Dentists' Voice Control: Effects on Children's Disruptive and Affective Behavior

Paul E. Greenbaum
University of Florida

Clara Turner
Department of Pediatric Dentistry
University of Florida

Edwin W. Cook III and Barbara G. Melamed
University of Florida

Voice control, a punishment technique based on loud commands, has been used widely in pediatric dentistry. This study examined whether (a) loudness is a necessary component of the technique, (b) voice control actually reduces children's disruptive behavior, and (c) after treatment, children's negative affect increases. Subjects were forty 3½- to 7-year-olds who posed potential behavior problems and who were scheduled for cavity restoration. Children were assigned randomly to either loud- or normal-voice groups. Children who were assigned to either group but who were not disruptive formed a nonexperimental control group. Prior to and after treatment, children reported their feelings using the Self-Assessment Mannequin. Disruptive behavior was scored using the Behavior Profile Rating Scale. Results indicated that, following loud, but not normal voice commands, children reduced their disruptive behavior ($p < .004$) and self-reported lower arousal ($p < .09$) and greater pleasure ($p < .10$). Theoretical and practical implications of these findings are discussed.

Key words: nonverbal communication, voice, dentistry, punishment, behavior management

Pediatric dentists frequently use aversive stimuli, in the form of loud commands, when children's behavior has become disruptive during dental treatment. Known as voice control, this technique is applied contingently, being used when children's behavior has caused ongoing treatment to stop. Typically, such patient behaviors include crying, kicking, flailing arms, and other overt fear responses that occur while children are seated in the dental chair. Dentists report that most pediatric patients respond to voice control by stopping their disruptive behavior (Brauer, 1964; Chambers, 1976; Pinkham & Paterson, 1985; Wright, 1975).

Although all theorists would agree that the purpose of any punishment technique is to remove an unwanted response from an individual's reper-

Requests for reprints should be sent to Paul E. Greenbaum, Department of Epidemiology and Policy Analysis, Florida Mental Health Institute, University of South Florida, 13361 Bruce B. Downs Boulevard, Tampa, FL 33612-3899.
toire, consensus as to what defines punishment does not exist. Instead, punishment has been defined using two different but not mutually exclusive criteria (Walters & Grusec, 1977; Zillman, 1979). One definition of punishment refers to the presentation of an aversive stimulus following the performance of a behavior. The second definition stresses the functional role of punishment—namely, “a reduction in the probability of a response as a result of immediate delivery of a stimulus for that response” (Azrin & Holz, 1966). Dentists’ voice control meets both criteria for classification as a punishment technique and can be viewed profitably from this perspective.

Use of aversive stimuli in clinical practice as part of therapeutic punishment has long been controversial. Within psychology, the debate has focused on whether aversive stimuli (i.e., shock therapy) should be used with severely disturbed children. However, the broader question of when aversive stimuli should be used, if at all, has also been raised. Recently, Skinner (1988), a leading critic of punishment in clinical practice, suggested that therapeutic punishment may be appropriate under some circumstances: “If brief and harmless aversive stimuli, made precisely contingent on self-destructive behavior, suppress the behavior and leave the [autistic] children free to develop in other ways... I believe it can be justified” (p. 22). Not surprisingly, however, due to potential abuse and negative side effects, punishment has been used infrequently in applied behavioral settings.

Voice control, as used by pediatric dentists, represents an exception. This practice is reportedly both long-standing and widespread (Brauer, 1964). In spite of the popularity of this technique, however, no prior research has been done to test its efficacy or to assess the psychological impact of voice control on children receiving dental treatment.

Most conclusions about punishment have been based on animal research using electric shock. Generalizing from this data base to other types of punishers used with humans had been problematic because different punishing stimuli used with different species may produce highly specific and quite varied effects (e.g., response facilitation). Among children, a variety of aversive stimuli such as loud noise, hand slapping, and mal-odorous scents have been studied, but with far less frequency than animal studies using electric shock as punishing stimuli. Currently, it is not known whether loud noises and electric shock have similar effects on behavior (Walters & Grusec, 1977).

