This preliminary study was designed to help two teachers who felt excessively self-critical to increase their positive self-thoughts and decrease their negative self-thoughts, i.e., to improve their self-esteem. An intensive experimental design involving four phases was employed. Each phase—Baseline, Thought Stopping, Positive Intervention, and Follow-Up—lasted approximately 2 weeks. During all phases of the experiment the teachers observed their positive and negative self-thoughts with the help of wrist counters. Classroom observers coded instances of overt positive and negative behavior by the teachers. A positive intervention, which used stimulus cues to prompt positive self-thoughts, was effective in increasing the positive thoughts of one teacher. The results of thought stopping (subvocalizing the word "stop" to inhibit unwanted thoughts) were mixed. The reported changes in self-thoughts were supported by two self-report measures and, to some extent, by the external observations of teacher and student classroom behavior. Some indication was found that self-observing does change the behavior being observed. (Authors)
CHANGING THE EVALUATIVE SELF-THOUGHTS OF TWO ELEMENTARY TEACHERS

James W. Hannum, Carl E. Thoresen, and David R. Hubbard, Jr.
Introductory Statement

The Center's mission is to improve teaching in American schools. Its work is carried out through five programs:

- Teaching Effectiveness
- The Environment for Teaching
- Teaching Students from Low-Income Areas
- Teaching and Linguistic Pluralism
- Exploratory and Related Studies

This paper reports on work in one area of the Program on Teaching Effectiveness.
Evidence suggests that internal events such as thoughts and images may follow the same principles of behavior as do external events (Bandura, 1969; Mahoney, 1971a). Thus attempts to modify internal or covert behavior should logically proceed in a manner similar to that of attempts to modify external or overt behavior. Concern with what may be termed covert or cognitive behavior modification has recently increased with efforts to examine a number of covert techniques (e.g., Thoresen & Mahoney, 1974). A growing number of studies have suggested that such modifications are possible (e.g., Mahoney, 1971b; Meichenbaum & Goodman, 1971; Todd, 1972) and that covert changes may lead to changes in overt behavior (Meichenbaum & Cameron, in press).

A personality construct long considered a significant determiner of a person's overt behavior is self-esteem. A person's attitude toward himself has been correlated with numerous other variables such as birth order, religious background, school achievement, and parents' self-esteem (Coopersmith, 1967; Rosenberg, 1965). If self-esteem is viewed as a class of covert behaviors concerned with a person's self-evaluations, then these self-evaluations may be altered by utilizing behavioral principles. Further, changes in internal behavior should presumably lead to identifiable changes in overt behavior. Clinical evidence (e.g., Rogers, 1961) suggests that changes in self-esteem may function as important antecedents to external behavior change, although

---

1 Now at the Counselling Centre, Memorial University, St. John's, Newfoundland, Canada.

2 Now at the School of Medicine, University of Connecticut.
some investigators (e.g., Hobbs, 1962) have argued that internal change, i.e., in self-esteem, probably follows rather than precedes overt change.

In this study a behavioral definition of self-esteem was developed drawing from the work of Marston (1965) and Homme et al. (1968). High self-esteem, or positive self-concept, was defined as a high frequency of positive (reinforcing) self-evaluations combined with a low frequency of negative (punishing) self-evaluations; low self-esteem (a negative self-concept) was viewed as the inverse, i.e., many negative evaluations and few positive ones. Hence changes in self-esteem could occur if self-thoughts (self-verbalizations) were modified directly. Of course, self-esteem can also be changed by altering the person's external environment.

Thought stopping (Cautela, 1969; Wolpe & Lazarus, 1966) was selected to reduce the frequency of negative self-thoughts, and a positive intervention derived from Premack's (1965) differential probability principle was used to increase positive self-thoughts. The effect of these interventions on other self-esteem indices and selected overt behaviors in the classroom was also explored.

Method

Subjects

The subjects were two elementary school teachers who volunteered to participate after hearing presentations by the authors at teacher's meetings. These teachers complained of being excessively critical of themselves and indicated that they would like to change this behavior. Both were women and both had had at least 16 years of teaching experience. Teacher One (T1) was 52 years old and Teacher Two (T2) was 55 years old. Two other teachers were originally involved in the study but were dropped due to a personal illness and a decision not to participate in the treatments.

