

*THE EFFECT OF TOKEN REWARDS ON  
"INTRINSIC" MOTIVATION FOR DOING MATH*

J. CHRISTOPHER MCGINNIS

FATHER FLANAGAN'S BOYS' HOME

PATRICK C. FRIMAN

FATHER FLANAGAN'S BOYS' HOME AND  
CREIGHTON UNIVERSITY SCHOOL OF MEDICINE

AND

WILLIAM D. CARLYON

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

This study used a multielement baseline design to analyze the effects of token rewards delivered contingent upon completion of math problems by 2 middle-school boys. Time spent on math and number of work pages completed increased (with high accuracy) during reward conditions and were maintained during fading and withdrawal. At follow-up, time spent and work pages completed remained well above baseline for 1 boy and fell below for the other, while accuracy remained high and ratings of liking math were the highest possible for both boys. Overall, the results are inconsistent with warnings about use of token rewards to motivate children.

DESCRIPTORS: reinforcement, motivation, token economy, mathematics

Motivating children to do math can be difficult, and contingent token rewards are sometimes used to help. The large literature asserting a decremental effect of rewards on "intrinsic" motivation is especially critical of token rewards used with children (e.g., Kohn, 1993). To date, attempts to address these criticisms have mostly involved critical reviews of the entire literature (e.g., Eisen-

berger & Cameron, 1996). Very few studies employing behavior-analytic methods with children have been reported. We found none published within the past 20 years, and only one that used math as a dependent variable (Vasta & Stirpe, 1979). It alternated two conditions, contingent rewards and baseline, and a decremental effect was not detected, although a transient below-baseline performance did occur for 2 of the 4 middle-school participants. The purpose of this study is to extend the work of Vasta and Stirpe by including a graduated withdrawal and subjective ratings analysis of token rewards used to motivate children to do math.

## METHOD

### *Participants*

The 2 participants included Charlie, an 11-year-old boy of average intelligence placed in residential care following numer-

---

This study was conducted in partial fulfillment of the first author's PhD requirements at The University of Southern Mississippi. J. Christopher McGinnis is now with the Family Psychology Clinic of Fort Myers, Florida. William D. Carlyon is now with Hillsborough County Public Schools, Tampa, Florida. Special thanks are extended to Jayne Arneil, Frank Bowers, Melissa Deutsch, and Michael Handwerk for their technical assistance, and to Dan Tingstrom, Ron Edwards, Joe Olmi, and John Koeppel for their assistance during the development of this study.

Correspondence and requests for reprints should be directed to Patrick C. Friman, Youthcare Building, Boys Town, Nebraska 68010 (E-mail: frimanp@boystown.org).

ous foster placements, and Michael, a 9-year-old boy of high average intelligence who was the son of administrative employees at the residential program.

### *Experimental Tasks*

Throughout all conditions, participants had a choice between three types of experimental work pages: (a) math, (b) letter coding, and (c) word unscrambling. Each math work page had five problems (addition, subtraction, and multiplication) chosen jointly by the first author and each child's classroom teacher. Answers were considered either correct or incorrect. Letter-code work pages consisted of two messages written in number form (10 to 20 characters) with a decoder key (i.e., 1 = A, 2 = B, 3 = C, etc.) at the top, and word-unscramble work pages consisted of three five-letter words with letters scrambled and a clue at the top.

### *Design*

This study used a multielement multiple baseline design across participants combined with an ABCDA sequence of conditions for Charlie and ABABCD A for Michael. Dependent measures were time spent working, pages completed, and accuracy (percentage correct). Ratings of *liking* math (range, 1 = *not at all* to 7 = *very much*) were obtained after each phase.

### *Procedure*

Participants attended three to five 15-min sessions per week and were free to choose any type of work page in every session, but only math pages were exposed to the reward contingency. After completion, participants gave the pages to the experimenter, who recorded session totals of pages completed, work time elapsed, and accuracy.

### *Conditions*

*Baseline.* Prior to the first baseline session, the experimenter explained the three tasks to

the participants and let them practice on sample pages from each type. At the beginning of all baseline sessions, the experimenter asked, "What type of page would you like to work on?" After each page was completed he asked, "What is your next choice?" The experimenter remained neutral at all times in order to avoid affecting choices.

*Rewards, withdrawal, follow-up.* During reward conditions, the boys were still offered a choice of any worksheet but were told, "Beginning today, you will earn a star for every three, four, or five math sheets that you complete." Stars could be exchanged for desired items according to a preset price (e.g., one star for a sticker, nine stars for a yoyo). Following the fixed-ratio 5, rewards were withdrawn entirely; 2 weeks after the final withdrawal session, two follow-up sessions were conducted.

### *Interscorer Agreement*

Interscorer agreement for the behavioral measures was assessed for all sessions by a person who was not an investigator. Agreement scores by participant were calculated with a point-by-point agreement ratio, with averages of 100% for all dependent variables for both boys except time spent working for Charlie, which was 97%.

Additional details about the methods used are available from the first and second authors (see also Vasta & Stirpe, 1979).

