ON TOKEN REINFORCEMENT AND STUTTERING
THERAPY: ANOTHER VIEW ON FINDINGS
REPORTED BY HOWIE AND WOODS (1982)

ROGER J. INGHAM

CUMBERLAND COLLEGE OF HEALTH SCIENCES
SYDNEY, AUSTRALIA

Howie and Woods (1982) have provided data that, they claim, indicate that a token reinforcement system is redundant in instating and shaping fluent speech within a stuttering treatment program developed by Ingham and Andrews (1973a, b). However, there were substantial procedural differences between the treatment programs referred to in both studies, as well as methodological weaknesses in Howie and Woods' study. These factors provide ample sources of explanation for Howie and Woods' failure to demonstrate benefits from their token reinforcement system.

DESCRIPTORS: stuttering treatment, token reinforcement, adults and adolescents, prolonged speech

This paper aims to raise some concerns regarding a study recently reported in the *Journal of Applied Behavior Analysis* by Howie and Woods (1982) on treating stuttering with the assistance of a token system. In order to place these concerns in context it is necessary to review some background information on the stuttering treatment program that Howie and Woods used and investigations that have established the efficacy of a token system in that context.

In 1965, Israel Goldiamond described a stuttering treatment procedure that founded one of the most commonly used therapies for this disorder. Initially subjects orally read under 250 msec delayed auditory feedback (DAF) and produced slow and unusual speech known as "prolonged speech." This "new fluent speech pattern," which mainly involves extended vowel production, reliably reduced stuttering. The pattern was then shaped toward ostensibly normal speech by systematically reducing the DAF level and increasing the subject's oral reading rate. Thereafter, various generalization strategies were recommended but, unhappily, they were never demonstrated to be reliable. This gave rise to a number of treatment programs designed to achieve this objective. One, developed by this

writer (Ingham, 1972; Ingham & Andrews, 1973a), used a variation of Goldiamond's procedure within a hospital-based token economy. The tokens purchased all food, drinks, and luxury items throughout the subjects' hospitalization.

The treatment program and experiments on the token system were described in the *Journal of Applied Behavior Analysis* by Ingham and Andrews (1973a). Subjects earned tokens for achieving target speech behaviors within 45-min "rating sessions" held regularly during each treatment day. In Stage A of their treatment, subjects (usually a group of four) earned tokens contingent on reduced percentages of syllables stuttered (%SS) for a prescribed number of spontaneously spoken syllables. In Stage B, an instatement phase, they negotiated a variation on Goldiamond's DAF procedure in order to instate and shape prolonged speech: Each subject had to produce a prescribed number of spontaneously spoken syllables with no stuttering and within a target syllable per minute (SPM) range in four consecutive rating sessions (two with and two without DAF). This was required first during 250 msec DAF speech rates and then during 200, 150, 100, 50, and 0 msec DAF. Nonstuttered speech at targeted speech

rates earned tokens but they were lost for stuttering or incorrect speech rate. Clinicians endeavored to monitor each subject's speech throughout nonsleeping hours in the hospital; tokens were then lost for each stuttering, or (for the entire group) for absence of speech at designated times. This tight control was then systematically withdrawn over Stage C, the transfer phase, when tokens were earned for maintaining the target speech in natural speaking situations. The treatment ended when the transfer phase tasks were completed or after 21 days treatment time.

Ingham and Andrews (1973a) described four experiments on Stages A and B of the treatment program. Three were time-series group experiments showing that variations of the token schedule altered the group's mean %SS scores. The fourth compared the number of rating sessions that two subject groups required to complete the Instatement Phase hierarchy. One group ($n = 12$) experienced the contingent token system schedule in this phase, while the other ($n = 11$) experienced a noncontingent schedule—subjects were awarded tokens merely for producing the prescribed syllables for a rating session, that is, without token penalties for stuttering or incorrect speech rate. This last group required significantly more rating sessions to complete this phase than did the first group. It is important to note that both groups also advanced or retreated on the instatement phase hierarchy according to whether or not they produced the target speech behavior for each step.

One of a number of variations on Ingham and Andrews' program has been referred to as the Prince Henry Program (Prince Henry Hospital, 1976). And it was mainly in the context of the Prince Henry Program that Howie and Woods (1982) conducted experiments that, they claim, demonstrated a token reinforcement system is "largely redundant" in influencing rating sessions required to complete the fluency-instatement schedule derived from Ingham and Andrews' (1973a) report. Because their findings failed to replicate those reported by Ingham and

Andrews, it is of some interest to compare both studies in order to find some possible reasons for this result—a result which brings into question a token system's relevance in this form of therapy.

