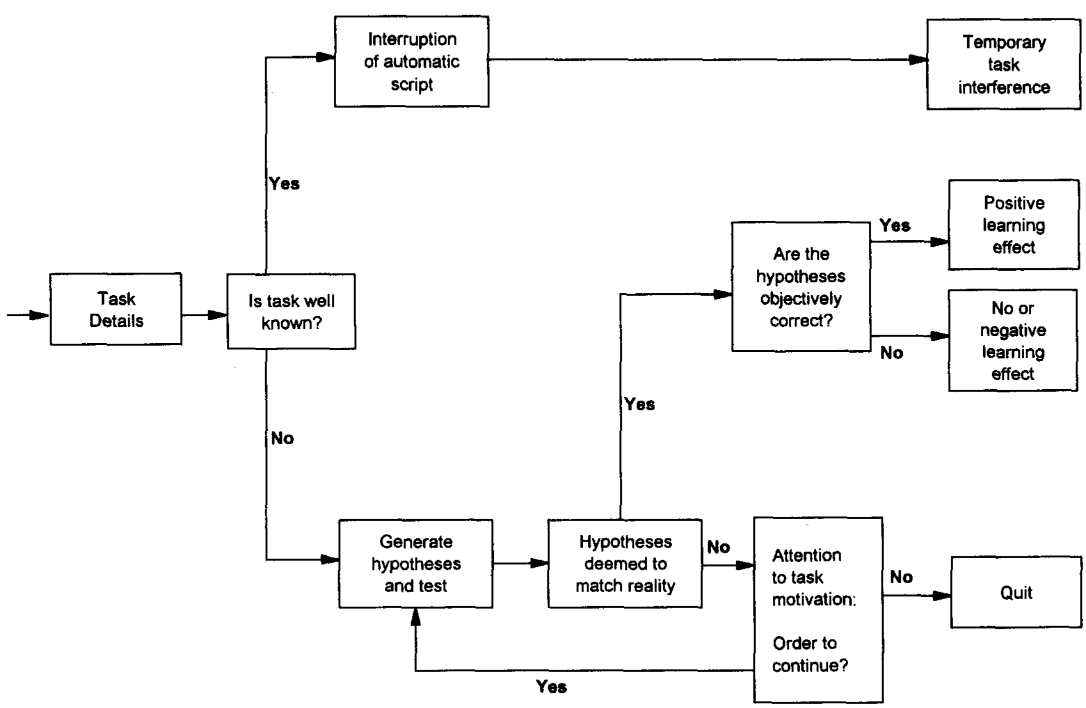

Figure 3. The effects of feedback intervention (FI)-induced attention on task-learning processes and their consequences for performance. (Putative FI-performance effects are illustrated by the boxes at the right-hand side of the figure.)

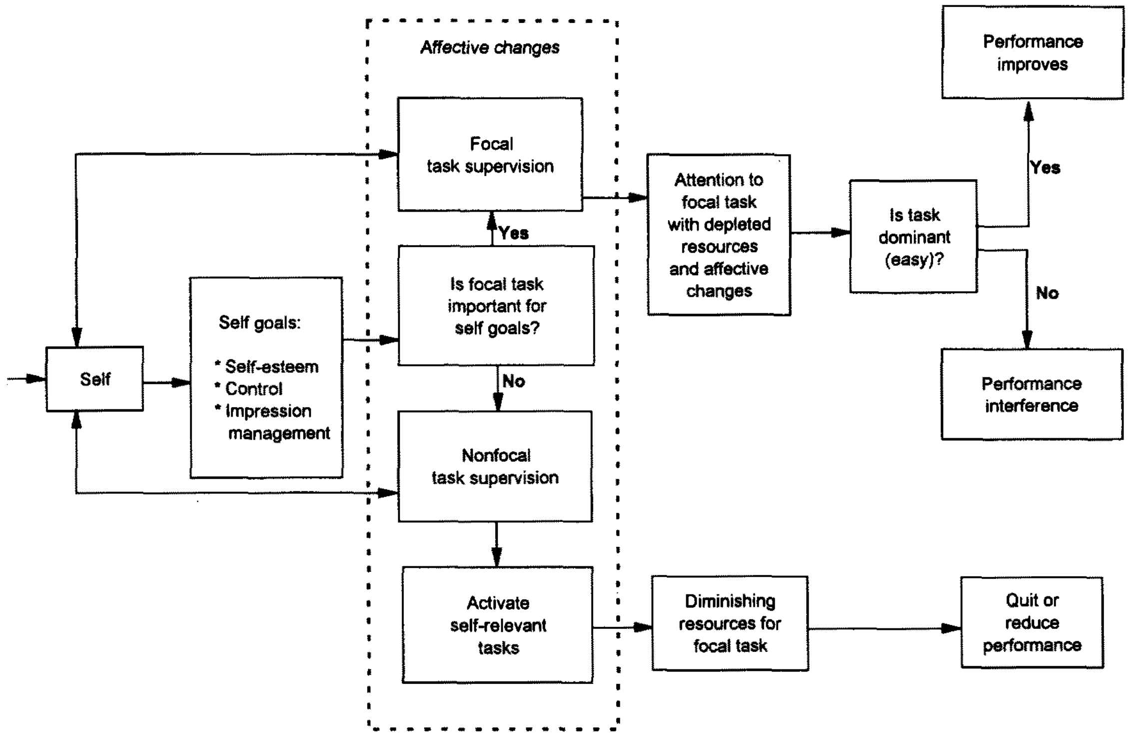

Figure 4. The effects of feedback intervention (FI)-induced attention on meta-task processes and their consequences for performance. (Putative FI-performance effects are illustrated by the boxes at the right-hand side of the figure.)

tasks (e.g., dichotomous MCPL), process FIs induce different strategies than outcome FIs without a noticeable difference in overall performance (Castellan & Swaine, 1977). Finally, even if attention is directed to learning processes, but the informational value of an FI is redundant with the preexisting knowledge, no FI effect on learning should be expected. As a summary, the possible FI-induced learning processes are depicted in Figure 3.

The MCPL literature suggests that for an FI to directly improve learning, rather than motivate learning, it has to help the recipient to reject erroneous hypotheses. Whereas correcting errors is a feature of some types of FI messages, most types of FI messages (see Meta-Analytic Moderator Analyses below) do not contain such information and therefore should not improve learning—a claim consistent with CAI research (reviewed above). Moreover, even in learning situations where performance seems to benefit from FIs, learning through FIs may be inferior to learning through discovery. Learning through discovery is a learning based on feedback from the task, rather than on feedback from an external agent. Task feedback may force the participant to learn task rules and recognize errors (e.g., Frese & Zapf, 1994), whereas FI may lead the participant to learn how to use the FI as a crutch, while shortcutting the need for task learning (cf. J. R. Anderson, 1987). Indeed, in one CAI experiment, it was found that FI was detrimental to the performance of transfer tasks, that is, tasks that were somewhat different than the task on which FI was provided (Carroll & Kay, 1988). This finding suggests that FIs may reduce the cognitive effort involved in task performance and therefore is detrimental in the long run. In FIT language, an FI may lead to the generation of a hypothesis designed to attain a goal of obtaining positive feedback, whereas no FI may lead to the generation of a hypothesis designed to attain a goal of performance improvement.

In summary, FIs affect the learning process by directing attention to discrepancies between the hypotheses (standards) regarding the details of task performance and the outcomes of acting on these hypotheses. If the FI is not accompanied with cues helping to reject erroneous hypotheses, it may cause the recipient to generate a multitude of hypotheses that can reduce consistency and hence decrease performance. Even when the FI is accompanied by useful cues, they may serve as crutches, preventing learning from errors (natural feedback) which may be a superior learning mode. However, an FI may also interfere with the ability to learn when it directs attention up, rather than down, the hierarchy and induces meta-task processes, which are discussed in the next section.

FI Effects on Meta-Task Processes

FI cues and the outputs of task processes may divert attention up the hierarchy and away from the details of the task. This shift of attention may activate at least four interdependent mechanisms: mode of resolving feedback–self discrepancies, attention to the self, depletion of cognitive resources for task performance, and affective processes. Each of these processes is complex and interdependent. However, for demonstration purposes, the possible effects of FI-induced meta-task processes are depicted in Figure 4. Each of these meta-task processes are discussed below.

Mode of Resolving Feedback–Self Discrepancies

The activation of negative-feedback-loop at the self level triggers motivation to reduce the self-related discrepancy. The self-related discrepancy can be reduced by resolving to continue to work on the task that induced the processes. Moreover, these meta-task processes may recognize an opportunity for attaining goals of the self (Frese & Zapf, 1994) and redirect the attention back down, possibly increasing the task-motivation standard. Yet, a self-related discrepancy can also be resolved by multitude of other routes available to a negative-feedback-loop so high in the hierarchy. For example, one may choose to engage in other (nonfocal) tasks or activities that would signal attainment of positive self view. If the latter option is chosen, the negative-feedback sign at the self level is reduced, but the task at hand may be abandoned or receive less attention. Alternatively, the higher loop may lower the standard for the task-motivational process, thereby reducing the negative-feedback sign at the task level and hence also at the self level.

The effects of activating self-related feedback loops are moderated by a host of variables that determine the likelihood of shifting attention back down to the task. One such variable is self-efficacy. Individuals high in self-efficacy are less likely to quit a task even in the face of failure relative to those low in self-efficacy. According to FIT, low self-efficacy is a meta-task mechanism that "releases" unresolved lower level feedback loops (Lord & Levy, 1994). Specifically, this mechanism is responsible for preventing endless attempts to reduce a feedback-standard discrepancy and may be activated whenever an interruption occurs in the lower level feedback loop (Carver & Scheier, 1981).

Another variable that determines the results of activating a self-related feedback loop is anxiety (Mikulincer, 1989a). Anxious participants whose self-related goals were activated are more likely to experience cognitive interference, that is, shifts of attention away from the task and toward the unmet goals of the self. A third variable that affects the decision to continue with the task is the velocity of the FIs. When more than one episode of FI is available, people can assess the rate of change in their performance. When the initial FI is very negative (i.e., large feedback-standard discrepancy), only a rapid rate of improvement leads to a decision to continue with the task, where a constant rate of improvement as well as a delayed improvement lead to a decision to withdraw (Duval, Duval, & Mulilis, 1992). Common to all variables that affect the decision to withdraw from the task is a low expectation to achieve the standard along with a shift of attention to meta-task processes. This shift interferes with performance. There are additional variables that interfere with task performance (see Carver & Scheier, 1981), but FIT assumes that their operations once activated by FI are not unique to the effects of FI and therefore are not discussed here in greater detail.

Attention to the Self

Many FI cues may direct attention to the self. For example, a normative FI is likely to divert attention away from the task to meta-task processes such as evaluating the utility of task performance for higher order goals (e.g., making a good impression; Vallacher & Wegner, 1987). Attention to the self is known to improve performance of dominant tasks and debilitate performance of nondominant tasks—as predicted by several theories such as objective self-awareness theory (Wicklund, 1975) and control theory (Carver & Scheier, 1981). Consistent with FIT, cues of both salient negative and salient positive FI have been implicated in shifting attention to the self.

For example, a devastating negative FI increased self-focused cognition and increased performance of a task relying on two memory cues but debilitated performance of a task relying on six memory cues (Mikulincer, Glaubman, Ben-Artzi, & Grossman, 1991). More important, the task on which the FI was provided was different from the task on which performance was measured, implicating that processes induced by the FI had effects beyond the task on which FI was provided. In addition, praise (a type of FI) impaired the performance of a cognitively demanding task but improved the performance on a simple task—findings interpreted in light of a self-attention model (Baumeister, Hutton, & Cairns, 1990). Moreover, these effects were obtained even when the praise was task irrelevant, thus implicating again, general processes that have effects beyond the task on which they were induced.

Depletion of Cognitive Resources for Task Performance

Regardless of the result of activating self-related loops, the mere shift of attention away from the task involves reallocation of cognitive resources. The attention diverted from a resource demanding activity to the nontask aspects of the intervention (meta-task processes) may cause performance loss due to competition for cognitive resources (e.g., Kanfer & Ackerman, 1989). Only if the task is automated, and therefore fewer resources are needed for its completion, then the motivation induced by the intervention may cause people to successfully work harder.

Affective Processes

Attention to the self is likely to activate affective reactions. Affective reactions may influence the way in which the available resources are used. Therefore, although mere resource competition may influence the amount of resources allocated to the task, FI-induced affect may influence the way the available resources are used.

According to FIT, most affective reactions are induced by evaluation of the feedback with respect to salient self goals—which may create several feedback signs. The feedback signs are then weighted and summed into a general feedback sign. This general feedback sign is then cognitively evaluated both for its harm–benefit potential for the self and for the need to take a new action (Kluger et al., 1994). The harm–benefit potential is a monotonic function of the feedback signs, whereas the need for a new action assessment is a function of the deviations of the signs from their standards (i.e., a curvilinear U-shaped function of the signs). The harm–benefit appraisal is reflected in the primary dimension of mood (i.e., pleasantness), whereas the appraisal of the need for action is reflected in a secondary dimension of mood (i.e., arousal).

The view that affect is partly the result of cognitive appraisal is consistent with several works on the *antecedents* of affect (Bandura, 1986; Higgins, 1987; Latham & Locke, 1991; Lazarus, 1991; S. E. Taylor, 1991). A common assumption in these works is that cognitive appraisals of progress toward, or maintenance of, an important goal result in affective reaction (for a review and criticism of appraisal theories of emotions, see Parkinson & Manstead, 1992). The juxtaposition of the two dimensions creates phenotypical reactions such as anxiety. Anxiety, for example, is an evaluation of a threat to the goals of the self, combined with a tendency to act to terminate the threat. These evaluations result in unpleasantness combined with high arousal. This view of affect suggests that some effects of FI-induced arousal may be common both to people who experience unpleasant moods (e.g., anxiety) and to people who experience pleasant moods (e.g., elation), as long as their arousal level is similar.