3 One teacher decided that he could not carry out the requirements of the study, in particular with regard to self-observing and thought stopping. He expressed strong doubts that he could ever really change his "personality."
Procedure

The study was based on an intensive clinical design (N=1) using repeated measures of each subject during each phase (Chassan, 1967; Sidman, 1960); these measures were compared against the subject's own baseline. The experiment involved four separate phases, as follows.

Baseline. Each teacher was provided with two wrist counters and was asked to self-observe positive and negative self-thoughts daily during a specified hour. Lists of common self-thoughts were generated for each teacher during an initial interview to aid in the self-observation. The scope and quality of each teacher's self-evaluations were also discussed during this interview. An example of a positive self-thought (+ST) was "I'm patient with the children." An example of negative self-thought (-ST) was "I'm just too old for teaching."

Classroom observers recorded two categories of teacher behavior during this same hour throughout the study. The overt behaviors observed were positive responses, either verbal or nonverbal (e.g., saying "That's good," or smiling), and negative responses, also either verbal or nonverbal (e.g., saying "Stop doing that!" or shaking a finger at a student). Single instances of each behavior were recorded for each 10-second interval; the daily totals were converted to rates by dividing by the total intervals observed. Interobserver reliability was assessed on 10 separate days throughout the study by means of another observer. The reliabilities for these variables were .93 and .99 respectively.

At the end of each day teachers recorded the frequency of their self-thoughts in a special notebook and completed a short report form for listing positive and negative events they may have been thinking about that day apart from their classroom activities. These outside events consisted of physical events (e.g., I have a headache), family or social events (e.g., today is my son's birthday), or job related but outside-the-classroom events (e.g., I didn't get my raise yet).

The Gough Adjective Check List (Gough & Heilbrun, 1965) was administered to each teacher at the beginning of the study. Two scales
derived from this measure were used as indications of self-acceptance and self-criticism. The baseline phase continued for a minimum of two weeks.

Thought Stopping (TS). Only T1 participated in this intervention, which began immediately following the baseline phase. T2 was called out of town for a family emergency; her baseline was therefore extended following her return (Baseline II) and thought stopping was not used. In thought stopping, the person learns to subvocalize the word "Stop" to reduce unwanted thoughts. Success in reducing the frequency of unwanted obsessive thoughts has been reported by Yamagami (1971), Cautela (1969), and others (cf. Wolpe & Lazarus, 1966). T1 received one thought-stopping session with a trainer on the first day of the phase and a short check-up contact several days later. She was asked to use the technique on her negative self-thoughts for a two-week period. She also continued to monitor the frequency of her positive self-thoughts.

Positive Intervention (PI). Both teachers participated in this intervention, which was designed to increase positive self-thoughts. The intervention was based in part on Premack's (1965) differential probability principle whereby behaviors that often occur can be used to increase less frequent behaviors. High frequency behaviors that each teacher performed in the classroom were identified (e.g., looking at the wall clock). A recognizable stimulus cue was developed (a small blue decal) and was placed (near the clock) to remind the teacher to think a positive self-thought before completing the high frequency behavior. This sequence of cueing a positive self-thought and engaging in a high probability response was expected to reinforce the positive self-thoughts. Clinical success in using variations of these procedures to increase positive self-thoughts has been reported (e.g., Jackson, 1972; Mahoney, 1971b; Todd, 1972). Each teacher received one session of instruction on the first day of this phase and a short check-up several days later. This technique was used over a two-week period. During this time the teachers continued to monitor their negative self-thoughts.
Follow-up. The positive intervention was withdrawn and the teachers continued to self-observe for 5 or 6 days. Following this period, several interviews were held with each teacher to assess her reactions to the study. The Gough Adjective Check List was readministered.

Results

The positive and negative self-thought data for T1 as well as the daily rates for positive and negative overt behavior are presented in Figure 1. Data for T2 are presented in Figure 2.

The daily thought frequencies and daily rates for positive and negative overt behaviors were analyzed for between phase mean changes using a one-way analysis of variance. Since the analysis of variance model rests upon assumptions of independence of the data, all variables with significant analysis of variance results were subsequently tested for the presence of a significant linear trend and lag 1 autocorrelation within each phase (Rao, 1967). These latter tests were performed to determine if the nonindependence of the data was a serious factor in the analysis of variance results (Thoresen & Elashoff, in press).