Howie and Woods (1982) reported three experiments that compared the number of rating sessions that different groups of adult stutterers required to complete their instatement phase. In Experiment 1, 10 stutterers were treated with a token reinforcement system outlined in the Prince Henry Program Manual (1976) and another group of 10 stutterers were treated with the same system except that the tokens had no purchasing power and were only used as "tangible representation of their [subjects'] progress." No significant differences appeared between the groups on their pretreatment %SS and SPM scores or total rating sessions required in the instatement phase. Experiment 2 compared two groups of five stutterers that had received a 1973 offering of the Ingham and Andrews (1973b) therapy program. One group earned tokens contingent on target speech in each step of the instatement phase. The other group was "paid a daily 'wage' of tokens at the beginning of each day regardless of . . . performance on the preceding day" (Howie & Woods, 1982, p. 59). Both groups also had similar pretreatment mean %SS and SPM scores plus similar numbers of rating sessions required in the instatement phase. Experiment 3 also compared two groups ($n = 10$; 10) matched on %SS and SPM pretreatment scores, and treated in accordance with the Prince Henry Program Manual (1976). One group received no tokens throughout the instatement phase and required about the same number of rating sessions in the instatement phase as the regularly treated group.

On first appearances these findings seem to justify a conclusion that the token reinforcement system used by Howie and Woods was quite impotent. But this impotence may say more about their experimental methodology and the token system they used rather than about the general applicability of a token system in such a pro-

gram. Indeed the methodological issues and the type of token system used in Howie and Woods' study may adequately explain the differences between their findings and those reported by Ingham and Andrews (1973a), differences that may also highlight some important considerations in designing token systems for this type of therapy.

METHODOLOGICAL ISSUES

Three methodological issues associated with Howie and Woods' study appear to jeopardize the validity of their experiments.

1. All three experiments are distinguished by a complete absence of reliability data. It is claimed that their clinicians achieved satisfactory measurement reliability in *other* studies, but neither intra- nor interrater reliability data were reported in *this* study.

2. The groups were matched on pretreatment telephone conversation data rather than rating session-type conversation data; that is, data derived from the experimental conditions. By contrast Ingham and Andrews' (1973b) subjects were assessed before entering the instatement phase in "probe" conditions that virtually matched rating session conditions. This is an important difference because stuttering may vary substantially across different speaking conditions. For example, Ryan (1974) reported that stuttering frequencies among a sample of 20 stutterers were much higher (though not consistently) during telephone conversations than during monologues or conversational speech. Thus, although Howie and Woods' subjects were equated on telephone tasks, there is no evidence that this resemblance extended to the experimental conditions.

3. There is a strong probability that Experiment 2 was confounded by uncontrolled retrospective subject selection and yet another experiment within their experiment. Post hoc and unspecified subject selection offers unlimited opportunities for experimental bias. This is especially problematic when Howie and Woods possibly chose from "60 new adult stutterers

(treated) each year" (Prince Henry Hospital, 1976, p. iii) by the same procedure. The fact that the Experiment 2 groups were treated in the same month does not justify ignoring other equally eligible experimental subjects. But more importantly the control group treated in 1973 by Ingham and Andrews' program (that is, by "contingent tokens with back up reinforcement") were also part of an experiment that assessed the effect of token economy modifications. This experiment involved all subjects treated in 1973 by "contingent tokens" and was the topic of an unpublished doctoral dissertation by Cinar (1976). It demonstrated the effect on speech performance of deflating and inflating the overall token economy. Such variations may have substantially influenced Howie and Woods' Experiment 2.¹

TOKEN SYSTEM ISSUES

Howie and Woods suggested that the Prince Henry Program and that reported by Ingham and Andrews (1973a, b) "differed only in minor respects" (Howie & Woods, 1982, p. 59), such as in their speech rate shaping procedures. Such differences would need to be minor because the

¹A lesser methodological issue concerns Howie and Woods' claim that Ingham and Andrews' (1973a) chose an inappropriate independent variable for comparing their contingent and noncontingent token system schedules. Ingham and Andrews established one difference between their contingent and noncontingent token system schedules: the requirement that tokens be earned contingent on achieving the target speech behavior for a rating session. However, both the contingent and the noncontingent token system groups were required to speak a prescribed number of syllables to earn either their "contingent" or "noncontingent" tokens for a rating session. Howie and Woods suggest that the poorer performance of the noncontingent group in Ingham and Andrews' report was because "clients may have risked stuttering and excessive speech rate by rushing to complete their syllables in the limited time" (Howie & Woods, 1982, p. 56). This overlooks the fact that the same time constraints applied for the contingent token system group. In spite of that constraint this group still achieved the target behavior more rapidly than the noncontingent group.

meaningfulness of Howie and Woods' study hinges on essential similarities between the two token systems. For, unlike Ingham and Andrews' study, Howie and Woods provided no data that show their token system actually modified behavior.