This interpretation can account for the numerous findings showing that FI-induced affect influences performance of tasks other than the one used to induce the affect. Further, it may suggest that the output of an activating affective process sets different standards for cognitive operation. For example, FI-induced pleasantness may induce a framing effect (Tversky & Kahneman, 1986), where negatively framed events promote risk seeking and positively framed events promote risk aversion. The framing effect may explain diverse findings showing that a negative FI sign is often followed by greater *variance* in both performance and standard setting than a positive FI sign. For example, Thorndike (as cited in Adams, 1978) found that the word *wrong* yielded lower performance consistency than the word *right* (leading both him and Skinner to concentrate on rewards rather than punishment). Similarly, Lewin et al. (1944) noted that participants receiving negative FIs set goals for the next performance episode with greater variability than those receiving positive FIs, possibly reflecting a risk-seeking strategy.

In summary, FIs induce meta-task processes by directing attention to standards that "supervise" task motivation. The attention to such standards may alter the task goals of the recipient. For example, if the attention was directed to a threat to the self, the gap may be resolved by avoiding the task that caused the threat to the self—this process is thought to be an act of releasing the cognitive structure from the attention given to the task goal (Lord & Levy, 1994). In addition, even if the recipient is capable of shifting attention back down the hierarchy, performance may be affected by prior attention to the self, by cognitive resource depletion, and by affective mechanisms that alter cognitive processing of task information.

FIT: Predicting the Effects of FI on Performance

The assumptions of FIT lead us to suggest that three classes of variables determine the effect of FI on performance: the cues of the FI message, the nature of the task performed, and situational (and personality) variables. FI cues determine which standards of the recipient will receive most attention and hence affect action (Cropanzano et al., 1993; Vallacher & Wegner, 1987). The nature of the task determines how susceptible it is to attentional shifts. Finally, situational and personality variables determine how the recipient chose to eliminate standard-FI gaps to which the FI brought attention. Figure 5 presents a graphical simplification of FIT, where the various effects of FI-induced processes on performance are shown. Direct effects are marked with single-headed arrows, reciprocal processes with double-headed arrows, and interactions with either dashed boxes or a boldface arrow. Each of these classes of variables are discussed next, along with specific research propositions.

FI Cues

Cues available in the FI message determine which level of action regulation will receive most attention. Some FI cues are likely to direct attention to meta-task processes. As such, they are also likely to shift attention away from the task toward other goals of the self and consequently may debilitate performance. (Attention to the self results in superior performance only if the task is very simple or if the recipient diverts attention back down to the task-motivation or -learning level.) An example of FI cues that direct attention to the self are classroom grades. In one study (Butler, 1987), grades increased ego involvement but did not affect performance relative to no-FI control, whereas task-focused FI (specific comments) increased task involvement and consequently performance. Another example is a contrast between a computerized FI and a FI given by the supervisor in a mail-order processing job (Earley, 1988). The computerized FI was more trusted, led to stronger feelings of self-efficacy, to more strategy development, and to better performance compared with an identical FI from a supervisor. According to FIT, the FI given from the supervisor directed attention to meta-task processes (evaluating the intentions of the supervisor and their implications for goals of the self), whereas the computerized FI directed attention to the task and to the task details.

More important, attention to the meta-task goals may lead to disengagement from the task even when the FI is positive. When ego involvement was elicited (FI cue) together with an unflattering (external) attribution for the success, a positive FI caused negative effects on intrinsic motivation and hence on performance (Koestner, Zuckerman, & Koestner, 1987). Similarly, a positive FI, when perceived to be controlling (another FI cue), has been found to reduce intrinsic motivation (Ryan, 1982).[9] From an FIT perspective, these results suggest that although the FI sign regarding the task may have been positive, it attracted attention to a negative FI sign regarding meta-task goals, such as protecting the self. This negative FI sign can be effectively eliminated by disengagement, to the extent possible, from the task. This tendency to disengage from the task however debilitates performance.

> Proposition 1: FI effects on performance are attenuated by cues that direct attention to meta-task processes (P1).

Such cues include normative FIs, person-mediated versus computer-mediated FIs, FIs designed either to discourage or praise the person, and any cue that may be perceived as a threat to the self.

[9] See Locke and Latham (1990), however, for a discussion of the measure of intrinsic motivation—free time spent on a task.

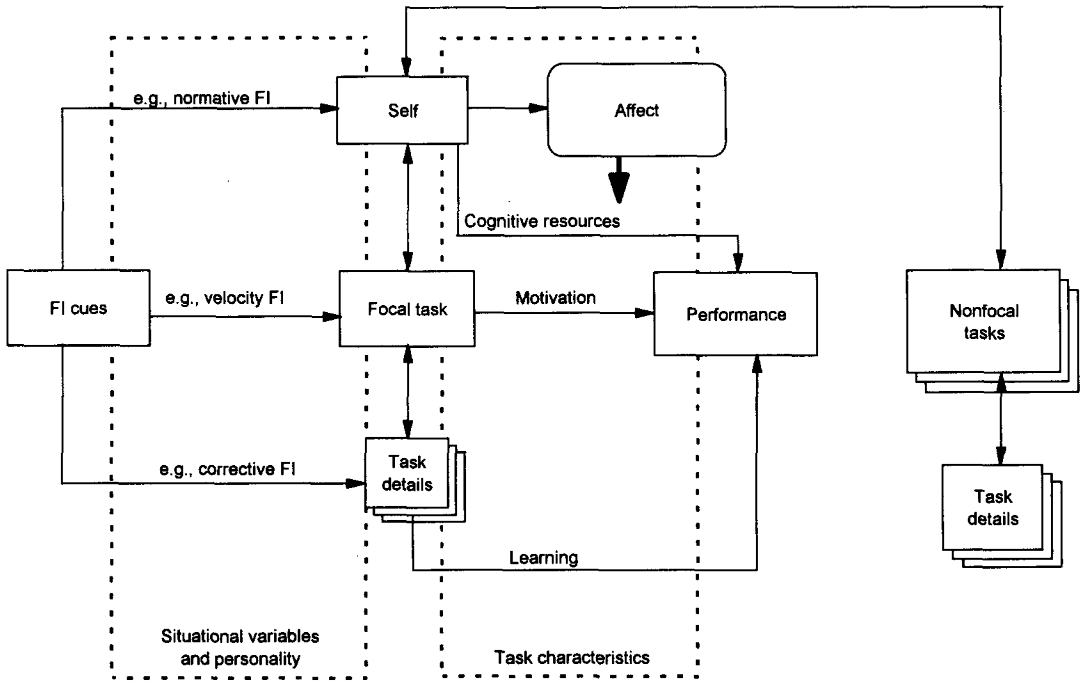

Figure 5. A schematic overview of feedback intervention (FI) theory.

The other side of the same coin is that FI cues that direct attention to the motivational level should improve performance. For example, a velocity FI directs attention to past performance level and may focus attention on task goals. People prefer to receive a positive velocity FI (you improved from the last trial), even at the expense of a lower objective FI sign (Carver & Scheier, 1990; Hsee & Abelson, 1991). This preference may reflect a congruence between FI cues and the normal regulation of action, where temporary goals are set and monitored as a mean for achievement of an overall task goal.

FI that directs attention to learning processes can also improve performance. However, for a direct learning effect, the cues must be sufficient to help the recipient to reject erroneous hypotheses. Therefore, FI messages containing corrective information (e.g., the correct answer is *a* because . . .) should improve performance. This seems to imply, as several FI researchers suggested, that FI specificity is correlated with learning. Yet, FI that is too specific can direct attention below the level necessary for performance, thus causing an interference (Vallacher & Wegner, 1987). In addition, the specific information may not match the natural way people represent the task cognitively and, therefore, attenuate some benefits of FI for learning (Ganzach, 1994). For example, FI with information about magnitude of error, as compared with FI with only the sign of the error (above or below the target), led to poorer performance as indicated by the single cue probability learning (CPL) measure of consistency, thereby showing a situation where "less can be more" (Ganzach, 1994). Yet, the less specific FI also led to lower estimates of variability, suggesting that specificity has complex effects on overall performance (for a discussion of the different facets of performance in CPL, see Lee & Yates, 1992). Therefore, it seems that the empirical data regarding specificity is not clear (Salmoni et al., 1984). However, according to FIT, the crucial aspect of FI that supports learning is its ability to point to erroneous hypotheses. (Of course, nonspecific FI cannot accomplish it, but specific FI does not guarantee it.)

> Proposition 2: FI effects on performance are augmented by (a) cues that direct attention to task-motivation processes and (b) cues that direct attention to task-learning processes coupled with information regarding erroneous hypotheses (P2).

An example of a cue that directs attention to the motivational level is a velocity FI, and a cue that direct attention to learning processes is an FI that contains the correct solution.

Task Characteristics

With very rare exceptions (e.g., Mikulincer et al., 1991; Baumeister et al., 1990), FI researchers have ignored the theoretical importance of task characteristics. An incredible effort was invested in manipulating the predictor (FI) characteristics, whereas very little attention was given to the criterion characteristics. This is surprising given Annett's (1969) comprehensive review of KR studies showing that the pattern of KR effects on motor, perceptual, and verbal tasks may be different. This may be partly due to a lack of sufficient task taxonomy, which was lamented by many authors who recognized that a taxonomy is needed to establish the boundaries of the predictive power of theories in various domains (Hammond, 1992; Wood, 1986). Indeed, Annett (1969) classified KR studies into "some kind of division," where "divisions are somewhat arbitrary" (p. 37).

However, when FI increases motivation, subjective task complexity would moderate the effect of FI because motivation improves performance mostly when the task requires little cognitive resources (Ackerman, 1987; Wood, Mento, & Locke, 1987). When performance is heavily dependent on cognitive resources, extra motivation cannot be translated into better performance. Furthermore, the cognitive resources allocated to the external pressure to perform (attention to meta-task processes) may debilitate performance (Kanfer & Ackerman, 1989). In a similar vein, if performance is measured only for a short duration, performance is largely dependent on existing resources (i.e., intelligence). When the task is sufficiently practiced, the role of intelligence in predicting performance diminishes (Hulin, Henry, & Noon, 1990), whereas the effect of FI-induced motivation may become more apparent.

> Proposition 3: In the absence of learning cues, the fewer cognitive resources needed for task performance, the more positive is the effect of FIs on performance (P3).

Situational Variables

There are several situational variables that might moderate the effects of FI, and these generally involve cues about externally provided goals. When the recipient has clear task goals and can easily compare the FI message with these goals, the presence of a goal-setting intervention may be superfluous. However, to the degree that the feedback-standard gap is ambiguous, a goal-setting intervention should both remove the ambiguity and direct attention to task processes, rather than to meta-task processes. Therefore, goal-setting interventions should augment the effect of FI on performance (Erez, 1977). The augmenting effect of goal-setting interventions should be most apparent when FI cues do not lend themselves to a clear interpretation.

> Proposition 4: Goal-setting interventions should augment the effect of FI on performance (P4).

Personality

There is no doubt that personality variables moderate the reaction to FI (Ilgen et al., 1979), but we provide only a general discussion of personality because we cannot test personality effects with the meta-analysis. Among the personality variables that are known to be involved in the reaction to FI are self-esteem (e.g., Ilgen et al., 1979), locus of control (e.g., Ilgen et al., 1979), tendency for cognitive interference (Kuhl, 1992; Mikulincer, 1989a), and altruism (Korsgaard, Meglino, & Lester, 1994). According to FIT and control theory interpretations (Cropanzano et al., 1993), personality variables are systems of goals and preferences of both attention allocation to different goals of the self and modes of resolving perceived feedback-standard discrepancies. For example, one such goal of the self is "avoiding negative stimuli," where people with this goal tend to have low self-esteem and high levels of anxiety (Cropanzano et al., 1993). Such people are much more likely to pay attention to this self-related goal as a result of receiving negative FI. Therefore, for people who have a salient goal of avoiding negative stimuli, task-level negative feedback is likely to send a signal of negative feedback to a higher level loop at the self level; whereas for people who do not emphasize this self goal, task-level negative feedback is more likely to be resolved by changing behavior or delegating the problem to a learning-level loop. For example, negative FI is more likely to direct attention to the self among participants low in self-esteem than among those high in self-esteem, but positive FI may have the opposite effect.

> Proposition 5: FI cues that match salient self goals of a given personality type direct attention to meta-task processes and therefore debilitate performance (P5).

Testing FIT With the Meta-Analytic Effects

Some of FIT's propositions can be tested on the meta-analytic effects. However, such tests should be considered both preliminary and partial for two important reasons: lack of information regarding the original studies and weaknesses of moderator analyses of meta-analytic effects.

First, original studies lack measures of the processes implied by FIT because they were not designed to test FIT. The fact that the included studies were not designed to test FIT is not a serious problem because meta-analysis has the advantage that it can be used to test new hypotheses never thought of by the original researchers. However, because of the limitation of the available data, we cannot test the *processes* suggested by FIT, but we can test the resultant FI effects on performance. For a similar reasons, we also cannot test any aspect of the last proposition regarding personality because most included studies did not measure the moderating effects of personality.

Second, FIT suggests many complex three-way interactions. For example, whereas the first proposition suggests that attention to the self debilitates performance, this proposition was further qualified by task type. This effect may disappear with a simple task that renders the task insensitive to cognitive resource depletion and with personality and situational factors that direct attention back to the task. The core of the proposition is an argument about a two-way interaction, that is, an interaction between FI presence (FI vs. no FI) and FI content (presence or absence of cues to the self). The qualification suggests a third class of variables that further moderate the two-way interaction. Similarly, the second proposition suggests that attention to task details augment performance. Yet, the second proposition is limited to situations where FI does not induce fruitless learning attempts. This qualification is also an argument about a three-way interaction. Three-way interactions have very low likelihood of being detected in a meta-analytic moderator analysis because of second-order sampling errors such as poor statistical power and extremely skewed and correlated predictors (too few studies populating some of the variable combinations implied by a three-way interaction). Therefore, the following test is largely limited to the simple two-way interactions (the propositions), recognizing a priori that many three-way interactions implied by FIT cannot be detected. However, where the condition of the third class of variables is either not common in the literature or when it has a weaker effect in nature, the simple, two-way interaction should still be detected. Therefore, the following test is a reasonable first step for establishing the viability of FIT.