Several prior studies have examined the effects of loud verbal stimuli on children, but results are equivocal for a variety of reasons. Several studies have confounded loud stimuli with other aversive stimuli. Johnson and Baumeister (1981) pointed out that, in many studies, verbal exclamations or reprimands have been paired initially with either shock or hand slapping, and, consequently, it is unclear what effect loud verbalizations may have
when used alone. Among studies in which loud voice tones were the sole
punishing stimuli, results have varied. Several researchers have found that
loud reprimands were by themselves sufficient to suppress unwanted
behavior (Baumeister & Forchand, 1972; Hall et al., 1971; McAllister,
Stochowiak, Baer, & Conderman, 1969). In contrast, O'Leary, Kaufman,
Kass, and Drabman (1970) found that a teacher's softly spoken commands
were more effective than loudly spoken commands. They suggested that,
among young children, aversive loudness may not enhance response
suppression. Saltz, Campbell, and Skoto (1983) found similar results, in
that 5-year-olds were better able to suppress motor responses when given
normal voice commands than when given loud voice commands.

Within dentistry, however, loud voice commands consistently have been
reported to suppress disruptive behavior among young children. Chambers
(1976) argued that the effective stimulus components of voice control were
the paralinguistic vocal qualities (i.e., sudden, loud, firm) rather than the
verbal message. He suggested that the same message, if shouted in a foreign
language, would be just as effective; how something is said rather than what
is said causes the patient "to stop . . . in whatever he is doing" (Chambers,
1976, p. 797). According to Ridley-Johnson and Melamed (1986), however,
punishment techniques in dentistry, including voice control, can produce
negative emotional arousal in children and should be used rarely. Early
unpleasant dental experiences are often cited as causes of later dental fear
and avoidance.

This study was designed to answer three specific research questions. First,
is loudness a necessary component of the voice control technique, or are
verbal commands issued in a normal tone of voice equally effective?
Second, does voice control, in the immediate, short term, reduce children's
behavior disruptive to dental treatment? Finally, what effect does voice
control have on children's self-reported affect following treatment? Specif-
ically, does fear increase when punishment is used? Answers to these
questions would provide data on a widely used technique to which children
are routinely exposed during their dental health care. More broadly, an
investigation of dentists' use of voice control permits an analysis of
punishment among normal children in a naturalistic setting. Such an
analysis may clarify the debate on when therapeutic punishment is approp-
riate in clinical practice.

**METHOD**

**Subjects**

Subjects were 40 (23 male, 17 female) 3½ to 7-year-olds who were dental
patients at the University of Florida Pediatric Faculty Practice Clinic. All
children were scheduled for a cavity restoration and were considered to pose potential behavior problems. Among children who had never been to the clinic before, children were judged to pose potential behavior problems if, during an initial exam by one of the participating dentists, the child's behavior when seated in the dental chair was considered fearful (e.g., crying, fidgeting, verbal complaints, reticence). This screening occurred at least 1 week prior to the experimental session. Among children who had been treated previously by participating dentists, an assessment was made by the dentist, based on the child's prior disruptiveness during treatment, that the child likely would be disruptive during the experimental session. Three pediatric dentists (one female, two male), all experienced in voice control, participated in the research. All three dentists used voice control routinely in their daily practices.

Procedure

Prior to arrival for treatment, children were assigned randomly by the toss of a coin to one of two conditions—loud- or normal-voice treatment groups. In the loud-voice condition, the dentist issued a sudden (i.e., contingent), loud (i.e., greater than normal voice volume), firm (i.e., duration of at least 2 sec) command (e.g., "Stop it, put your hands down to your side"). This was done when the child's behavior disrupted treatment. Treatment interruption was operationalized as when the dentist either could not continue treatment (e.g., could not place an instrument in the patient's mouth as needed) or when the patient started sustained crying (i.e., not a momentary outburst). This procedure corresponded to the dentist's usual application of voice control. In the normal-voice condition, the dentist delivered similar verbal commands for patient compliance when the child interrupted treatment, except in this case the dentist refrained from using a loud voice, speaking instead in a normal tone. During both conditions, dentists were instructed to deliver their commands using whatever words they normally would use in daily practice. Standardized scripts were not used because such a procedure was believed to compromise external validity by changing the in vivo application of the voice control technique. Children who did not interrupt treatment and who received no voice commands were assigned to a third, nonexperimental control group. This group provided a

---

1Voice control as used in this study was operationalized to include the three paralinguistic qualities described by Chambers (1976); sudden, loud, and firm. Between loud- and normal-voice groups, however, only voice loudness differentiated the groups, because experimental procedures had dentists say similar things at similar times (i.e., sudden or contingent upon treatment stopping) to both groups. For purposes of brevity, the terms loud voice and sudden, loud, and firm voice are used interchangeably in this article to refer to voice control.
comparison of how voice control effects behavior and self-reported affect in relation to children who do not engage in disruptive behavior.