Even a cursory reference to the Prince Henry Program Manual (Prince Henry Hospital, 1976) shows that Howie and Woods disingenuously referred to differences between the programs as "minor." For example, unlike Ingham and Andrews' subjects who were hospitalized throughout the instatement and transfer phases, Howie and Woods' subjects were only hospitalized for the first 5 treatment days. This fact alone might be expected to reduce the subjects' necessity to earn and save tokens. Furthermore, Ingham and Andrews' stutterers were monitored throughout the instatement phase from the time they awoke until they went to sleep, whereas Howie and Woods' subjects only attended clinic sessions from 8:00 a.m. to 8:00 p.m. During this time their between-rating-sessions speech was not subject to token control conditions and was only monitored informally. More importantly, other arrangements existed to control their subjects' speech over the instatement phase. For instance, the clinician could set daily speech performance goals for all subjects that had to be reached (by unspecified means) before the treatment day ended (see Prince Henry Hospital, 1976, p. 8). In short, if token control was not sufficient to achieve a daily goal then the clinician could intervene to provide additional assistance!

In general, therefore, Howie and Woods' token system was combined with many supplementary controls over their subjects' speech behavior, which caused the program to differ in important ways from Ingham and Andrews' (1973a, b) program. Furthermore, these supplementary controls might have provided an entirely different source of management over speech behavior. The manual fails to detail the consequences of not meeting the daily treatment "goals" or not completing the instatement phase within the first 5 days. However, because

it is stipulated that these goals *must* be met, it follows that the stutterer might even expect to be removed from treatment if they were not reached. Faced with these exigencies, it may not be surprising that Howie and Woods' stutterers showed less responsiveness to the token system than to the therapy format schedule. At the very least, the impact of such supplementary conditions needs to be isolated in order to evaluate the token system's effects.

The apparent impotence of Howie and Woods' token system is even less surprising in view of some other differences between their system and that used by Ingham and Andrews. For instance on the first treatment day each subject was donated 100 tokens. Because a token's value is not specified it is difficult to estimate the impact of these unearned tokens on first day performance. But if their value resembled Ingham and Andrews' tokens then there is reason to believe that such "savings" may have weakened Howie and Woods' token system. During investigations on the Ingham and Andrews' program some data were collected on subjects whose savings either exceeded 50 tokens or were less than 20 tokens prior to speaking tasks in the transfer phase. These data (Ingham, 1975) show that the contingent token system had virtually no effect on %SS scores when savings exceeded 50 tokens, but was effective when savings were 20 tokens or less. Another oddity in Howie and Woods' token system is that savings may have been removed at the end of each day because the Manual states that "All patients begin each day afresh on an equal standing as far as token payments are concerned. . . ." (Prince Henry Hospital, 1976, p. 10). Thus each patient might expect that tokens acquired toward the end of a treatment day would be valueless on the following day. Finally, the number of tokens that could be earned changed progressively over each day; for instance, in session 1 subjects could earn five tokens but by session 10 they could earn 14 tokens. This earning schedule differs radically from that used by Ingham and Andrews wherein all instatement phase rating sessions had equal

token earning possibilities. At the very least the combination of the token donation, daily loss of savings, and variable token rewards across daily rating sessions means that Howie and Woods' token system was quite unlike Ingham and Andrews' system. Moreover, unlike the latter system, there is no evidence that Howie and Woods' token system ever achieved functional control over stuttering behavior.

A final concern relates to the steps in the instatement phase and rating sessions. In Ingham and Andrews' procedure, subjects could make only one attempt in a rating session to pass an instatement phase step. As a result, the number of rating sessions in this phase reflected the number of attempts a subject required to complete all instatement phase steps. But in Howie and Woods' study this nexus was broken, because subjects could have more than one attempt within a rating session at passing a step on the instatement phase schedule (see Prince Henry Hospital, 1976, p. 10). In Ingham and Andrews' (1973a) study, the consequence for stuttering in a rating session was both token loss and failure to complete an instatement phase step. However, in Howie and Woods' study, stuttering and token loss did not necessarily prevent the subject completing an instatement phase step during the same rating session. At least two implications flow from this difference. The first is a generally beneficial implication; subjects should require fewer rating sessions to complete the instatement phase. This appears to be true as the majority of Howie and Woods' subjects required relatively fewer sessions in this phase. The second is more problematic because it means that the token system was not always permitted to influence directly the number of rating sessions needed to complete the instatement phase. Consequently, using the number of rating sessions in the instatement phase as a dependent variable would not necessarily reflect effects attributable to the contingent token system when compared with a noncontingent token system. One possible (though not necessarily exact) indicator of the token system's effect on the number of attempts

subjects made at instatement phase steps might be total stutterings recorded throughout this phase. From this measure, it is evident in Howie and Woods' Experiment 3 that the no-token group may have required 48% more attempts than the "contingent tokens" group to pass through the instatement phase.