Meta-Analytic Moderator Analyses

Moderators

Most of the moderators considered were derived from the first four propositions of FIT. In addition, though, several modera-

Table 1
FI Moderators: Descriptive Statistics, Correlations with d values, and Intercorrelations (Before Exclusions)

Moderator	K	M	SD	d	d'	d''	d'''	1	2	3	4	5	6	7	8	9	10	11
1. FI sign	596	3.73	1.31	*24**	*09*	*−01*	−01	(82)										
2. Correct–incorrect	596	0.59	0.48	*−13**	*−03*	*−02*	01	−25	(91)									
3. Correct solution (P2)	596	0.25	0.42	**23***	**19***	**18***	**19***	10	35	(87)								
4. Attainment level (P4)	596	0.30	0.44	07	−01	08	02	13	−27	−10	(85)							
5. Velocity (P2)	596	0.13	0.33	**18***	**15***	**23***	**14***	11	−21	−06	45	(89)						
6. Normative information (P1)	596	0.13	0.32	02	−04	−01	−04	23	−37	−20	12	02	(78)					
7. Norms (P1)	596	0.08	0.26	02	−02	−04	−05	03	−26	−06	19	12	34	(79)				
8. Discouraging FI (P1)	596	0.25	0.42	**−31***	**−16***	**−21***	**−20***	−71	18	−34	−25	−22	−08	−11	(92)			
9. Praise (P1)	596	0.16	0.36	−03	−09	−18*	**−17***	62	−28	−24	−06	02	34	08	−10	(94)		
10. Verbal FI (P1)	596	0.56	0.48	−13*	−02	−12*	**−12***	−12	−01	−27	−27	−15	05	03	42	18	(89)	
11. Written FI (P2)	596	0.24	0.41	09	02	07	06	07	−06	19	30	13	10	04	−17	00	−47	(84)
12. Graphical FI (P2)	596	0.12	0.32	*11**	*07*	*04*	−08	07	−08	06	18	26	−06	02	−17	−08	−20	−06
13. Computer FI (P2)	596	0.18	0.37	*09*	*04*	*12**	12	06	16	24	−01	06	−13	−12	−28	−20	−53	−13
14. Public FI (P1)	596	0.11	0.30	*20**	*17**	*17**	06	11	−30	−03	34	55	08	16	−21	−04	−06	−07
15. Group FI (P2)	596	0.06	0.22	*14**	*13**	*17**	−02	11	−20	−07	31	45	11	02	−12	04	−03	12
16. FI frequency	548	3.83	2.75	−07	07	13*	**15***	−29	42	18	−24	−11	−37	−24	23	−26	−07	−09
17. Task novelty (P3)	597	3.67	1.52	00	01	03	02	00	21	04	−17	−16	−04	−07	09	05	18	−19
18. Task complexity (P3)	597	3.07	1.03	−01	−08	−08	−11	08	−11	01	02	06	−06	−03	−08	08	13	−08
19. Time constraint (P3)	596	0.55	0.48	*−11**	*00*	*03*	08	−27	00	−13	10	01	03	09	28	−13	06	−06
20. Time duration (P3)	444	−0.30	3.08	*21**	*12*	*20**	04	25	−25	18	29	43	13	−02	−39	00	−21	19
21. Creativity (P3)	597	1.59	0.86	*14**	*15**	*09*	07	13	−27	−20	12	02	18	01	−04	14	22	−05
22. Quantity–quality	464	1.36	0.76	08	12	07	04	−15	19	23	−17	−18	−23	−22	16	−13	−02	17
23. Performance rating	597	0.10	0.29	*17**	*15**	*15**	03	12	−25	−05	23	26	04	−01	−17	−04	00	02
24. Transfer measure	597	0.23	0.41	*−27**	*−08*	*−05*	−04	−40	25	−29	−34	−18	06	−14	62	04	39	−27
25. Latency measure	597	0.11	0.30	−01	−01	07	09	−07	06	07	−11	−02	−10	−07	01	−08	−15	−23
26. Physical task	597	0.12	0.32	**−16***	**−23***	**−29***	**−28***	04	−25	−22	09	−02	−08	12	11	24	07	−02
27. Reaction time	597	0.04	0.20	01	−01	02	04	04	−05	−01	05	22	−09	00	−11	−07	−14	−09
28. Memory task	597	0.20	0.38	04	16*	10	**14***	−13	14	15	−24	−13	−03	−04	12	−02	06	−06
29. Knowledge task	597	0.46	0.48	02	05	03	08	−05	17	11	−01	−19	−08	−16	−02	−17	−01	20
30. Following rules	597	0.23	0.41	01	−07	−07	**−16***	03	−11	06	01	24	07	07	−04	16	00	09
31. Vigilance task	597	0.19	0.37	−08	01	08	11	−16	15	−05	−14	−03	−06	02	20	−07	05	−06
32. Goal setting (P4)	552	0.08	0.27	*11**	*08*	*16**	11	08	−24	−12	28	36	03	08	−15	06	−13	25
33. Threat: Self-esteem (P1)	597	2.88	1.26	−08	−10	−09	**−14***	−23	−08	−09	11	06	00	−03	19	−16	02	12
34. Rewards–punishments	597	2.12	1.33	*12**	*05*	*02*	−08	18	−26	−07	13	22	21	10	−26	08	−13	10
35. Experimental control	500	0.78	0.42	*−13**	*−06*	*−09*	−01	−26	33	06	−33	−23	−15	−04	30	−12	13	−13
36. Lab–field	562	0.93	0.26	*−17**	*−15**	*−17**	−04	−09	25	08	−30	−49	−08	−04	17	−01	07	−11

Note. Decimals were omitted for all correlations. A series of boldface numbers indicates that the effect is significant regardless of exclusions. Single boldface of effect sizes; d = correlations with d; d' = correlations with d, excluding Mikulincer's studies; d'' = correlations with trimmed d'; d''' = correlations with d'', excluding
* $p < .01$.

tors that have been commonly discussed in the literature were included as well for the sake of completeness. All of these moderators can be classified into four groups: FI cues, task characteristics, situational variables, and methodological variables.

Each effect was rated on the putative moderators by two graduate students. Before the final ratings of the data, the raters were familiarized with the rating scales and practiced rating independently until they reached acceptable interjudge correlations (see the diagonal of Table 1). The correlations among the putative moderators (Table 1) suggest that the moderators are not independent, which call for caution in interpreting the results. For each moderator, the average rating of the two raters was used to predict the effect size (d). The number (K) of valid ratings for each moderator is given in the first column of Table 1. A summary version of the moderator definitions given to the raters is presented below. The definitions are followed, where appropriate, by the predictions of FIT including the relevant proposition number (noted in parentheses with the letter P). All other moderators are presented without predictions. The moderators are presented in the order that they were given to the raters.

FI Cues

FI sign. This is a measure of how positive or negative, on a 7-point scale, the FI was for the average recipient, where 1 is *very negative*, 4 *mixed*, and 7 *very positive*. Note that some studies manipulated task difficulty and therefore changed the proportion of participants that received positive–negative FI. FIT suggests that we cannot predict the effect of an FI sign without an a priori knowledge of how the feedback-standard gap is likely to be resolved (see *The Limitation of the Feedback-Standard Comparisons Argument* from earlier).

FI content. The following 14 measures where rated on a yes(1)–no(0) scale and reflected the type and format of information present in the FI: (a) *Correct–incorrect* feedback; (b) *Correct solution* should augment FIs effects on performance because it may help to reject erroneous hypotheses (P2); (c) *Attainment level* (number or things produced) should attenuate FI effects on performance because of its ambiguity for assessing the feedback-standard gap (see P4); (d) *Velocity* (change from previous trials) should augment FI effects because it directs attention to the task-motivation level and has a clear reference (P2); (e) *Normative information* (comparison with others, e.g., grades), should debilitate performance because it directs attention to meta-task processes such as the self and away from the task (P1); (f) *Norms* or information about the performance of others (P1); (g) FIs *designed to discourage* (P1; i.e., the experimenter designed a destructive message or cues that discouraged the recipient); (h) FIs *designed to praise* directs attention to

12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
(89)																								
−07	(88)																							
33	−06	(83)																						
33	02	65	(83)																					
−13	27	−20	−15	(85)																				
01	07	−16	−01	−12	(79)																			
06	07	10	21	−19	38	(70)																		
−25	07	−11	−15	14	04	−12	(88)																	
41	11	62	70	−20	−12	27	−52	(99)																
07	−22	19	22	−26	12	29	01	18	(67)															
−20	06	−08	−04	22	−01	−16	−09	01	−11	(64)														
40	−08	40	44	−26	01	13	−19	58	50	08	(84)													
−10	−26	−18	−14	25	03	−24	34	−42	03	−10	−19	(96)												
−05	37	−12	−10	26	26	08	28	−10	−17	−11	−13	06	(74)											
08	−18	−03	−07	−20	03	32	−25	−03	−03	−20	−10	−20	−13	(97)										
−06	31	−05	−06	11	00	−15	17	−11	−15	−04	−08	−12	40	−08	(90)									
−15	−01	−11	−14	15	−13	−14	08	−26	−08	07	−19	27	−05	−16	−12	(76)								
−24	−05	00	−10	07	−04	−13	−09	−08	06	30	−06	03	−17	−35	−21	−24	(87)							
23	03	22	33	−30	21	35	−18	53	−02	−07	31	−21	09	24	−12	−17	−43	(87)						
−01	01	−15	−12	32	−14	−13	40	−28	−16	−10	−16	22	01	−17	13	27	−44	−10	(79)					
13	−04	32	41	−10	−13	01	−06	41	22	−06	37	−09	−07	−02	−03	−15	00	16	−08	(77)				
09	−08	28	28	−09	−06	21	−18	39	17	02	26	−09	−07	09	−12	−23	18	20	−25	17	(65)			
27	08	47	49	−19	−27	14	−21	66	20	−07	41	−24	−10	−10	−14	−16	02	30	−15	32	46	(72)		
−26	12	−34	−44	25	14	−04	30	−57	−40	13	−38	23	16	02	08	13	−06	−03	20	−38	−25	−52	(78)	
−35	03	−71	−78	17	15	−18	21	−78	−26	08	−45	16	11	00	06	16	10	−32	14	−45	−38	−63	53	(86)

numbers indicate that the effect becomes significant after exclusions. Italic numbers indicate that the effect disappears after exclusions (may be an artifact). *K* = number time series effects; FI = feedback intervention; P# = propositions.

meta-task processes (P1); (i) FIs *provided verbally* may direct attention to meta-task processes because of the salience of the FI provider (P1); (j) FIs *provided in writing* should augment FI effects because it is less likely to invoke meta-task processes (P2); (k) FIs *provided graphically* (same as j; P2); (l) FIs *provided from a computer* (P2); (m) FIs made *public* should debilitate FI effects because it may direct attention to meta-task processes (P1); (n) FIs referred to a *group* performance should augment performance because it diverts attention away from the self (P2).

FI frequency. This is a measure of the number of times FI was provided, that is, the number of FI episodes for each effect. This variable was positively skewed. To normalize it, we used a square-root transformation. FI frequency was reported to augment FI effects on performance (Ilgen et al., 1979; Salmoni et al., 1984). Conditions however under which FI frequency debilitate performance are also known (e.g., higher relative FI frequency impairs performance on learning transfer of motor tasks; Salmoni et al., 1984; frequent outcome FI impairs cognitive consistency; see a discussion of MCPL studies above and also in Ilgen et al., 1979). Therefore, we did not have a clear prediction regarding the simple moderating effect of this variable.

Task Characteristics

Novelty. This is a measure of the subjective novelty of the *rules* of the task on a 7-point scale, where 1 means *familiar rules* and 7 *novel rules.* Unless FI is directed at learning (most of the FI effects in the meta-analysis are not), task novelty should attenuate the effect of FI (P3).

Complexity. This measures the objective level of task complexity, reflecting the number of actions, the dependencies among actions, and the temporal dependencies of actions needed for successful task performance. The complexity is a characteristic of the task independent of the participants who perform it (Wood, 1986) and was rated on a 7-point scale, where 1 means *very simple* and 7 *very complex.* The effect of complexity should be similar to the effect of novelty (P3).

Time constraint. A yes(1)-no(0) measure regards the imposition of time constraint on performance (i.e., all participants were required to perform the task during the same time). Time constraint may increase the role of intelligence in observed performance and therefore diminish any motivational effects of FI (P3).

Task duration. This measures time in weeks, days, hours, or minutes of the duration of the task reflected in the performance measure. We used a log transformation of task duration to normalize this variable. If the FI is not directed at learning, task duration should augment the effect of FI (P3).

Creativity. This measures the degree to which successful performance requires creativity on a 7-point scale, where 7 means *high need for creativity.* Creativity should have the same effect as novelty (P3).

Quality–quantity. Performance measures reflecting quality were coded 1, whereas measures reflecting quantity measurement were coded 0.

Ratings–objective performance. This measure indicates whether the performance is based on objective data (0) or on performance ratings (1). FIT has no predictions.

Transfer. A transfer task (yes-1; no-0) is where the effect of FI on one task was measured on another tasks.

Latency. This measure indicates whether the performance reflects latency or speed (1) or not (0).

In the absence of task taxonomy, for the following task moderators, FIT has no predictions.

Task type. The following six measures where rated on a yes(1)–no(0) scale and reflected the task type: (a) *physical* task, (b) *reaction time* task, (c) *memory* task, (d) *knowledge* task, (e) *following rules* task, and (f) *vigilance* task.

Situational and Methodological Variables

Goal setting. Some studies manipulated goal setting in conjunction with FI. In these goal-setting studies, both the FI and control groups received the same type of goal-setting manipulation. Hard and difficult goals where coded 1, whereas moderate, easy, "do your best," and no goals were coded 0 (five levels of goal setting were collapsed because of a small proportion of the studies that contained goal-setting and FI interventions). Goal setting should increase attention to the task-motivation level and therefore augment FI effects (P4).