After arrival at the clinic, each child was administered the 15-item Dental subscale (DS) of the Children's Fear Survey Schedule (Scherer & Nakamura, 1968) and was asked to self-report his or her affect about "being at the dentist" using the Self-Assessment Mannequin (SAM; Lang, 1980). The SAM consists of three sets of drawings of a schematized human figure. Each set of drawings represents a 5-point bipolar scale and measures one of three independent dimensions of emotion: pleasure, arousal, and dominance. This version of the SAM has been designed to measure affect in children, with fear being characterized by high arousal, low pleasure, and low dominance (Silverman, 1975). Both instruments, the DS and the SAM, provide measures of children's dental fear; the DS, however, represents a relatively stable, more trait-like index, whereas the SAM measures more situational fluctuations in affect. In earlier studies, Klingman, Melamed, Cuthbert, and Hermecz (1984) found the test-retest reliability of the DS to be \( r = .86 \); Silverman (1975) found that children's SAM ratings varied across situations.

After completing these questionnaires, the child was escorted into the operatory, where the dentist and assistant began cavity restoration treatment. This procedure involved injection, rubber dam placement, use of a high-speed drill, and completion of an amalgam restoration. During this operatory period, a wall-mounted videocamera recorded the child's behavior. Videotapes were later scored using the Behavior Profile Rating Scale (BPRS; Melamed, Hawes, Helby, & Glick, 1975). The BPRS was designed to provide a measure of the child's overt fear and disruptive behavior during dental treatment. The durations of 22 different behaviors that lead to treatment problems (i.e., crying, kicking, refusal to open mouth, flailing arms, etc.) are scored.\(^2\) Each behavior has been weighted by a factor from 1 to 5 based on dentists' judgments of each behavior's degree of disruption (see Melamed & Lumley, 1988, for a complete description, including psychometric properties of this scale).

Only one interruption per subject was studied. For the normal-voice group, this was the first occurrence of an interruption that necessitated the dentist's use of verbal commands, which were delivered in a normal voice. To permit an adequate measure of predisruption behavior, a further requirement was that the interruption had to occur either at least 60 sec after the onset of treatment or if a BPRS behavior occurred within the first 10 sec of a 60-sec preperiod. For the loud-voice group, in addition to the required delay, the disruptive incident also had to elicit a loud vocal.

---

\(^2\)In the current study, one BPRS behavior, "eyes closed," was not scorable due to the angle of view of the stationary videocamera.
command from the dentist. Usually this was also the first disruption; in a total of 5 cases of 13, however, the selected interruption was not the first.

To assess reliabilities for BPRS judgments, two independent raters, a rater blind to the hypotheses of the study and one of the authors (Greenbaum), scored each type of behavior’s time of onset and offset for disruptive episodes from eight children. BPRS scores used in this study are based on both 1- and 10-sec time intervals; therefore, agreement scores were estimated by Pearson’s $r$ for both 1- and 10-sec intervals. Interrater agreement was adequate for measures based on either time interval; for 10 sec, $r = .94$; for 1 sec, $r = .84$.

At the end of treatment, dentists rated the child’s fear and cooperation on 7-point scales ranging from *not at all* (1) to *very much* (7). After leaving the operatory, children were asked again to self-report their feelings, using SAM, about being at the dentist.