CONCLUSIONS

In general, therefore, it seems far wiser to look beyond Howie and Woods' data when assessing the merits of a token system in stuttering treatment programs similar to that described by Ingham and Andrews (1973a). A combination of methodological problems and an unusual therapy format in Howie and Woods' study provide a pool of possible explanations for the differences between their findings and those obtained by Ingham and Andrews. It is almost certain that Howie and Woods' program contained features that were sufficient to nullify their token system's effects on stuttering. Nevertheless, their report has rightly drawn attention to a number of features in Ingham and Andrews' instatement procedure that may be sufficient (though not necessarily more efficient) to control a stutterer's speech performance during similar therapy programs. Indeed, in current applications of this treatment program, this writer (see Ingham, 1980, 1081) has also reported its use without a token system. There is every indication that the program's stepwise task hierarchy, which is also designed to maintain beyond-clinic gains, is sufficient to manage treatment. Certainly, reported investigations on the program's maintenance schedule (Ingham, 1980) have shown that its performance-contingent requirements may modify both maintenance and generalization. But it is premature to conclude from Howie and Woods' study that a token system may not contribute to this form of stuttering therapy. Howie and Woods suggested that their data reflect reasons why clinicians have shown "cavalier disregard" for the use of token reinforcement in applying Ingham and An-

draws' procedure. This is a debatable observation because most other reports on the procedure within the stuttering treatment literature have included use of a monetary or token reinforcement schedule (see, for example, Boberg, 1976, 1980; Carroll, Campbell, Limburn, & Pratt, 1979). There is more justification for concluding that it is Howie and Woods who have shown "cavalier disregard"—in their case for procedural matters that are necessary to describe and investigate a stuttering treatment program.

REFERENCES

- Boberg, E. Intensive group therapy program for stutterers. *Human Communication*, 1976, 1, 29-42.
- Boberg, E. Intensive adult therapy program. *Seminars in Speech, Language and Hearing*, 1980, 1, 365-373.
- Carroll, P., Campbell, H., Limburn, A., & Pratt, S. An approach to group therapy for stammerers. *The Bulletin of the College of Speech Therapists*, February 1979, pp. 2-3.
- Cinar, E. M. K. Tests of the strong axiom of revealed preference theory and variations in consumer equilibrium demand in a token economy. Doctoral Dissertation, Texas A&M University, 1976.
- Goldiamond, I. Stuttering and fluency as manipulatable operant response classes. In L. Krasner & L. P. Ullmann (Eds.), *Research in behavior modification*. New York: Holt, Rinehart and Winston, 1965.
- Howie, P. M., & Woods, C. L. Token reinforcement during the instatement and shaping of fluency in the treatment of stuttering. *Journal of Applied Behavior Analysis*, 1982, 15, 55-64.
- Ingham, R. J. The development, application and analysis of a token system for the treatment of adult stutterers. Doctoral Dissertation, University of New South Wales, 1972.
- Ingham, R. J. Operant methodology in stuttering therapy. In J. Eisensohn (Ed.), *Stuttering: A second symposium*. New York: Harper and Row, 1975.
- Ingham, R. J. Modification of maintenance and generalization during stuttering treatment. *Journal of Speech and Hearing Research*, 1980, 23, 732-745.
- Ingham, R. J. Evaluation and maintenance in stuttering treatment: a search for ecstasy with nothing but agony. In E. Boberg (Ed.), *Maintenance of fluency*. New York: Elsevier, 1981.
- Ingham, R. J., & Andrews, G. An analysis of a token economy in stuttering therapy. *Journal of Applied Behavior Analysis*, 1973, 6, 219-229. (a)
- Ingham, R. J., & Andrews, G. Details of a token economy stuttering therapy programme for adults. *Australian Journal of Human Communication Disorders*, 1973, 1, 13-20. (b)
- Prince Henry Hospital. *Therapist manual for the stuttering treatment program*. Division of Communication Disorders, 1976.
- Ryan, B. P. *Programmed therapy for stuttering in children and adults*. Springfield, Ill.: Charles C Thomas, 1974.

Received August 10, 1982

Final acceptance June 13, 1983