Threat to self-esteem. This measures the degree that the level of performance may have psychological consequences for the participant. On the low end (1), there are very little consequences. For example, participants said that the purpose of the task is to calibrate the task. On the high end (7) are situations in which participants may perceive the task as a reflection of their intelligence, their career prospects, and alike. This is an additional situational cue that may direct attention to meta-task processes (P1).

External rewards–punishments. This measures the degree that the level of performance may have personal consequences for the participant on a 7-point scale, where 1 means *no consequences* and 7 means *meaningful consequences.*

Experimental control. True randomized experiments (1) were contrasted with quasiexperiments (0).

Lab–field. Laboratory participants (usually students) were studied outside of their normal environments (1) or field participants (workers or students) were studied in their normal environment (0). Kopelman (1986) reported that field studies tend to have more positive FI effects.

Eliminated Variables

Several other potential moderators were considered and measured as well. These were not included in any of the analyses because of the low frequency of either the occurrence of the moderator or information regarding the moderator. These moderators are discussed here for the sake of completeness.

Process–outcome. Process FI conveys information about *how* one performs the task (not necessarily how well). Outcome FI conveys information regarding *how well* one performs the task. Only six effects were available for process FI.

Performance reliability. Only two studies reported reliability estimates for the performance measure.

Analyses

Moderator analyses is typically done by comparing the mean of d for one value of a putative moderator with the mean of d in another level of the moderator (Hunter & Schmidt, 1990). For example, we could compare the average d in studies that provided positive FI with the average d in studies that provided negative FI. However, for many putative moderators, such an analysis would amount to splitting a continuous predictor at its mean and thus drastically reducing statistical power (e.g., Stone-Romero & Anderson, 1994). Therefore, we used the two approaches. First, we determined whether a variable significantly correlates with d, and for the significant moderators, we also reported the values of d for the extreme levels of the moderator. These procedures both circumvent the loss of power problem and yet present the results in a second way that aids interpretation. For the sake of consistency, we reported all moderator effects as correlations, even when the moderators are dichotomies. Hence, a positive correlation means that when the moderator has a high value, the estimated effect of FI is greater than average ($d = .41$, when all effects are included). If the moderator is a dummy code, a positive correlation means that when the condition of the dummy code is met (e.g., the FI provided a correct solution), the effect of FI is greater than the mean; when the condition is not met (e.g., the FI did not provide a correct solution), the effect of FI is smaller than the mean. The converse is true for negative correlations.

Before searching for moderators, few very extreme outliers were capped, such that if $d > 4.5$, then it was set to $d = 4.5$. This was done because a search for moderators in a meta-analysis is prone to second-order sampling error (Hunter & Schmidt, 1990). In addition, three more threats to second-order sampling error existed in the data. First, a violation of the assumption of independence was apparent in the data. For example, the correlation of FI sign with d is .24 in all the studies but only .09 without all the studies by Mikulincer. Second, despite the capping, outliers still produced spurious effects and masked some others. Finally, the quasi-d effects (the 17 time series effects) have unknown statistical properties. Therefore, we analyzed the correlations of each moderator once with all the data (d) and then with each threat removed at the time.

Excluding the Mikulincer studies reduced the sample by 91 effects (d'). Then we trimmed 5% of the data from each end of the distribution of d (46 effects; Hunter & Schmidt, 1990)[10] and Mikulincer's studies. Many of the outliers were Mikulin-

[10] "Tukey (1960) and Huber (1980) recommended deletion of the most extreme 10% of data points—the largest 5% and the smallest 5% of values" (Hunter & Schmidt, 1990, p. 207).

cer's studies, and therefore removing both threats from data entailed deleting only 20 extreme effects on top of the 91 Mikulincer's effects (d''). Whereas trimming may appear to commit an undesirable range restriction, the removal of outliers is likely to remove error from the data, as evidenced below in the few cases where the moderators appear stronger in the trimmed data. Finally, we excluded the quasi-*d* effects (the 17 time series effects) from the latter set to assess the impact of these effects on our conclusions (d''').

Overall, 470 effect sizes survived all exclusions (d''') and contained enough information to be rated on at least one moderator. Of these effects, 32% were negative. The average FI effect was .38 with a variance of .45 (drastically reduced because of the trimming of the outliers), whereas the expected variance was .09. Again, this suggests both that the estimate of FI mean effect is robust and that even for the trimmed data set, most of the variances cannot be accounted for by sampling error.

In the absence of a precise significance test—because of the violation of the assumption of independence—a conservative approach was taken to estimate significance for the moderators, that is, a Type I error was set to .01, and all tests were two tailed. Therefore, if a moderator correlated significantly at $p < .01$ with *d*, especially if it survived all the exclusions, it was considered robust or "significant."

Results

All four FIT propositions tested here received some support from the moderator analyses (Table 1). However, the various exclusions influenced the results and suggest four types of moderators: moderators that were significant regardless of exclusions, moderators that were significant after all the exclusions, moderators that are significant before the exclusions but turned nonsignificant after the exclusions, and nonsignificant moderators. Each class of these moderators suggests some unique conclusions.

First, four moderators showed systematic relationship with *d* regardless of the exclusions and suggest that discouraging FIs attenuate FI effects (consistent with P1), that velocity FIs and correct solution FIs augment FI effects (consistent with P2a and P2b), and that FI effects on performance of physical tasks are lower than FI effects on other tasks—a finding that we did not anticipate.

Second, six moderators became significant after all the exclusions. Because the exclusions were made to remove sources of error from the data, these six effects are likely to reflect true effects. These moderators suggest that praise, FIs threatening self-esteem, and verbal FIs attenuate FI effects (consistent with P1) that FIs with frequent messages augment FI effects (not predicted), and that FI effects are stronger for memory tasks and weaker for following rules tasks (not predicted). In addition, three moderators almost reached our significance criteria of .01 (i.e., $ps < .05$): Computerized FI yielded stronger FI effects (consistent with P2); FIs on complex tasks yielded weaker effects (P3); and FIs were more effective with a goal-setting intervention (P4). Yet, these effects should be treated with extra caution because of the reasons that led us to set alpha at .01 above.

Third, the exclusions rendered several moderators nonsignificant. First, the exclusion of Mikulincer's studies that violated the assumption of independence rendered, most notably, an FI sign and transfer measures nonsignificant. Second, the exclusion of outliers increased some estimates slightly (e.g., velocity FIs and discouraging FI), suggesting that outliers attenuated some estimates. Finally, five effects were rendered nonsignificant as a result of excluding the quasi-*d* effects (the time series): public FI, group FI, time duration, performance rating, and lab versus field studies.

To facilitate the interpretation of the effects that survived all exclusions, the weighted mean and variance of *d* for each level of each significant (and marginally significant) moderators are presented in Table 2. For each significant variable, Table 2 shows the mean and variance of *d* for two extreme levels. Miss-

Table 2
Feedback Intervention (FI) Effects by Levels of Significant FI Moderators After All Exclusions

Moderator	K	$\bar{d}$	σ_d
Correct solution (P2)			
Yes	114	.43	.38
No	197	.25	.44
Velocity (P2)			
Yes	50	.55	.46
No	380	.28	.40
Discouraging FI (P1)			
Yes	49	−.14	.52
No	388	.33	.37
Praise (P1)			
Yes	80	.09	.38
No	358	.34	.39
Verbal FI (P1)			
Yes	194	.23	.40
No	221	.37	.42
Computer FI (P2)			
Yes	87	.41	.40
No	337	.23	.42
FI frequency			
Top quartile	97	.32	.31
Bottom quartile	171	.39	.34
Task complexity (P3)			
Top quartile	107	.03	.46
Bottom quartile	114	.55	.39
Physical task			
Yes	65	−.11	.39
No	378	.36	.38
Memory task			
Yes	43	.69	.54
No	357	.30	.39
Following rules			
Yes	100	.19	.52
No	320	.36	.37
Goal setting (P4)			
Yes	37	.51	.40
No	373	.30	.45
Threat to self-esteem (P1)			
Top quartile	102	.08	.30
Bottom quartile	170	.47	.48

Note. Levels of dichotomous moderators are presented only for effects for which both judges agreed on the ratings. Levels for continuous moderators are presented for the approximate extreme quartile groups, where the categorization was done at the closest available point to the quartile value. K = number of effect sizes; $\bar{d}$ = N-weighted mean effect size; σ_d = variances; P# = propositions.

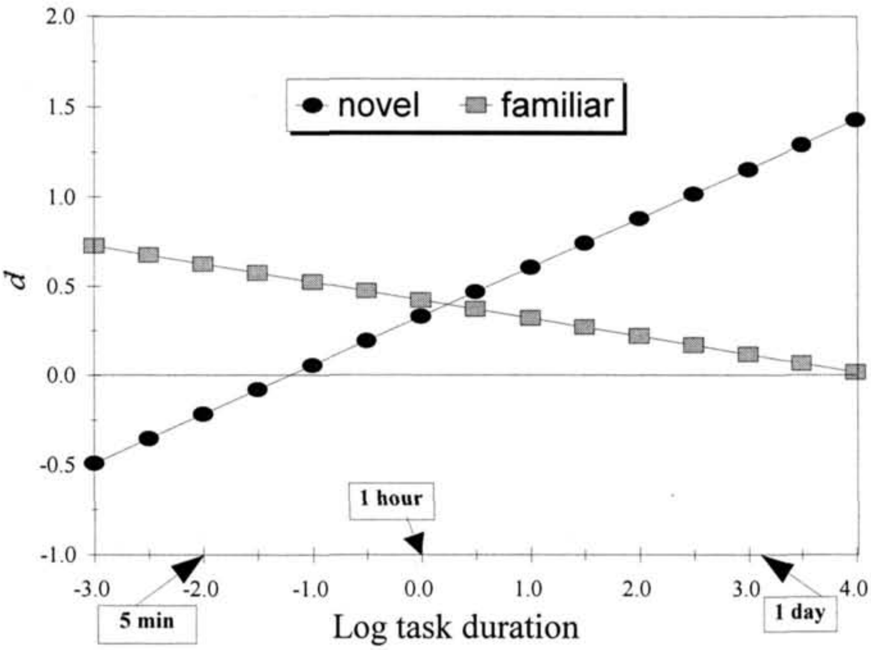

Figure 6. Regression-predicted values of d as a function of task duration for low (1) and high (7) novelty tasks.

ing effects sizes are, as in Table 1, a result of missing information in the original paper. For dichotomous moderators, only effects that were unanimously rated as either "yes" (1) or "no" (0) by the judges were used. (In the correlational analyses, disagreement was averaged to .50.) For the continuous moderators, the effects of the top and bottom quartile are presented. Quartile membership is approximate and was determine by the closest available empirical value to the quartile location. Table 2 merely reflects the results of Table 1. Only in the case of FI frequency are the results in the opposite direction, reflecting the poor distribution of this variable that is already transformed to normalize it. Therefore, FI frequency effect may be an artifact. Table 2 also suggests that even within each level of the moderators, there is a large portion of unexplained variance of FI effects.

Last, several predicted moderators were nonsignificant (e.g., task novelty and normative FIs). Some of the nonsignificant effects may reflect three-way interactions that are consistent with FIT. Tests of three-way interaction on our data suffered from an excessive Type II error because of second-order sampling error, the poor split of several dichotomous moderators, and the low statistical power of any complex interaction test. However, we tested two such interactions, where the distribution of predictors provided some variance. First, the lack of task novelty effect may be due to an interaction between novelty and task duration. Novel tasks become familiar across time, and hence the debilitating effect of FI on performance should disappear across time. Indeed the interaction between novelty and task duration was significant (see Figure 6) with both all studies (ΔR^2 due to the interaction = .026), $F(1, 443) = 12.3$, $p < .001$, and Mikulincer's studies excluded ($\Delta R^2 = .029$), $F(1, 352) = 10.7$, $p < .01$. Similar results were obtained with the trimmed data and with marginal results, once the time series effects were removed ($p < .07$; but see the disappearance of the main effect of time duration once the time series effects are removed, above). In all cases, the form of the interaction suggests that FIs debilitate performance on novel tasks performed for short duration but improve performance of tasks measured for a long time (consistent with P3), whereas the effect of FIs on performance of familiar tasks is not sensitive to the task duration.[11]

Also, goal setting had an interaction with attainment FIs ($\Delta R^2 = .021$), $F(1, 549) = 12.2$, $p < .001$, for all studies; ($\Delta R^2 = .036$), $F(1, 458) = 17.3$, $p < .001$, without Mikulincer's studies, but this interaction was marginal in the trimmed data sets ($ps < .06$). This interaction suggests that the FI effects are especially augmented by goal setting when the FI message is not interpretable (i.e., expressed in terms of attainment such as "you produced 200 units"; consistent with P4).

We performed two supplementary analyses to test the effect of the FI sign. First, we tested for curvilinear relationship between the FI sign and d. Second, we tested the effects of the FI sign on the ratio of the experimental group variance to the control group variance. Neither effect was significant.

Discussion and Conclusion

Our first two goals were to provide a historical review and to conduct a meta-analysis to demonstrate the large variability FI effects on performance. The historical review showed large and

[11] A Novelty × Complexity interaction was also tested but yielded nonsignificant results, which given the low reliability of the complexity measure is not surprising.

often ignored variability in FI effects. The meta-analysis showed, not surprising, that FIs improve performance by approximately .4 of a *SD* (a finding similar to a limited meta-analysis of FIs by Guzzo, Jette, & Katzell, 1985). However, we demonstrated a large variability of FI effects such that in over one third of the cases FIs reduced performance. Most of the observed variability cannot be explained by sampling or other errors. As such, it provides a strong empirical support for the conclusion of FI researchers who are identified with various theoretical approaches (see epigraph): FIs are double-edged swords (U.S. Congress, Office of Technology Assessment, 1987) because FIs do not always increase performance and under certain condition are detrimental to performance.