**RESULTS**

Background variables of age, sex of child, self-reported dental fear (DS score), and prior treatment experience (*yes–no*) were included for analysis. Based on random subject assignment, it was expected that groups would not differ significantly on these variables. Pretreatment equivalence among treatment groups was tested for each background variable using the appropriate chi-square test or analysis of variance (ANOVA). The groups did not differ significantly on any background variable. Mean age of children was 60.00 months, $F(2, 37) = 0.29$, $p > .74$, with approximately half the children in each group having never received previous dental treatment, $\chi^2(2, N = 40) = 0.53$, $p > .70$. Fifty-eight percent of the children were boys, with similar proportions of boys and girls in each group, $\chi^2(2, N = 40) = 0.35$, $p > .80$. Group means on the DS did not differ, $F(2, 32) = 0.18$, $p > .83$. Moreover, as expected, based on initial screening, DS scores indicated higher-than-average dental fear ($M = 43.37$, $SD = 12.99$) when compared to norms for 5- to 8-year-olds ($M = 30.55$), $t(34) = 5.84$, $p < .01$ (Cuthbert & Melamed, 1982). Thus, prior to the experimental intervention, all groups were similar in age, sex, rated fear of dentistry, and treatment experience.

Children's affect, as measured by the SAM, was also tested for

---

3 In 5 cases, the selected interruption was the second ($n = 2$), third ($n = 1$), or fourth ($n = 2$).

4 The DS was added to the research protocol after data from 5 subjects—1 from the normal-voice condition, 4 from the loud-voice condition—had already been collected; therefore, $n = 35$ for the DS analysis.
pretreatment equivalence. Consistent with nonsignificant differences between groups on the DS, prior to treatment, groups did not differ on any of the three affect dimensions: Pleasure ($M = 2.58$), $F(2, 37) = 0.04$, $p > .96$, Arousal ($M = 3.25$), $F(2, 37) = 1.33$, $p > .27$, and Dominance ($M = 2.12$), $F(2, 37) = 1.08$, $p > .35$.

Counterbalancing of dentists across conditions also was achieved through randomization of subjects' group assignment. All dentists saw at least one subject in each of the three conditions (i.e., Dentist 1—7 loud voice, 4 normal voice, 8 no disrupt; Dentist 2—2 loud voice, 3 normal voice, 3 no disrupt; Dentist 3—4 loud voice, 1 normal voice, 8 no disrupt).

**Manipulation Checks**

Several tests were conducted to assess group comparability on the use of voice commands in the two experimental conditions. First, dentists' voice tone was compared by having a naive rater listen to the voice commands and sort them into loud- and normal-voice episodes. All cases were sorted into their assigned experimental condition.

Verbal content, specifically the punishing-reinforcing quality of what dentists actually said when delivering voice commands, was judged from written text of the spoken words by two independent raters using a 7-point bipolar scale ranging from extremely reinforcing (−3) to extremely punishing (+3). Interrater reliability was measured by Pearson $r$ and equaled .72. Raters' mean scores of the verbal content was between neutral and somewhat punishing (Rater 1, $M = 0.14$; Rater 2, $M = 0.48$). An ANOVA indicated that the two experimental groups did not differ on rated punishing-reinforcing quality of the verbal commands, $F(1, 19) = 1.46$, $p > .24$. Duration of the verbal commands also was tested for group comparability, and no significant difference between groups was found (loud voice, $M = 8.68$ sec; normal voice, $M = 8.25$ sec), $F(1, 19) = .04$, $p > .84$. Additionally, the experimental groups did not differ either on the total number of interruptions per session (loud voice, $M = 4.77$; normal voice, $M = 6.50$), $F(1, 19) = 0.98$, $p > .33$, or on whether the selected interruption occurred early (i.e., before or during the injection; 57%) or late (i.e., after the injection; 43%) in the dental procedure (Fisher's exact test, ns).

**Overt Fear Behavior**

Figure 1 shows graphically the effects of voice control on children's overt fear behavior. Children's BPRS scores are plotted in 1-sec intervals for 10 sec before and after the dentist's verbalization. A two-factor (Group × Time) mixed-model ANOVA of BPRS scores during the 10 sec prior to the
dentist's intervention indicated a significant linear increase among both loud- and normal-voice groups' scores across the preperiod, from a mean of 1.19 at 10 sec preperiod to a mean of 2.86 sec at 1 sec preperiod—time main effect, $F(9, 171) = 2.61, p < .04$, adjusted by Huynh–Feldt, epsilon = .50,\(^5\) linear trend, $F(1, 19) = 13.33, p < .01$. Preverbalization increases in BPRS scores would be expected if dentists were applying the intervention according to operationalized procedures—that is, voice commands occurred when disruption increased to the point at which treatment stopped. Moreover, as the pattern of increased disruption was not significantly different between loud- and normal-voice groups—group main effect, $F(1, 19) = 0.17, p > .68$, Group $\times$ Time interaction, $F(9, 171) = 0.84, p > .50$—these data provide another manipulation check that dentists were consistent across groups, applying the different interventions during equivalent disruptions in each group's treatment.