The data from the teacher’s positive and negative daily reports were dichotomized into days on which a report was made and those for which no report was made. A comparison of the changes in these frequencies between phases for both positive reports and negative reports was done using the Fisher Exact Probability Test (Siegal, 1956). The pre and post scores on the Gough Adjective Check List were converted to Self-Acceptance (total favorable adjectives checked over total adjectives checked) and Self-Criticism (total unfavorable adjectives checked over total adjectives checked) scores. These raw scores were then converted to percentile scores using the population values provided by Crowne, Stephens, and Kelly (1961).

Teacher One

T1 showed significant increases in positive self-thoughts ($p < .001$); the major change in means occurred between the thought stopping and positive intervention phases (see Fig. 1). Unexpectedly, positive self-
Fig. 1. Self-observation data and overt classroom responses plotted against days of observation for Teacher One. The horizontal lines represent the phase means.
Fig. 2. Self-observation data and overt classroom responses plotted against days of observation for Teacher Two. The horizontal lines represent the phase means.
TABLE 1

Analysis of Variance for Positive and Negative Self-Thoughts and Positive and Negative Overt Responses Between and Within Intervention Phases: Teacher One

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive self-thoughts</td>
<td>Between phases</td>
<td>3</td>
<td>520.1</td>
<td>43.46</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>36</td>
<td>12.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative self-thoughts (log)</td>
<td>Between phases</td>
<td>3</td>
<td>7.0</td>
<td>23.59</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>35</td>
<td>.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive overt responses</td>
<td>Between phases</td>
<td>3</td>
<td>.0007</td>
<td>.58</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>38</td>
<td>.0012</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative overt responses</td>
<td>Between phases</td>
<td>3</td>
<td>.0031</td>
<td>6.20</td>
<td>&lt;.005</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>38</td>
<td>.0005</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

thoughts (+ST) decreased during thought stopping to a mean of 4.9 per day from a baseline mean of 14.7 per day. Negative self-thoughts (-ST) decreased significantly (p < .001), with the major decrease in means occurring between the baseline and TS phases. The first day's score for -ST was omitted from the analysis because it was clearly an extreme point. A log transformation of the data was conducted to help equalize the interphase variances.

Negative overt responses also decreased significantly (p < .005); the major change in means occurred between the baseline and TS phases. Positive overt responses did not change significantly between the treatment phases, although there were small increases during the PI phase. All tests for linear regression and lag 1 autocorrelation were insignificant.

The teacher's positive daily reports showed a significant increase (p < .01) during the PI phase and an overall, though not significant,
increase across all phases. Negative daily reports showed a near significant decrease during TS; in general this decrease was found across all phases. On the adjective checklist the Self-Acceptance score increased (pre-post) from the 83rd to the 88th percentile while the Self-Criticism score decreased from the 33rd to the 11th percentile.

T1 showed improvement in self-esteem with significant changes in positive and negative self-thoughts. The self-thought changes were supported by two other self-report measures, as well as by overt negative responses. However, little effect on positive overt responses was shown.

Teacher Two

Both positive self-thoughts and negative self-thoughts increased significantly during the baseline phase. Inspection of the frequency graph (Figure 2) indicates that in both cases these positive trends were due to extremely low frequencies reported during the initial 4 to 6 days of observation. After this period, the frequencies stabilized at somewhat higher values. It seemed reasonable to assume that this observed shift was due to adaptation, i.e., the teacher's getting used to the self-observation procedures. Therefore these first few days were omitted from the analysis (the first 6 days for +ST and the first 4 days for -ST). This changed the baseline slopes and provided a more stable indication of the teacher's actions before self-observation was started. Since there appeared to be no change in levels between the Baseline I and II phases, the data for these were combined to produce a single baseline phase (for +ST and -ST only).

The analysis of variance demonstrated significant changes in both positive and negative self-thoughts for T2 (see Table 2), with the major shift in the means occurring between the baseline and PI phases.