Our third goal was to account for some of the FIs variability through a preliminary theory—FIT. We showed that the preliminary FIT can integrate vast and seemingly unrelated literatures. Most important, FIT can encompass task-related learning, task-related motivation, and self-related and other meta-task processes. We believe that as such FIT is valuable in organizing the large bodies of relevant data and related theories, regardless of whatever evidence may presently exist for its validity.

Our last goal was to test the preliminary FIT through moderator analyses of the meta-analytic effects. These analyses provided partial support to FIT. Below we discuss the major conclusions that the moderator analyses suggest, the limitations and the contributions of FIT, the limitations of the meta-analysis, and the applied implications of our work.

Moderator Analyses: Major Conclusions

The moderator analyses suggest two major conclusions. First, several FI cues that seem to direct attention to meta-task processes *attenuate* FI effects on performance, whereas several FI cues that seem to direct attention to task-motivation or task-learning processes *augment* FI effects on performance. This pattern of findings provides reasonable support for the first two propositions. Specifically, both praise and FI designed to discourage were postulated to increase attention to meta-task processes and were found to attenuate FI effects. Furthermore, both the attenuating effect of praise and the nonsignificant effect of an FI sign (which is discussed later in this section) are not easily predicted by most FI-related theories. The debilitating effects of praise on performance received some direct experimental support both in the laboratory and in the field and were explained, respectively, by a model of self-attention (Baumeister et al., 1990) and by control theory (Waldersee & Luthans, 1994).

These findings are also consistent with a review of field studies (many of which did not qualify for the meta-analysis) that concluded that "praise may not be widely effective as a reinforcer" (Balcazar et al., 1985 p. 79). In addition, verbal FI that involves the saliency of another person was related to lower FI effects, whereas computerized FI that is likely to focus attention on the task was marginally related to higher FI effects. Also, FIs that threatened self-esteem had lower FI effectiveness. Furthermore, velocity FI that was postulated to direct attention to the task augmented FI effects on performance. According to FIT, velocity FIs create a very clear feedback-standard discrepancy at the task level. Finally, FIs that supply the correct solution were related to stronger FI effects. In summary, four out of seven variables testing P1 and three out of six variables testing P2 yielded significant results. None of the variables testing these proposition showed significant results opposite the predicted direction. This pattern of results is largely consistent with FIT's argument that, *ceteris paribus* (everything else being equal), FI cues affect performance by changes in locus of attention: The lower in the hierarchy the FI-induced locus of attention is, the stronger the benefit of an FI for performance.

The second major conclusion is that FI effects are moderated by the nature of the task. However, the exact task properties that moderate FI effects are still poorly understood. The third proposition of FIT suggested five variables that should moderate FI effects. None of these moderators showed a clear effect. Yet, three of the five variables were weakly or interactively related to FI-moderating effects. First, simple-task performance benefitted from FIs (marginally) more than complex-task performance. Task complexity had relatively low interjudge reliability (.70), reflecting perhaps the difficulty in conceptualizing task complexity (and other task dimensions) and therefore suggesting that the effect that we observed is an underestimate. Indeed, when we investigated the meaning of the weak correlational effect of task complexity with differences in mean FI effect between the extreme quartiles of task complexity (Table 2), a large effect of task complexity appeared. (Of course, this effect appears large because we looked at the extreme quartiles, yet it helps to demonstrate the implication of the weak correlation.) Second, the performance of novel tasks seemed to be debilitated when performance was measured for a short time (i.e., performance in the initial stages of task acquisition). This effect implicates meta-task processes that render the performance of subjectively complex tasks susceptible to interference from interventions such as FI. This conclusion is consistent with results regarding other interventions such as goal setting (Kanfer & Ackerman, 1989) and inappropriate task labeling (Vallacher & Wegner, 1987). Therefore, our findings provide only weak evidence for P3 that task type and its level of mastery play an important role in determining the effect of FI on performance.

Although the moderating effects of task features identified by FIT received weak support, several task dimensions moderated FI effects unexpectedly: Physical tasks and following rules tasks yielded weaker FI effects, and memory tasks yielded stronger FI effects. Our results strongly suggest that task type places a serious boundary condition on the knowledge of effectiveness of various interventions designed to improve performance (cf. Hammond, 1992). Therefore, the lack of a valid task taxonomy that can be used across vastly different tasks (e.g., vigilance, memory, and adherence to regulations) poses a serious obstacle for FI research. Moreover, even within similar types of tasks (MCPL), the "effects of feedback seem to be very sensitive to the task environment [difficulty]" (Castellan & Swaine, 1977, p. 118).

In addition, P4 received marginal support, that is, the effects of FI seem to be augmented by goal setting, especially when the FI contains information that is difficult to evaluate without an externally set goal. The effect of FI in addition to a goal-setting intervention appear weaker (it was marginal) than the effects implied in a literature survey of this question by Locke and Latham (1990; Figure 8-4). Locke and Latham showed that, almost invariably, adding goals to FI augment that FI effect. However, there are several features of our approach that explain the difference. First, we used partially nonoverlapping reports.[12] Second and most important, we tested the question of goal setting as a moderator of FI effects in a data set that contained mostly studies that did not manipulate goals at all (Table 2). In contrast, in their Figure 8-4, Latham and Locke directly compared goals having FI with goals alone. Therefore, our effect size of this moderator depends not only on the true effect of adding goals to FI but also on the small proportion of studies having goals alone as a control group. Therefore, our weak result here represents, we believe, more a weakness of the test of this moderator than divergence in conclusions.

In summary, the moderator analyses have lent reasonable support to the major propositions of FIT. The propositions are supported, however, only with an overall pattern of results and not with detailed evidence at the single-variable level. This type of support is what one can expect with the complexity of FI effects and the limitation of second-order sampling error, which has poor statistical power for testing three-way interactions. However, this type of support, we believe, justifies accepting FIT as a preliminary theory that now requires further validation with primary and detailed research. Such research should investigate both the processes suggested by the propositions and explore additional issues suggested by some of the nonsignificant moderators.

Several nonsignificant moderators are conspicuous, most notably FI sign. FI sign has a curvilinear effect on arousal (Kluger et al., 1994). Therefore, we also tested whether the square of FI sign would moderate the effect of FI on performance and found no such effect. The absence of an FI sign moderating effect does not mean that FI sign is inconsequential, however. According to FIT, FI sign signals a discrepancy, and therefore it is likely to receive attention and be acted on. Yet, at present, there is no FI-related theory that can predict a priori the effects of all the important moderators that determine how feedback sign affects performance. This difficulty is illuminated by FIT which suggests that every feedback-standard discrepancy can be resolved in several ways. The most investigated area of strategy choice in reducing a feedback-standard gap involves gaps monitored by meta-task processes. For example, in face of a negative FI sign and attention directed to the self, an internal attribution for the cause of the failure may result in choosing a strategy of disengagement from the task (B. Weiner, 1974). However, cues that foster high self-efficacy (Bandura, 1986) direct attention back to the task and cause people to invest more effort. The cues that affect the choice between types of discrepancy reduction may be available in the environment or may be derived from the personality disposition of the recipient.

Extensive research by Mikulincer has shown that personality variables such as attributional style (1988b), cognitive interference (1989a), and need for structure (Mikulincer, Yinon, et al., 1991) moderate the reaction to severe negative FI. Therefore, it seems that the general mechanism implicated in these findings is an ability to redirect attention to the task. This capacity may be related to an inhibition mechanism (Cropanzano et al., 1993; Lord & Levy, 1994) that protects the working memory space from competing demands on the system. A tendency to blame the self, think about the self, or doubt the self may interfere with the inhibition required for task performance and may "release" the actor from the pressure to eliminate feedback-standard discrepancy at the task level (Lord & Levy, 1994). These personality variables could not be included in our meta-analyses because most studies do not report the moderating effect of personality on the FI–performance relationship. However, personality should continue to be studied in future FIT investigations, if we are to understand the effects of FI sign.

Limitations

Limitations of the Preliminary FIT

FIT in its present form lacks very detailed and specific predictions. In this form, FIT runs the risk of being unfalsifiable. For example, we did not find an effect for normative FIs (which could be a function of measurement error). Should we therefore reject FIT altogether? We think that the judgment should be withheld until the processes implied by FIT are tested directly. For example, FIT implicates meta-task processes such as attention to the self, where attention to the self was not directly measured as a dependent variable. Future FIT research should test processes implied by FIT. An example of such a hypothesis is that FI directs attention to the self as a function of feedback-standard discrepancy at the task–process level. An extreme deviation from task standard (both above and below) may require the intervention of a higher process. Therefore, participants who receive FI similar to their task standards should show less attention to the self than those who receive very discrepant FI.

Limitations of the Meta-Analysis

Our empirical conclusions are based on a meta-analysis. Our meta-analysis exposed several limitations of both the literature and the technique. However, these limitations are extremely instructive in guiding future FI research. First, the existing FI literature contains only a meager proportion of studies that reported a well-controlled FI experiment. The most common reason for excluding papers from the meta-analysis was the lack of a control group. Unfortunately, many researchers of FIs still implicitly assume that FIs increase performance, and therefore they limit their studies to comparisons of several types of FIs (e.g., a positive vs. negative FI). Without control groups, we may know more about the relative merits of several types of FI messages, but we have no idea if they are better, equal, or inferior to no intervention. This state of affairs is alarming, although it was already noticed in the MCPL literature:

[12] We found six papers that compared FI having goals with goals alone that were not reported by Latham and Locke (1991). We did not include some of the studies they used in the meta-analysis because of two reasons. First, some studies did not report sufficient details to allow the computation of *d* and therefore could not be used in a meta-analysis. Second, two papers that Latham and Locke reviewed estimated the effect of FI by providing FI first and FI removal second. We did not include such effect sizes because they may reflect not the effect of feedback removal as much as the effect of removing the intervention per se, regardless of the intervention content.

> Results indicating that subjects who received no feedback . . . perform better than subjects who receive feedback certainly poses problems for MCPL investigators who have assumed that subjects were learning as a result of feedback and have consequently neglected no-feedback controls. (Schmitt, Coyle, & Saari, 1977, p. 326)

The practice of omitting a control group is especially disturbing when the researcher(s) strongly recommend FIs where there is no evidence for the benefits of the intervention. We hope that the results of this meta-analysis will convince future FI researchers to include a control group. The lack of control groups in the present literature may bias our results to an unknown degree.

Second, removing Mikulincer's studies, which created a violation of the assumption of independence from the meta-analysis, had a small effect on the estimate of the mean FI effect but a large effect on the estimate of the presence of a few moderators. This suggests that other unknown violations of the assumption of independence may exist in our data and bias the results to an unknown degree. It also raises a question regarding the uniqueness of Miklincer's studies. Apart from the effect of removing Miklincer's studies on FI sign discussed earlier, it changed the estimates of FI frequency from a (nonsignificant) negative effect to a positive effect, and it changed the estimate of task transfer measure from very negative to nonsignificant. These variables reflect the paradigm of learned helplessness experiments: Frequent and negative FIs on one task and measurement of performance on another serves as the experimental tool to induce and measure learned helplessness. Learned helplessness is "*the experience of uncontrollable failure to solve a problem originally perceived as solvable* [sic]" (Mikulincer, 1994, p. 13). Neither FI sign nor frequency are theoretically essential to induce uncontrollability. However, it is difficult to experimentally create a sense of uncontrollability in other ways (e.g., with random success; Mikulincer, 1994). Therefore, we can conclude that while a repeated negative FI *can* be detrimental to performance, it is not necessarily detrimental to performance (cf. Mikulincer, 1994) and that other FI moderators can induce performance deficit (e.g., praise).

Third, the systematically positive FI effects that were found among the 17 time series studies suggest that the typical exclusion of the time series effects from the meta-analyses may again bias the conclusions to an unknown degree. The time series effects were largely obtained from the behavior modification tradition (e.g., Komaki et al., 1980). Removal of these effects rendered several moderators nonsignificant, thus identifying the unique characteristics of the behavior modification time series studies of FI: field studies that measured rated performance for a long time and provided public and group FIs. It would be however a mistake to conclude that these moderators are artifacts or that behavior modification always yields positive FI effects (Balcazar et al., 1985). Rather, we cannot determine their unique role with the present methodology and view these findings as a source for hypotheses that need to be tested in the future.

Contributions

The limitation of our study should be evaluated in the light of its contributions. Although recent writings repeatedly suggested that the effects of FI on performance are not automatic and far from being understood (Balcazar et al., 1985; Ilgen et al., 1979; Locke & Latham, 1990; Salmoni et al., 1984), we provided, through FIT, a positive identification of several moderators. Moreover, our empirical results may provide the necessary information to mitigate the persistent and unwarranted belief that FI always improves performance (e.g., Pritchard et al., 1988).

We see three major types of explanations for the persistence of this view: psychological, economical, and theoretical. First, feedback is psychologically reassuring, and people like to obtain feedback, although they may refrain from seeking it when the cost (social and otherwise) of obtaining it is prohibitive (Ashford & Cummings, 1983). Otherwise, people appear to seek feedback about their performance even if it does not affect their performance (Kluger & Adler, 1993; Kluger, Adler, & Fay, 1992). Furthermore, when costs are low, people may seek feedback repeatedly, even when they are informed that it cannot benefit their performance (Kroll, Levy, & Rapoport, 1988). This suggests that FI about task performance is often used to assess progress toward goals of the self (meta-task processes),[13] which may be independent from the goal of achieving an objectively superior performance. Therefore, we believe that researchers and practitioners alike confuse their feelings that feedback is desirable with the question of whether FI benefits performance. This may indicate that FIs have benefits other than positive effects on immediate performance. For example, feedback may increase satisfaction (Fried & Ferris, 1987) and may contribute to long-range persistence on the focal task—a variable that was not studied in the FI literature.

Second, many intervention techniques that are sold by practitioners in educational and work settings are based on the assumption that FIs improve performance, so there is an economic incentive for some practitioners to continue to laud the positive effects of FI. An early example is Pressy's (1950) self-scoring device, which was largely beneficial for learning but yielded detrimental effects as well. Clearly, those who have a financial stake in the assumption that FI always improves performance would have very little interest in carefully testing this assumption.