In contrast, a similar analysis of postintervention behavior, during the 10 sec following the dentist's commands, found that the two groups had significant differences in postintervention BPRS scores. The loud-voice group was less disruptive, having a lower mean score than the normal-voice group ($M_s = 0.83$ and 2.74, respectively), $F(1, 19) = 10.68, p < .004$. Less

\(^5\)All within-subject terms in reported analyses have been adjusted by Huynh–Feldt's epsilon, as significant covariance heterogeneity was found when covariance matrices were tested by Mauchly's sphericity test ($p < .0001$).
disruption was observed during the first postintervention second (loud voice, $M = 1.92$; normal voice, $M = 2.88$) and reached significance at 2 sec ($Ms = 1.38$ and $3.00$, respectively), $F(1, 19) = 4.63$, $p < .05$. Seconds 3 and 4 were marginally lower for the loud-voice group ($ps \leq .13$ and $.065$, respectively), with the remaining 6 sec being significantly lower for the loud-voice group (all $ps < .04$).

To discern more long-term effects, BPRS scores also were analyzed over a 2-min postperiod. Figure 2 shows these data graphically. In this analysis, BPRS 10-sec interval scores were analyzed across 3 min—1 min prior to intervention and 2 min postintervention. As in the earlier analysis, both groups showed an increase in BPRS scores prior to the dentist’s intervention—preintervention time main effect, $F(5, 95) = 5.98$, $p < .003$, adjusted by Huynh–Feldt, epsilon = .50, with significant linear and quadratic trends over the six 10-sec time intervals, $F(1, 19) = 9.15$, $p < .01$, and $F(1, 19) = 4.42$, $p < .05$. Visual inspection of the preintervention scores indicated that the linear and quadratic terms reflected the change from a relatively constant low level of disruption in the first 50 sec to a sharp increase in the last 10 sec. As in the earlier 1-sec analysis, the 10-sec analysis indicated significantly lower BPRS scores for the loud-voice group during the postperiod, $F(1, 19) = 7.58$, $p < .02$ ($Ms = 0.72$ and 2.85, respectively). Thus, the data indicated that the loud-voice group displayed significantly fewer overt fear behaviors within 2 sec after receiving voice control, and this pattern continued for at least 2 min postintervention.

![Diagram showing BPRS scores over time](image)

**FIGURE 2** Children’s disruptive behavior as measured by mean BPRS scores at 10-sec intervals during 60 sec before and 120 sec after dentist's voice control.
Self-Reported Affect

Posttreatment children’s affect ratings (SAM) were analyzed using each child’s pretreatment SAM score as a covariate in a between-groups analysis of covariance. This technique is identical to a pre-post difference score analysis (Cohen & Cohen, 1983). Planned contrasts of the Pleasure dimension indicated that, at posttreatment, loud-voice subjects tended to report greater pleasure than normal-voice subjects (loud voice, $M = 2.87$; normal voice, $M = 1.74$), $F(1, 36) = 3.03, p < .10$. Analysis of the Arousal dimension also indicated that loud-voice subjects tended to report lower arousal than normal-voice subjects (loud voice, $M = 3.19$; normal voice, $M = 4.38$), $F(1, 36) = 3.23, p < .09$. Additional comparisons revealed that the loud-voice group’s arousal did not differ from the no-disruption group’s ratings ($Ms = 3.19$ and $3.35$, respectively), $F(1, 36) = 0.09, p > .77$, and both the no-disruption and the loud-voice groups tended to report less arousal than the normal-voice group ($Ms = 3.28$ and $4.38$), $F(1, 36) = 3.60, p < .07$. Analysis of the other SAM dimension, Dominance, indicated no significant group differences, $F(2, 44) = 1.61, p > .21$.