For overt responses (both positive and negative) the Baseline II levels were considerably lower than Baseline I levels, thus presenting a problem of interpretation. Analysis of variance across all four phases indicated a significant decrease for both variables (see Table 2). Simultaneous 95% confidence intervals for all the contrasts (Dixon &
TABLE 2

Analysis of Variance for Positive and Negative Self-Thoughts (Baseline I and II combined) and Positive and Negative Overt Responses (Baseline I and II separate) Between and Within Intervention Phases: Teacher Two

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive self-thoughts</td>
<td>Between phases</td>
<td>2</td>
<td>13.88</td>
<td>5.12</td>
<td>&lt;.025</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>22</td>
<td>2.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative self-thoughts</td>
<td>Between phases</td>
<td>2</td>
<td>25.70</td>
<td>6.02</td>
<td>&lt;.01</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>24</td>
<td>4.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive overt responses</td>
<td>Between phases</td>
<td>3</td>
<td>.0057</td>
<td>8.14</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>29</td>
<td>.0007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative overt responses</td>
<td>Between phases</td>
<td>3</td>
<td>.0284</td>
<td>10.50</td>
<td>&lt;.0005</td>
</tr>
<tr>
<td></td>
<td>Within phases</td>
<td>29</td>
<td>.0027</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Massey, 1969) showed that the significant differences took place between Baseline I and all three following phases for overt negative responses, and between Baseline I and Baseline II, and Baseline I and Follow-up for overt positive responses. When Baseline I was omitted from the analysis of variance, overt negative responses showed an insignificant decrease ($F_{2, 17} = 1.58$), while overt positive responses showed a significant increase ($F_{2, 17} = 5.50$, $p < .025$). Further tests for regression showed a significant negative slope for overt negative responses during the PI phase, i.e., the teacher was increasingly less negative with students when she was using stimulus cues to encourage positive self-thoughts. All other results as well as tests for autocorrelation were insignificant.

Positive daily reports significantly increased during PI ($p < .01$) as well as overall ($p < .01$). Negative daily reports showed insignificant decreases during PI and overall. T2's Self-Acceptance score re-
mained at the 99th percentile, obviously due to a ceiling effect; Self-
Criticism remained unchanged at the 16th percentile.

In general T2 increased positive self-thoughts and decreased nega-
tive self-thoughts—changes which were supported by increases in
positive daily reports. With regard to overt behavior, however, the
results are unclear. Unquestionably, T2's absence for two weeks be-
tween the Baseline I and II periods had a strong effect on her class-
room behavior. Beyond that it appears that the positive intervention
did result in an increase in overt positive behavior, although the changes
were only temporary since levels returned to Baseline II levels by
follow-up. Overt negative behaviors were decreasing (as evidenced by
the significant negative slope during PI), and remained at lower levels
during follow-up.

Posttreatment interviews indicated that both teachers were pro-
foundly influenced by the interventions, referring to their partici-
pation as among the most significant events of their lives. An impor-
tant component of the strategies seemed to be the experiencing of
self-control—of demonstrating to the teachers that they themselves
could change the ways they felt and thought.

Discussion

The following conclusions seem warranted:

1. Both interventions (thought stopping and the positive inter-
vention) were associated with changes in the frequency of self-thoughts.

2. Changes in the self-thought frequencies were supported by
similar changes in two other self-report measures.

3. Corresponding decreases in overt negative behavior and increases
in overt positive behavior were found in each case, but they were much
less than the changes in covert behavior. However, anecdotal posttreat-
ment interview data suggested that other overt behavior changes had
occurred which were not measured by the observation procedures used in
this study.
4. The techniques of self-observation used in this study were feasible for teachers to use in the classroom during ongoing activities. Questions can be raised regarding the utility of concentrating on self-thoughts as a means of changing self-esteem rather than teaching a person better coping behaviors. Low self-esteem (high self-criticalness and low self-positiveness) can be viewed in two ways, and the intervention chosen may vary accordingly. In one case, the self-criticalness may be the result of actual problems and hence a realistic response to one's situation. Here an intervention focused upon learning more facilitative behaviors would take precedence over dealing with the self-critical thoughts. Self-criticalness would be expected to decrease as coping behaviors increased.

But self-criticalness may occur even though a person appears to be coping well with the environment if the person's evaluative standards are far too strict or harsh (Bandura, 1971). In this latter case, intervention on overt behaviors is unnecessary. The intervention should focus rather on the self-critical thoughts themselves, as in this study, seeking to change the patterns of self-evaluation. The two teachers in this study were chosen not because of classroom problems, but because they indicated they were much too critical of themselves and wanted to change. Hence the interventions on self-evaluations. Interestingly, the self-acceptance and self-criticism scores of these teachers, as measured by the Gough Adjective Check List, were not consistent with their views of themselves. Indeed one teacher was at the 99th percentile in self-acceptance.