Finally, we believe that the major culprit is a lack of a general theory regarding the effects of FI on performance. Without a comprehensive theory, there is no way to integrate the vast and inconsistent empirical findings. Furthermore, the vague nature of most views about how FI works are not likely to lead to any theoretical revisions because they are not articulated well enough to begin with. Without a clear theoretical expectation, it is not possible to interpret findings as posing serious contradictions, and thus there is no progress. Even worse, in Popper's (as cited in Robinson, 1986) terminology, the vague hypotheses about FI were unfalsifiable and, hence, nonscientific. We believe that with appropriate caution, FIT offers a remedy for this situation.

---

[13] A. Rapoport (personal communication, March 13, 1995) suggested that people are inclined to seek feedback repeatedly even when it appears useless to verify that the environment remained stable. Therefore, in the long run, such a strategy may be adaptive.

Implications

Before we conclude, we must reflect on the applied implication of our study. The identification of a number of moderators suggests that in certain situations, FI can yield a large and positive effect on performance. Specifically, an FI provided for a familiar task, containing cues that support learning, attracting attention to feedback-standard discrepancies at the task level (velocity FI and goal setting), and is void of cues to the meta-task level (e.g., cues that direct attention to the self) is likely to yield impressive gains in performance, possibly exceeding 1 *SD*. However, even such an FI is not always an efficient intervention. Even when FI has considerable positive effects on performance, its utility may be marginal or even negative. When an FI increases performance through an increase in task motivation, the effect may depend on a continuous FI. Removal of such an FI may result in a reversal as some field studies have demonstrated (Komaki et al., 1980). Therefore, the cost of maintaining a continuous intervention should be considered in evaluating such an intervention. If, however, FI affects performance through task-learning processes, the effect may create only shallow learning and interfere with more elaborate learning. Lack of elaborate learning affects the ability to use the learned material in transfer tasks where the task is similar but not identical (e.g., Carroll & Kay, 1988). Moreover, the evidence for any learning effect here was minimal at best. Indeed, in the MCPL literature, several reviewers doubt whether FIs have any learning value (Balzer et al., 1989; Brehmer, 1980) and suggest alternatives to FI for increasing learning, such as providing the learner with more task information (Balzer et al., 1989). Another alternative to an FI is designing work or learning environments that encourage trial and error, thus maximizing learning from task feedback without a direct intervention (Frese & Zapf, 1994). These considerations of utility and alternative interventions suggest that even an FI with demonstrated positive effects on performance should not be administered whenever possible. Rather, additional development of FIT is needed to establish the circumstance under which positive FI effects on performance are also lasting and efficient and when these effects are transient and have questionable utility. This research must focus on the processes induced by FIs and not on the general question of whether FIs improve performance—look at how little progress 90 years of attempts to answer the latter question have yielded.

References

Studies preceded by an asterisk were included in the meta-analysis.

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3–27.

Adams, J. A. (1978). Theoretical issues for knowledge of results. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 229–240). New York: Academic Press.

Adelman, L. (1981). The influence of formal, substantive and contextual task properties on the relative effectiveness of different forms of feedback in multiple-cue probability learning tasks. *Organizational Behavior and Human Performance, 27,* 423–442.

Albright, M. D., Levy, P. E., & Williams, J. R. (1992). *Reactions to feedback discrepancies from multiple raters.* Paper presented at the 7th annual convention of the Society for Industrial and Organizational Psychology, Montreal, Quebec, Canada.

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *Journal of General Psychology, 54,* 279–299.

*Anderson, D. C., Crowell, C. R., Doman, M., & Howard, G. S. (1988). Performance posting, goal setting, and activity-contingent praise as applied to a university hockey team. *Journal of Applied Psychology, 73*(1), 87–95.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94,* 192–210.

Anderson, K. J. (1990). Arousal and the inverted-U hypothesis: A critique of Neiss's "Reconceptualizing Arousal." *Psychological Bulletin, 107,* 96–100.

Anderson, S., & Rodin, J. (1989). Is bad news always bad?: Cue and feedback effects on intrinsic motivation. *Journal of Applied Social Psychology, 19,* 449–467.

Annett, J. (1969). *Feedback and human behaviour.* Harmondsworth, Middlesex, England: Penguin Books.

*Anshel, M. H. (1987). The effect of mood and pleasant versus unpleasant information feedback on performing a motor skill. *Journal of General Psychology, 115*(2), 117–129.

*Arkin, R. M., & Schumann, D. W. (1984). Effects of corrective testing: An extension. *Journal of Educational Psychology, 76,* 835–843.

*Arkin, R. M., & Walts, E. A. (1983). Performance implications of corrective testing. *Journal of Educational Psychology, 75,* 561–571.

Arps, G. F. (1920). Work with knowledge of results versus work without knowledge of results. *Psychological Monographs, 28,* 1–41.

Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance, 32,* 370–398.

*Austin, M. F., & Grant, T. N. (1981). Interview training for college students disadvantaged in the labor market: Comparison of five instructional techniques. *Journal of Counseling Psychology, 28*(1), 72–75.

Azuma, H., & Cronbach, L. J. (1966). Cue-response correlations in the attainment of a scalar concept. *American Journal of Psychology, 79,* 38–49.

Babad, E. (1990). Measuring and changing teachers' differential behavior as perceived by students and teachers. *Journal of Educational Psychology, 82,* 683–690.

*Baechie, C. L., & Lian, M. J. (1990). The effects of direct feedback and practice on metaphor performance in children with learning disabilities. *Journal of Learning Disabilities, 23*(7), 451–455.

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of performance feedback. *Journal of Organizational Behavior Management, 7,* 65–89.

Balzer, W. K., Doherty, M. E., & O'Connor, R., Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106,* 410–433.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1991). Social cognitive theory of self regulation. *Organizational Behavior and Human Decision Processes, 50,* 248–287.

*Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45,* 1017–1028.

*Baron, R. A. (1988). Negative effects of destructive criticism: Impact on conflict, self-efficacy and task performance. *Journal of Applied Psychology, 73*(2), 199–207.

Baumeister, R. F., Hutton, D. G., & Cairns, K. J. (1990). Negative effects of praise on skilled performance. *Basic and Applied Social Psychology, 11,* 131–148.

*Bellucci, G., & Hoyer, W. J. (1975). Feedback effects on the performance of self-reinforcing behavior of elderly and young adult women. *Journal of Gerontology, 30*(4), 456–460.

*Bethge, H., Carlson, J. S., & Wiedl, K. H. (1982). The effects of dynamic assessment procedures on Raven Matrices performance, visual search behavior, test anxiety and test orientation. *Intelligence, 6,* 89–97.

*Betz, N. E., & Weiss, D. J. (1976). *Effects of immediate knowledge of results and adaptive testing on ability test performance* (Research Rep. No. 76-3). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.

Bilodeau, E. A. (1969). Supplementary feedback and instructions. In E. A. Bilodeau (Ed.), *Principles of skill acquisition* (pp. 235–253). New York: Academic Press.

*Bilsky, L. H., Gilbert, L., & Pawelski, C. E. (1978). Facilitation of class-inclusion performance in mildly retarded adolescents: Feedback and strategy training. *American Journal of Mental Deficiency, 83*(2), 177–184.

Bless, H., Bohner, G., Schwarz, N., & Strack, F. (1990). Mood and persuasion: A cognitive response analysis. *Personality and Social Psychology Bulletin, 16,* 331–345.

*Boggiano, A. K., & Barrett, M. (1985). Performance and motivational deficits of helplessness: The role of motivational orientations. *Journal of Personality and Social Psychology, 49,* 1753–1761.

Book, W. F., & Norvell, L. (1922). The will to learn, an experimental study of incentives in learning. *The Pedagogical Seminary, 29,* 305–362.

*Brabender, V., & Boardman, S. K. (1977). Sex differences in self-confidence as a function of feedback and social cues. *Psychological Reports, 41,* 1007–1010.

Brand, J. E. (1905). The effect of verbal suggestion upon the estimation of linear magnitudes. *Psychological Review, 12,* 41–49.

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica, 45,* 223–241.

*Bridgeman, B. (1974). Effects of test score feedback on immediately subsequent test performance. *Journal of Educational Psychology, 66,* 62–66.

Brown, F. J. (1932). Knowledge of results as an incentive in schoolroom practice. *Journal of Educational Psychology, 23,* 532–552.

*Bustamante, J., Moreno, P., Rehbein, L., & Vizueta, A. (1980). Effects of feedback and reinforcement in tachistoscopic training on a fault-detection task. *Perceptual and Motor Skills, 51,* 987–993.

*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.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58,* 1–14.

*Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology, 78,* 210–216.

*Calpin, J. P., Edelstein, B., & Redmon, W. K. (1988). Performance feedback and goal setting to improve mental health center staff productivity. *Journal of Organizational Behavior Management, 9*(2), 35–58.

Campion, M. A., & Lord, R. G. (1982). A control systems conceptualization of the goal-setting and changing process. *Organizational Behavior and Human Performance, 30,* 265–287.

*Carroll, J. M., & Kay, D. S. (1988). Prompting, feedback and error correction in the design of a scenario machine. *International Journal of Man–Machine Studies, 28,* 11–27.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self regulation: A control theory to human behavior.* New York: Springer-Verlag.

Carver, C. S., & Scheier, M. F. (1990). Origins and function of positive and negative affect: A control-process view. *Psychological Review, 97,* 19–35.

Castellan, N. J., Jr., & Swaine, M. (1977). Long term feedback and differential feedback effects in non-metric multiple-cue probability learning. *Behavioral Science, 22,* 116–128.

Casti, J. L. (1989). *Alternate realities: Mathematical models of nature and man.* New York: Wiley.

*Chhokar, J. S., & Wallin, J. A. (1984). A filed study of the effect of feedback frequency on performance. *Journal of Applied Psychology, 69*(3), 524–530.

Christianson, S. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112,* 284–309.

*Chung, K. H., & Dean, V. (1976). Relative effectiveness and joint effects of three selected reinforcements in a repetitive task situation. *Organizational Behavior and Human Performance, 16,* 114–142.

*Church, R. M., & Camp, D. S. (1965). Change in reaction-time as a function of knowledge of results. *American Journal of Psychology, 78,* 102–106.

*Clark, C. R., Guskey, T. R., & Benninga, J. S. (1983). The effectiveness of mastery learning strategies in undergraduate education courses. *Journal of Educational Research, 76*(4), 210–214.

Crafts, L. W., & Gilbert, R. W. (1935). The effect of knowledge of results on maze learning and retention. *Journal of Educational Psychology, 26,* 177–187.

Crawley, S. L. (1926). An experimental investigation of recovery from work. *Archives of Psychology, 13*(85).

Cropanzano, R., James, K., & Citera, M. (1993). A goal hierarchy model of personality, motivation, and leadership. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 15, pp. 267–322). Greenwich, CT: JAI Press.

Cusella, L. P. (1987). Feedback motivation and performance. In F. M. Jablin, L. L. Putnam, K. H. Robertsand, & L. W. Pooter (Eds.), *Handbook of organizational communication* (pp. 624–678). Newbury Park, CA: Sage.

*DeGregorio, M., & Fisher, C. D. (1988). Providing performance feedback: Reactions to alternative methods. *Journal of Management, 14*(4), 605–616.

Deputy, E. C. (1929). Knowledge of success as a motivating influence in college work. *Journal of Educational Resources, 20,* 327–334.

*Dossett, D. L., Latham, G. P., & Mitchell, T. R. (1979). Effects of assigned versus participatively set goals, knowledge of results, and individual differences on employee behavior when goal difficulty is held constant. *Journal of Applied Psychology, 64*(3), 291–298.

Duval, T. S., Duval, V. H., & Mulilis, J. (1992). Effects of self-focus, discrepancy between self and standard, and outcome expectancy favorability on the tendency to match self to standard or to withdraw. *Journal of Personality and Social Psychology, 62,* 340–348.

*Earley, P. C. (1986). Trust, perceived importance of praise and criticism, and work performance: An examination of feedback in the United States and England. *Journal of Management, 12*(4), 457–473.

Earley, P. C. (1988). Computer-generated performance feedback in the magazine-subscription industry. *Organizational Behavior and Human Decision Processes, 41,* 50–64.

Earley, P. C., Connolly, T., & Ekegren, G. (1989). Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology, 74,* 24–33.

Easterbrook, J. A. (1959). The effect of emotion of cue utilization and the organization of behavior. *Psychological Review, 66,* 183–201.

Elwell, J. L., & Grindley, G. C. (1938–1939). The effect of knowledge of results on learning and performance. *British Journal of Psychology, 39,* 39–54.

*Erez, M. (1977). Feedback: A necessary condition for the goal setting–performance relationship. *Journal of Applied Psychology, 62,* 624–627.

Eysenck, M. W. (1992). *Attention and arousal: Cognition and performance.* Berlin: Springer-Verlag.

*Fischer, G. W. (1982). Scoring-rule feedback and the overconfidence syndrome in subjective probability forecasting. *Organizational Behavior and Human Performance, 29,* 352–369.

Forgas, J. P. (1993). On making sense of odd couples: Mood effects of the perception of mismatched relationships. *Personality and Social Psychology Bulletin, 19,* 59–70.

Forgas, J. P., Bower, G. H., & Moylan, S. J. (1990). Praise or blame? Affective influences on attributions for achievement. *Journal of Personality and Social Psychology, 59,* 809–819.

*Foushee, H. C., Davis, M. H., Stephan, W. G., & Bernstein, W. M. (1980). The effects of cognitive and behavioral control on post-stress performance. *Journal of Human Stress, 6,* 41–48.

*Fowler, B. (1981). The aircraft landing test: An information processing approach to pilot selection. *Human Factors, 23*(2), 129–137.

Frese, M., & Stewart, J. (1984). Skill learning as a concept in life-span developmental psychology: An action theoretic analysis. *Human Development, 27,* 145–162.

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 4, pp. 271–340). Palo Alto, CA: Consulting Psychologists Press.

Fried, Y., & Ferris, G. R. (1987). The validity of the job characteristics model: A review and meta-analysis. *Personnel Psychology, 40,* 287–322.