Thus, among children receiving voice control, posttreatment arousal levels were not significantly different from those of children who did not disrupt treatment and tended to be lower than those of disruptive children who received only normal voice commands. Posttreatment pleasure ratings also tended to be higher for loud- versus normal-voice subjects.

Dentists’ Ratings

Results from dentists’ ratings of patient fear are consistent, with disruptive children being more fearful than nondisruptive children. A between-groups (loud voice, normal voice, no disruption) ANOVA indicated that the groups differed significantly on rated fear, $F(2, 37) = 14.14, p < .0001$. Planned comparisons between loud- and normal-voice groups further indicated that the two voice groups were not rated significantly different ($Ms = 4.85$ and $5.75$, respectively), $F(1, 37) = 1.69, p > .19$. Additionally, both disruptive groups (loud and normal voice) were rated more fearful than the no-disruption group (disruptive, $M = 5.19$; no disruption, $M = 2.68$), $F(1, 37) = 27.04, p < .0001$.

A similar between-groups ANOVA of patient cooperation ratings showed that the no-disruption group was judged more cooperative than the loud-voice group (no disruption, $M = 6.16$; loud voice, $M = 4.23$; normal voice, $M = 5.25$), $F(2, 37) = 6.79, p < .004$; however, loud- and normal-voice groups did not differ in rated cooperation, $F(1, 37) = 2.43, p > .13$. 
CONCLUSIONS

Voice control proved to be highly effective in reducing children's disruptive behavior during dental treatment. Response suppression occurred very quickly (i.e., within 1 to 2 sec) after the dentist delivered a loud command and continued throughout the 2-min postperiod. Importantly, after treatment, increased fear or negative affect, as measured by self-reported pleasure and arousal, was not observed among loud-voice patients. Moreover, reported arousal and pleasure levels among these children did not differ from those of patients who were not disruptive. In contrast, disruptive patients who received normal voice commands persisted in their disruptive behavior and, after treatment, tended to self-report elevated arousal and lower pleasure when compared to loud-voice subjects. Voice control, when used with young children (3½- to 7-year-olds) who were initially fearful of dentistry and who subsequently disrupted treatment, quickly reduced unwanted behavior without increasing negative affect.

The specific research questions posed earlier can now be assessed. First, did the loudness of the voice commands actually reduce disruptive behavior? The data are unequivocal; a rapid and sustained decline in overt fear behavior was associated with loud voice interventions. When verbal commands were delivered without loud voice, disruptive behavior was not suppressed. With regard to the question of whether paralinguistic voice quality or verbal content was the effective stimulus component of voice control, the data provide a similarly clear answer. Only when dentists used loud voice did children respond to commands to cease disruption. Normal voice commands were followed by unabated, or higher levels of, disruption.

Finally, the study assessed the impact of voice control on children's affect. After treatment, children who received loud rather than normal voice commands tended to self-report less arousal and more pleasure. These findings contradict widespread expectations that undesirable emotional side effects accompany punishment. However, findings from numerous studies of brief and harmless punishment have similarly failed to confirm that negative emotional responses occur as a sequel (Johnston, 1972; Matson & DiLorenzo, 1984; Walters & Grusec, 1977; Zillman, 1979).

In conclusion, it appears that a therapeutic punishment procedure, voice control, applied contingently upon children's disruptive behavior, was an effective technique for suppressing children's disruptive behavior during dental treatment. Moreover, undesirable emotional effects, as measured by children's self-reports, were not found. Aside from practical implications for using voice control in dental treatment, these results are consistent with a broad body of literature suggesting that punishment, in the form of brief and harmless aversive stimuli, may be an appropriate way to suppress unwanted behavior without causing affective distress. It should be noted
that more long-term effects—such as increased disruption, dental fear, and avoidance—have not been addressed by this study. Future research needs to answer these long-term questions before recommending routine use of voice control. Punishment, even when used effectively, must still be monitored for dysfunctional effects.

ACKNOWLEDGMENTS

This study was supported by National Institute of Dental Research Grants 5ROI1DE05305 and 5T32DE07133.

We thank Frank Courts, Greg Jerrell, and Bruce Riggs for their assistance in data collection.

Edwin W. Cook III is now with the Department of Psychology, University of Alabama at Birmingham.

REFERENCES