The thought stopping resulted in negative and positive thought reductions for T1 (see Fig. 1), which was surprising since only negative self-thoughts were the focus of the intervention. In the follow-up interviews, the teacher provided a possible explanation for this finding. She reported thinking negative self-thoughts, using the "Stop" procedure, and then switching to a positive self-thought. Thus she may have inadvertently made her negative self-thoughts the dominant cue for positive self-thoughts, thereby reducing positive self-thoughts as negative self-thoughts decreased. Additional training sessions might
have facilitated the effectiveness of thought stopping (cf. Thoresen & Mahoney, 1974). Clearly, additional studies are needed to evaluate the value of thought stopping as an effective covert control strategy.

The positive intervention may have inadvertently used the high probability behaviors (HPB) as stimulus cues (discriminative stimuli) for positive self-thoughts rather than as reinforcing stimuli. The sequence reversal, i.e., engaging in the high probability response before the target behavior, however, appeared effective. Teachers reported the reinforcing effects of engaging in positive self-thoughts; using the wrist counter to record a +ST may have functioned as a reinforcing event. The temporal sequence of a low probability behavior followed by a high probability response may not be essential in all conditions of human learning (cf. Ashem & Donner, 1968). Use of a high probability response as a discriminative stimuli may be sufficient without providing overt reinforcement since covert self-reinforcement may occur, as was suggested in the interview data with the teachers.

A limitation of the present study involves the problem of experimenter demand factors as an alternative explanation of the results (cf. Jeffrey, in press). Conceivably, changes in +ST may have covaried with the nonspecific effects of engaging in a "treatment" per se and not necessarily with the PI treatment. This rival hypothesis and others should be examined in subsequent studies where (a) some subjects are mislead about the possible effects of treatment, and (b) a reversal or multiple baseline design is employed, or (c) a comparative group design is used (Thoresen, 1972). While comparison of each subject with his own baseline performance does answer the question of relative change for that person, it fails to explain fully why change may have occurred. Additional studies seem warranted in examining how interventions can be used to alter covert actions that relate to positive self-esteem and self-concept.

The reactive effect of self-observation was evident with TI. She reported being shocked at the number of negative self-thoughts she engaged in the first day (58), and her counts decreased immediately thereafter. This reaction suggested that some form of self-punishment
may have been operating to reduce negative self-thoughts. In contrast, the increases in negative and positive self-thoughts for T2 after the first 4 to 6 days were probably not due to reactivity, but to increases in the accuracy of her self-observation. This teacher reported taking the first week to arrive at a stable definition of what thoughts she was to observe. Further studies are needed to separate the reactive and reliability effects of self-observation, primarily dealing with overt behavior (Kazdin, in press). The reliability of covert self-observation is more difficult since an external observer is not possible. Concurrent observations of overt actions related to internal actions, however, can provide an indirect method.

The results of the overt behavioral changes were inconclusive. While PI was associated in both cases with increases in positive overt responses, the magnitude of change was not comparable to the changes in self-thoughts. Further, the changes were not maintained during follow-up. Negative overt behavior was more responsive to change in negative self-thoughts, and decreases were maintained during follow-up. It seems likely that covert behavior toward oneself and overt behavior toward others, while sharing some of the same stimuli and reinforcing conditions, also has separate controlling conditions. The self/other distinction seems crucial. It may be more appropriate to compare changes in covert self-evaluations with changes in overt self-evaluations, or changes in overt evaluations of others with covert evaluations of others. Such comparisons may yield a clearer relationship between overt and covert variables. For example, changes in self-thoughts could be obtained from a disguised task which measured the amount of self-reinforcement in a structured setting (e.g., Marston, 1965).

In general, the present study suggests that covert events are both observable and modifiable, and need not remain outside the realm of effective intervention and self-control. Further work is needed, however, to establish the functional relationships between overt and covert behaviors.
References


Homme, L. E. Perspectives in psychology: XXIV. Control of coverants, the operants of the mind. Psychological Record, 1965, 15, 501-511.


Mahoney, M. J. Toward an experimental analysis of coverant control. Behavior Therapy, 1971, 2, 510-521. (a)

Mahoney, M. J. The self-management of covert behavior: A case study. Behavior Therapy, 1971, 2, 575-578. (b)