*Fulmer, R. S., & Rollings, H. E. (1976). Item-by-item feedback and multiple choice test performance. *Journal of Experimental Education, 44,* 30–32.

Ganzach, Y. (1994). Feedback representation and prediction strategies. *Organizational Behavior and Human Decision Processes, 59,* 391–409.

*Gardner, R. M., Sandoval, Y., & Reyes, B. (1986). Signal detection analysis of recognition memory of obese subjects. *Perceptual and Motor Skills, 63,* 227–234.

Garfield, E. (1983). How to use citation analysis for faculty evaluation, and when is it relevant? Pt. 2. *Current Contents, 45,* 5–14.

Gates, A. I. (1917). *Recitation as a function in memorizing* (*Archives of Psychology* Monograph No. 40). New York: Columbia University Press.

*Geen, R. G. (1981). Effects of being observed on persistence at an insolvable task. *British Journal of Social Psychology, 20,* 211–216.

Gilliland, A. R. (1925). The effect of practice with and without knowledge of results in grading handwriting. *Journal of Educational Psychology, 16,* 532–536.

*Glover, J. A. (1989). Improving readers' estimates of learning from text: The role of inserted questions. *Reading Research and Instruction, 28*(3), 68–75.

Gollwitzer, P. M., Heckhausen, H., & Ratajczak, H. (1990). From weighing to willing: Approaching a change decision through pre- or postdecisional mentation. *Organizational Behavior and Human Decision Processes, 45,* 41–65.

*Goltz, S. M. (1990). *Investigating self, task and supervisory feedback effects* (Research Paper Series 90-01). Notre Dame, IN: University of Notre Dame, Center for Research in Business.

*Goltz, S. M., Citera, M., Jensen, M., Favero, J., & Komaki, J. L. (1989). Individual feedback: Does it enhance effects of group feedback? *Journal of Organizational Behavior Management, 10*(2), 77–92.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology, 38,* 275–291.

Hammond, K. R. (1992). Onward. *Organizational Behavior and Human Decision Processes, 53,* 278–283.

Hammond, K. R., & Summers, D. A. (1972). Cognitive control. *Psychological Review, 79,* 58–67.

*Hammond, K. R., Summers, D. A., & Deane, D. H. (1973). Negative effects of outcome–feedback in multiple-cue probability learning. *Organizational Behavior and Human Performance, 9,* 30–34.

*Hanna, G. S. (1976). Effects of total and partial feedback in multiple-choice testing upon learning. *Journal of Educational Research, 69,* 202–205.

Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancies effects: 31 meta-analyses. *Psychological Bulletin, 97,* 363–386.

Higgins, T. E. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319–340.

*Hill, R. P., & Ward, J. C. (1989). Mood manipulation in marketing research: An examination of potential confounding effects. *Journal of Marketing Research, 26,* 97–104.

Hogarth, R. M., Gibbs, B. J., McKinzie, C. R. M., & Marquis, M. A. (1991). Learning from feedback: Exactingness and incentives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 734–752.

*Hornblum, J. N., & Overton, W. F. (1976). Area and volume conservation among the elderly: Assessment and training. *Developmental Psychology, 12,* 68–74.

Hsee, C. K., & Abelson, R. P. (1991). Velocity relation: Satisfaction as a function of the first derivative of outcome over time. *Journal of Personality and Social Psychology, 60,* 341–347.

*Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin, 107,* 328–340.

*Hundal, P. S. (1969). Knowledge of performance as an incentive in repetitive industrial work. *Journal of Applied Psychology, 53*(3), 224–226.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Hyland, M. E. (1987). Control theory interpretation of psychological mechanisms of depression: Comparison and integration of several theories. *Psychological Bulletin, 102,* 109–121.

Ilgen, D. R. (1971). Satisfaction with performance as a function of the initial level of expected performance and the deviation from expectations. *Organizational Behavior and Human Performance, 6,* 345–361.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organization. *Journal of Applied Psychology, 64,* 349–371.

Ilgen, D. R., & Hamstra, B. W. (1972). Performance satisfaction as a function of the difference between expected and reported performance at five levels of reported performance. *Organizational Behavior and Human Performance, 7,* 359–370.

Irion, A. L. (1969). Historical introduction. In E. A. Bilodeau (Ed.), *Principles of skill acquisition* (pp. 1–31). New York: Academic Press.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 203–253). New York: Academic Press.

*Ivancevich, J. M. (1982). Subordinates' reactions to performance appraisal interviews: A test of feedback and goal-setting techniques. *Journal of Applied Psychology, 67*(5), 581–587.

*Ivancevich, J. M., & McMahon, J. T. (1982). The effects of goal setting, external feedback, and self-generated feedback on outcome variables: A field experiment. *Academy of Management Journal, 25*(2), 359–372.

Jacoby, J., Mazursky, D., Troutman, T., & Kuss, A. (1984). When feedback is ignored: Disutility of outcome feedback. *Journal of Applied Psychology, 69*(3), 531–545.

Johanason, A. M. (1922). The influence of incentive and punishment on reaction time. *Archives of Psychology, 8,* 53.

Jones, G. M. (1910). Experiments on the reproduction of distance as influenced by suggestions of ability and inability. *Psychological Review, 17,* 269–278.

Judd, C. H. (1905). Practice without knowledge of results. *Psychological Review Monographs, 2*(7, Suppl.), 185–198.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integration/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74,* 657–690.

Kernan, M. C., & Lord, R. G. (1991). An application of control theory to understanding the relationship between performance and satisfaction. *Human Performance, 4,* 173–185.

*Kim, J. S., & Hamner, W. C. (1976). Effect of performance feedback and goal setting on productivity and satisfaction in an organizational setting. *Journal of Applied Psychology, 61*(1), 48–57.

*Kistner, J. A. (1985). Attentional deficit of learning-disabled children: Effects of rewards and practice. *Journal of Abnormal Child Psychology, 13*(1), 19–31.

*Kluger, A. (1989, April). *Negative and positive effects of feedback on motivation: The role of control and individual differences.* Paper presented at the 4th annual convention of the Society for Industrial and Organizational Psychology, Boston, MA.

Kluger, A. N., & Adler, S. (1993). Person- versus computer-mediated feedback. *Computers in Human Behavior, 9,* 1–16.

*Kluger, A. N., Adler, S., & Fay, C. (1992). Computerized feedback effects on feedback seeking, performance and motivation. In M. I. Nurminen & G. R. S. Weir (Eds.), *Human jobs and computer interfaces* (pp. 131–145). Amsterdam: North-Holland.

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.

*Knight, J. L., & Salvendy, G. (1981). Effects of task feedback and stringency of external pacing on mental load and work performance. *Ergonomics, 24*(10), 757–764.

Koestner, R., Zuckerman, M., & Koestner, J. (1987). Praise, involvement, and intrinsic motivation. *Journal of Personality and Social Psychology, 53,* 383–390.

*Komaki, J. L., Collins, R. L., & Penn, P. (1982). The role of performance antecedents and consequences in work motivation. *Journal of Applied Psychology, 67*(3), 334–340.

*Komaki, J., Heinzmann, A. T., & Lawson, L. (1980). Effect of training and feedback: Component analysis of a behavioral safety program. *Journal of Applied Psychology, 65*(3), 261–270.

Kopelman, R. (1986). Objective feedback. In E. A. Locke (Ed.), *Generalizing from the laboratory to field settings* (pp. 119–146). Lexington, MA: Lexington Books.

Korsgaard, A., Meglino, B. M., & Lester, S. W. (1994, August). *The virtue of being altruistic: The role of the value of helping and concern in individuals' reactions to feedback from others.* Paper presented at the 1994 meeting of the Academy of Management, Dallas, TX.

*Kratochwill, T. R., & Brody, G. H. (1976). Effects of verbal and self-monitoring feedback on Wechsler Adult Intelligence Scale performance in normal adults. *Journal of Consulting and Clinical Psychology, 44,* 879–880.

Kroll, Y., Levy, H., & Rapoport, A. (1988). Experimental tests of the separation theorem and the capital asset pricing model. *American Economic Review, 78,* 500–519.

Kuhl, J. (1992). A theory of self-regulation: Actions vs. state orientation, self-discrimination, and some applications. *Applied Psychology: An International Review, 41,* 97–129.

*Lacher, M. B. (1983). Effects of feedback, instruction, and initial performance level upon training and persistence of verbal rehearsal. *Journal of General Psychology, 108,* 43–54.

*Larson, J. R., Jr., & Callahan, C. (1990). Performance monitoring: How it affects work productivity. *Journal of Applied Psychology, 75,* 530–538.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes, 50,* 212–247.

*Latta, R. M. (1976). Differential tests of two cognitive theories of performance: Weiner versus Kukla. *Journal of Personality and Social Psychology, 34,* 295–304.

*Latta, R. M. (1978). Interactive effects of initial achievement orientation and prior success feedback on the mastery of subsequent difficult and easy tasks. *American Educational Research Journal, 15*(1), 17–24.

Lazarus, R. S. (1991). Progress on a cognitive–motivational–relational theory of emotion. *American Psychologist, 46,* 819–834.

Lee, J. W., & Yates, J. F. (1992). How quantity judgment changes as the number of cues increases: An analytical framework and review. *Psychological Bulletin, 112,* 363–377.

*Leherissey, B. L., O'Neil, H. F., & Hansen, D. N. (1970). *Effects of memory support on state anxiety and performance in computer-assisted learning* (Contract No. N00014-68-A-0494). Washington, DC: U.S. Office of Naval Research, Psychological Sciences Division, Personnel & Training Research Programs.

Lepper, M., & Gurtner, J. L. (1989). Children and computers: Approaching the twenty-first century. *American Psychologist, 44,* 170–178.

*Levy, J. (1974). Social reinforcement and knowledge of results as determinants of motor performance among EMR children. *American Journal of Mental Deficiency, 78*(6), 752–758.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. M. Hunt (Ed.), *Personality and the behavior disorders* (pp. 333–377). New York: Ronald Press.

Lewinsohn, S., & Mano, H. (1993). Multi-attribute choice and affect: The influence of naturally occurring and manipulated moods on choice processes. *Journal of Behavioral and Decision Making, 6,* 33–51.

*Lewis-Beck, J. A. (1978). Locus of control, task expectancies, and children's performance following failure. *Journal of Educational Research, 71,* 207–210.

*Light, K. C., & Obrist, P. A. (1980). Cardiovascular response to stress: Effects of opportunity to avoid, shock experience, and performance feedback. *Psychophysiology, 17*(3), 243–252.

*Lindahl, L. G. (1945). Movement analysis as an industrial training method. *Journal of Applied Psychology, 29,* 420–436.

Lindell, M. K. (1976). Cognitive and outcome feedback in multiple-cue probability learning tasks. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 739–745.

Locke, E. A. (1967). Motivational effects of knowledge of results: Knowledge or goal settings? *Journal of Applied Psychology, 51,* 324–329.

*Locke, E. A., & Bryan, J. F. (1968). Goal-setting as a determinant of the effect of knowledge of score on performance. *American Journal of Psychology, 79,* 398–406.

Locke, E. A., & Bryan, J. F. (1969). Knowledge of scores and goal level as determinants of work rate. *Journal of Applied Psychology, 53,* 59–65.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice Hall.

*Locke, E. A., & Shaw, K. N. (1984). Atkinson's inverse-U curve and the missing cognitive variables. *Psychological Reports, 55,* 403–412.

Lord, R. G., & Levy, P. E. (1994). Moving from cognition to action: A control theory perspective. *Applied Psychology: An International Review, 43,* 335–366.

*Lucas, R., Heimstra, N., & Spiegel, D. (1973). Part-task simulation training of drivers' passing judgments. *Human Factors, 15*(3), 269–274.

*Lyhle, K. G., & Kulhavy, R. W. (1987). Feedback processing and error correction. *Journal of Educational Psychology, 79,* 320–322.

Mace, C. A. (1935). *Incentives: Some experimental studies* (Medical Research Council, Industrial Health Research Board, Rep. No. 72). London: Her Majesty's Stationery Office.

Mano, H. (1991). Assessing the structure and intensity of noninduced and induced affect. *Multivariate Behavioral Research, 26,* 389–411.

Mano, H. (1992). Judgments under distress: Assessing the role of unpleasantness and arousal in judgment formation. *Organizational Behavior and Human Decision Processes, 52,* 216–245.

*Manzer, C. W. (1935). The effect of knowledge of output on muscular work. *Journal of Experimental Psychology, 8,* 80–90.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253.

*McClintock, C. G., & Van Avermael, E. (1975). The effects of manipulating feedback upon children's motives and performance: A propositional statement and empirical evaluation. *Behavioral Science, 20,* 101–116.

*McCormack, P. D., Binding, F. R. S., & Chylinski, J. (1962). Effects on reaction-time of knowledge of results of performance. *Perceptual and Motor Skills, 14,* 367–372.

*McCormack, P. D., Binding, F. R. S., & McElheran, W. G. (1963). Effects on reaction time of partial knowledge of results of performance. *Perceptual and Motor Skills, 17,* 279–281.

*McFarlin, D. B., Baumeister, R. F., & Blascovich, J. (1984). On knowing when to quit: Task failure, self-esteem, advice, and nonproductive persistence. *Journal of Personality, 52*(2), 138–155.

*Mesch, D. J., Farh, J., & Podsakoff, P. M. (1989). *Effects of feedback sign on group goal setting, strategies, and performance: An empirical examination of some control theory hypotheses.* Unpublished manuscript, Indiana University, Bloomington.

Midrash Rabbah: Ecclesiastes. (1939). In H. Freedman & M. Simon (Eds.) and A. Cohen (Trans.), *Midrash Rabbah* (p. 39, ch. 1, verse 13). London: Soncino Press.

*Mikulincer, M. (1988a). A case study of three theories of learned helplessness: The role of test importance. *Motivation and Emotion, 12,* 371–383.

*Mikulincer, M. (1988b). Reactance and helplessness following exposure to unsolvable problems: The effects of attributional style. *Journal of Personality and Social Psychology, 54,* 679–686.

*Mikulincer, M. (1989a). Cognitive interference and learned helplessness: The effects of off-task cognitions on performance following unsolvable problems. *Journal of Personality and Social Psychology, 57,* 129–135.

*Mikulincer, M. (1989b). Coping and learned helplessness: Effects of coping strategies on performance following unsolvable problems. *European Journal of Personality, 3,* 181–194.

*Mikulincer, M. (1989c). Learned helplessness and egotism: Effects of internal/external attribution on performance following unsolvable problems. *British Journal of Social Psychology, 28,* 17–29.

*Mikulincer, M. (1990). Joint influence of prior beliefs and current situational information on stable and unstable attributions. *Journal of Social Psychology, 130,* 739–753.

Mikulincer, M. (1994). *Human learned helplessness: A coping perspective.* New York: Plenum Press.

Mikulincer, M., Glaubman, H., Ben-Artzi, E., & Grossman, S. (1991). The cognitive specificity of learned helplessness and depression deficits: The role of self-focused cognitions. *Anxiety Research, 3,* 273–290.

*Mikulincer, M., Kedem, P., & Zilkha-Segal, H. (1989). Learned helplessness, reactance and cue utilization. *Journal of Research in Personality, 23,* 235–247.

*Mikulincer, M., & Nizan, B. (1988). Causal attribution, cognitive interference, and the generalization of learned helplessness. *Journal of Personality and Social Psychology, 55,* 470–478.

*Mikulincer, M., Yinon, A., & Kabili, D. (1991). Epistemic needs and learned helplessness. *European Journal of Personality, 5,* 249–258.

*Montague, W. E., & Webber, C. E. (1965, April). Effects of knowledge of results and differential monetary reward on six uninterrupted hours of monitoring. *Human Factors,* 173–180.

*Morgan, C. L., & Morgan, L. V. (1935). Effects of immediate awareness of success and failure upon objective examination scores. *Journal of Experimental Education, 4*(2), 63–66.

*Murray, J., & Epstein, L. H. (1981). Improving oral hygiene with videotape modeling. *Behavior Modification, 5*(3), 360–371.

Murray, N., Sujan, H., Hirt, E. R., & Sujan, M. (1990). The influence of mood on categorization: A cognitive flexibility interpretation. *Journal of Personality and Social Psychology, 59,* 411–425.

Nadler, D. (1977). *Feedback and organizational development: Using data-based methods.* Reading, MA: Addison-Wesley.

*Neenan, D. M., & Routh, D. K. (1986). Response cost, reinforcement, and children's Perteus Maze qualitative performance. *Journal of Abnormal Child Psychology, 14*(3), 469–480.

Neiss, R. (1990). Ending arousal's reign of error: A reply to Anderson. *Psychological Bulletin, 107,* 101–105.

*Newman, M. I., Williams, R. G., & Hiller, J. H. (1974). Delay of information feedback in an applied setting: Effects on initially learned and unlearned items. *Journal of Experimental Education, 42*(4), 55–59.

*Oner, N. P. (1977). Impact of teacher behavior and teaching technique on learning by anxious children. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety* (Vol. 4, pp. 171–188). Washington, DC: Hemisphere.

*Onyehalu, A. S. (1983). Feedback and performance of Piagetian conservation tasks in a developing country. *American Journal of Psychology, 96*(1), 65–73.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge, England: Cambridge University Press.

*Ozolins, D. A., & Anderson, R. P. (1980). Effects of feedback on the vigilance task performance of hyperactive and hypoactive children. *Perceptual and Motor Skills, 50,* 415–424.

Parkinson, B., & Manstead, A. S. R. (1992). Appraisal as a cause of emotion. In M. S. Clark (Ed.), *Emotions: Review of personality and social psychology* (pp. 122–149). Newbury Park, CA: Sage.

*Payne, R. B., & Hauty, G. T. (1955). Effect of psychological feedback upon work decrement. *Journal of Experimental Psychology, 50,* 343–352.

Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology, 71,* 211–218.

*Pearson, C. A. L. (1991). An assessment of extrinsic feedback on participation, role perceptions, motivation, and job satisfaction in a self-managed system for monitoring group achievement. *Human Relations, 44*(5), 517–537.

*Perkins, V. L. (1988). Feedback effects on oral reading errors of children with learning disabilities. *Journal of Learning Disabilities, 21*(4), 244–248.

*Philips, C. W., Paterson, C. E., & Pettijohn, T. F. (1984). Feedback conditions and motor skills performance. *Journal of General Psychology, 112*(1), 53–57.

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.

Powers, W. T. (1973, January 26). Feedback: Beyond behaviorism. *Science, 179*, 351–356.

Pressy, 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.

*Prestwood, J. S., & Weiss, D. J. (1978). *The effects of knowledge results and test difficulty on ability test performance and psychological reactions to testing* (Research Rep. No. 78-2). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.

*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.

*Reber, R. A., & Wallin, J. A. (1984). The effects of training, goal setting, and knowledge of results on safe behavior: A component analysis. *Academy of Management Journal, 27*(3), 544–560.

*Rebok, G. W., & Balcerak, L. J. (1989). Memory self-efficacy and performance differences in young and old adults: The effect of mnemonic training. *Developmental Psychology, 25*, 714–721.

*Rees, P. J. (1986). Do medical students learn from multiple choice examinations? *Medical Education, 20*, 123–125.

*Reid, L., Lefebvre-Pinard, M., & Pinard, A. (1988). Generalization of training speaking skills: The role of overt activity, feedback, and child's initial level of competence. *Perceptual and Motor Skills, 66*, 963–978.

*Ringel, B. A., & Springer, C. J. (1980). On knowing how well one is remembering: The persistence of strategy use during transfer. *Journal of Experimental Child Psychology, 29*, 322–333.

Robinson, D. (1986). *An intellectual history of psychology.* Madison: University of Wisconsin.

*Rocklin, T., & Thompson, J. M. (1985). Interactive effects of test anxiety, test difficulty, and test feedback. *Journal of Educational Psychology, 77*, 368–372.

*Rosenbaum, M., & Ben-Ari, K. (1985). Learned helplessness and learned resourcefulness: Effects of noncontingent success and failure on individuals differing in self-control skills. *Journal of Personality and Social Psychology, 48*, 198–215.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Ross, C. C. (1933). The influence of achievement of a knowledge of progress. *Journal of Educational Psychology, 24*, 609–619.

Rothstein, H. G. (1986). The effects of time pressure on judgment in multiple cue probability learning. *Organizational Behavior and Human Decision Processes, 37*, 83–92.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 36*, 1152–1168.

Russell, J. A. (1991). Categories of emotions. *Psychological Bulletin, 110*, 426–450.

*Rust, J. O., Strang, H. R., & Bridgeman, B. (1977). How knowledge of results and goal setting function during academic tests. *Journal of Experimental Education, 45*, 52–55.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450–461.

*Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology, 45*, 736–750.

Salmoni, 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.

*Schloss, P. J., Wisniewski, L. A., & Cartwright, G. P. (1988). The differential effect of learner control and feedback in college students' performance on CAI modules. *Journal of Educational Computing Research, 4*(2), 141–150.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47*, 1173–1181.

*Schmitt, N., Coyle, B. W., & Saari, B. B. (1977). Types of task information feedback in multiple-cue probability learning. *Organizational Behavior and Human Performance, 18*, 316–328.

*Schuldt, W. J., & Bonge, D. (1979). Effects of self-imposition and experimenter imposition of achievement standards on performance. *Psychological Reports, 45*, 119–122.

Seifert, T. L. (1991). Determining effect sizes in various experimental designs. *Educational and Psychological Measurement, 51*, 341–347.

*Shavit, H., & Rabinowitz, A. (1978). Locus of control and effects of failure on performance and perceived competence. *Journal of Personality Assessment, 42*(3), 265–271.

*Simensen, R. J. (1973). Acquisition and retention of a motor skill by normal and retarded students. *Perceptual and Motor Skills, 36*, 791–799.

Smith, F. D. (1933). Repetition without knowledge of results as a factor in learning. *Psychological Bulletin, 30*, 673–674.

*Sobal, J., & Juhasz, J. B. (1977). Sex, experimenter, and reinforcement effects in verbal learning. *Journal of Social Psychology, 102*, 267–273.

*Sonnenschein, S. (1986). Developing referential communication: Transfer across novel tasks. *Bulletin of the Psychonomics Society, 24*(2), 127–130.

Spencer, L. T. (1923). Minor studies from the psychological laboratory of Yale University. *American Journal of Psychology, 34*, 107–111.

*Stein, J. R. (1976). Effects of feedback on performance of introverted and extraverted delinquents and non-delinquents. *Criminal Justice and Behavior, 3*(4), 371–378.

Steinman, D. O. (1974). Transfer of lens model training. *Organizational Behavior and Human Performance, 12*, 1–16.

Sternberg, R. J. (1992). *Psychological Bulletin*'s top 10 "hit parade." *Psychological Bulletin, 112*, 387–388.

*Stinson, M., & Belmont, J. M. (1979). Effects of feedback manipulation and feedback preference upon children's performance. *Journal of Genetic Psychology, 135*, 159–160.

Stone-Romero, E. F., & Anderson, L. E. (1994). Relative power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology, 79*, 354–359.

*Strang, H. R. (1983). The effects of knowledge of results upon reaction time performance. *Journal of General Psychology, 108*, 11–17.

*Strang, H. R., Lawrence, E. C., & Fowler, P. C. (1978). Effects of assigned goal level and knowledge of results on arithmetic computation: A laboratory study. *Journal of Applied Psychology, 63*(4), 446–450.

*Strang, H. R., & Rust, J. O. (1973). The effects of immediate knowledge of results and task definition on multiple-choice answering. *Journal of Experimental Education, 42*(1), 77–80.

Taylor, M. S., Fisher, C. D., & Ilgen, D. R. (1984). Individuals' reactions to performance feedback in organizations: A control theory perspective. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (pp. 81–124). Greenwich, CT: JAI Press.

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. Volume I: The original nature of man.* New York: Columbia University, Teachers College.

*Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39,* 212–222.

*Tindale, R. S., Kulik, C. T., & Scott, L. A. (1991). Individual and group feedback and performance: An attributional perspective. *Basic and Applied Social Psychology, 12*(1), 41–62.

*Titus, T. G. (1973). Continuous feedback in recognition memory. *Perceptual and Motor Skills, 37,* 771–776.

*Trowbridge, M. H., & Cason, H. (1932). An experimental study of Thorndike's theory of learning. *Journal of General Psychology, 7,* 245–258.

Tubbs, M. E., Boehne, D. M., & Dahl, J. G. (1993). Expectancy, valence, and motivational force functions in goal-setting research: An empirical test. *Journal of Applied Psychology, 78,* 361–373.

*Turner, R. R. (1975). Effect of information feedback on matrices learning for differing socioeconomic levels. *Journal of Educational Psychology, 67,* 285–295.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59*(4, Pt. 2), 251–278.

U.S. Congress, Office of Technology Assessment. (1987). *The electronic supervisor: New technology, new tensions* (Publication No. OTA-CIT-333). Washington, DC: U.S. Government Printing Office.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94,* 3–15.

*Vitulli, W. F. (1982). Effects of immediate feedback on computer-assisted testing of ESP performance. *Psychological Reports, 51,* 403–408.

*Wade, M. G., & Newell, K. M. (1972). Performance criteria for stabilimeter learning. *Journal of Motor Behavior, 4*(4), 231–239.

*Wade, T. C. (1974). Relative effects on performance and motivation of self-monitoring correct and incorrect responses. *Journal of Experimental Psychology, 103*(2), 245–248.

Waldersee, R., & Luthans, F. (1994). The impact of positive and corrective feedback on customer service performance. *Journal of Organizational Behavior, 15,* 83–95.

*Warm, J. S., Epps, B. D., & Ferguson, R. P. (1974). Effects of knowledge of results and signal regularity on vigilance performance. *Bulletin of the Psychonomic Society, 4*(4A), 272–274.

Waters, R. H. (1933). The specificity of knowledge of results and improvement. *Psychological Bulletin, 30,* 673.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98,* 219–235.

Weiner, B. (1974). An attributional interpretation of expectancy-value theory. In B. Wiener (Ed.), *Cognitive views of human motivation* (pp. 51–70). New York: Academic Press.

Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 233–237). New York: Academic Press.

*Wiener, E. L. (1974). An adaptive vigilance task with knowledge of results. *Human Factors, 16*(4), 333–338.

*Wiener, E. L. (1975). On simultaneous monitoring and tracking. *Journal of Applied Psychology, 60*(1), 100–105.

*Wiener, E. L., & Attwood, D. A. (1968). Training for vigilance: Combined cuing and knowledge of results. *Journal of Applied Psychology, 52*(6), 474–479.

*Winett, R. A., & Vachon, E. M. (1974). Group feedback and group contingencies in modifying behavior of fifth graders. *Psychological Reports, 34,* 1283–1292.

*Wise, S. L., Plake, B. S., Pozehl, B. J., Barnes, L. B., & Lukin, L. E. (1989). Providing item feedback in computer-based tests: Effects of initial success and failure. *Educational and Psychological Measurement, 49,* 479–486.

Wolf, F. M. (1986). *Meta-analysis: Quantitative method for research synthesis* (Sage University Papers Series of Quantitative Applications in the Social Sciences, Series No. 07-059). Beverly Hills, CA: Sage.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes, 37,* 60–82.

Wood, R. E., & Locke, E. A. (1990). Goal setting and strategy effects on complex tasks. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 12, pp. 73–109). Greenwich, CT: JAI Press.

Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects: A meta-analysis. *Journal of Applied Psychology, 72,* 416–425.

Wright, W. R. (1906). Some effects of incentive on work and fatigue. *Psychological Review, 13,* 23–34.

*Zohar, D., Cohen, A., & Azar, N. (1980). Promoting increased use of ear protectors in noise through information feedback. *Human Factors, 22*(1), 69–79.

Received October 11, 1993
Revision received April 27, 1995
Accepted April 30, 1995 ■