## RESULTS AND DISCUSSION

The moderate amounts of time Charlie and Michael spent on math during baseline increased to 100% with contingent rewards and remained high during fading and withdrawal. Performances diverged in follow-up, however, with Michael at 100% and Charlie below baseline (see Figures 1 and 2). Data on pages completed reflect a similar pattern, whereas accuracy was high in baseline for both (with some fluctuation for Michael)

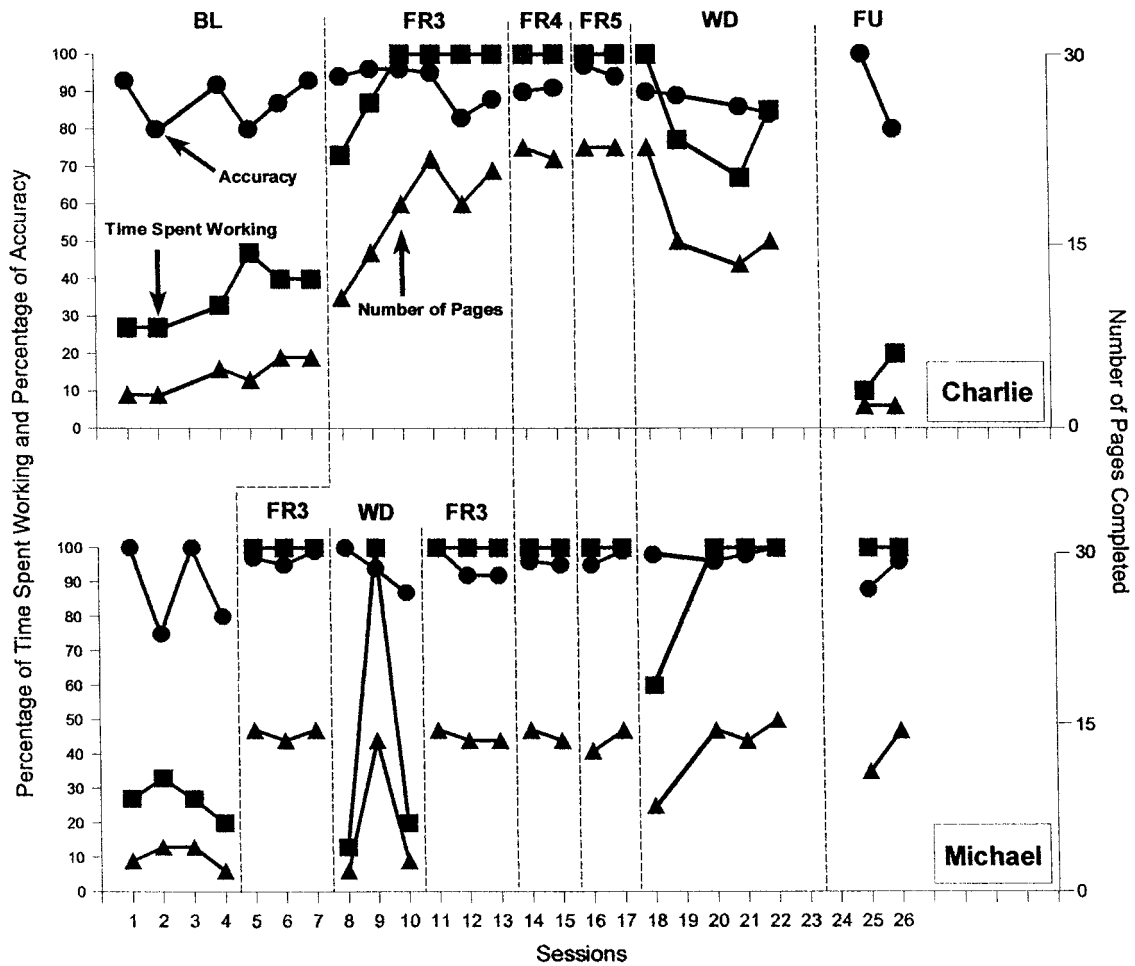


Figure 1. Mean percentages of time spent working (squares), number of pages completed (triangles), and accuracy (circles) attained on the math task. Note that the data indicating number of pages completed are scaled along the far right ordinate.

and remained relatively constant throughout all conditions (see Figure 1). Note that coding and word-scramble data include only time spent (see Figure 2). All Charlie's *liking* ratings were 7, and Michael's increased to 7 in follow-up.

A core assertion in the literature on decremental effects of rewards is that rewarding behavior with unrewarded performance levels above zero results in a rapid decline in performance to below unrewarded levels when the rewards are withdrawn. This literature places special emphasis on the likelihood of declines in the quality of postre-

ward performance and in measures of subjective interest (Kohn, 1993). These claims are frequently employed to warn teachers against use of rewards to motivate children's performance in school. Yet in this study, the reward-based improvements across three dimensions of math performance were maintained during fading and withdrawal for both participants and during follow-up for Michael. In follow-up, Charlie completed fewer pages and spent less time on math than in baseline, but these reductions occurred 2 weeks after reward withdrawal. In addition, as with Michael, Charlie's accuracy

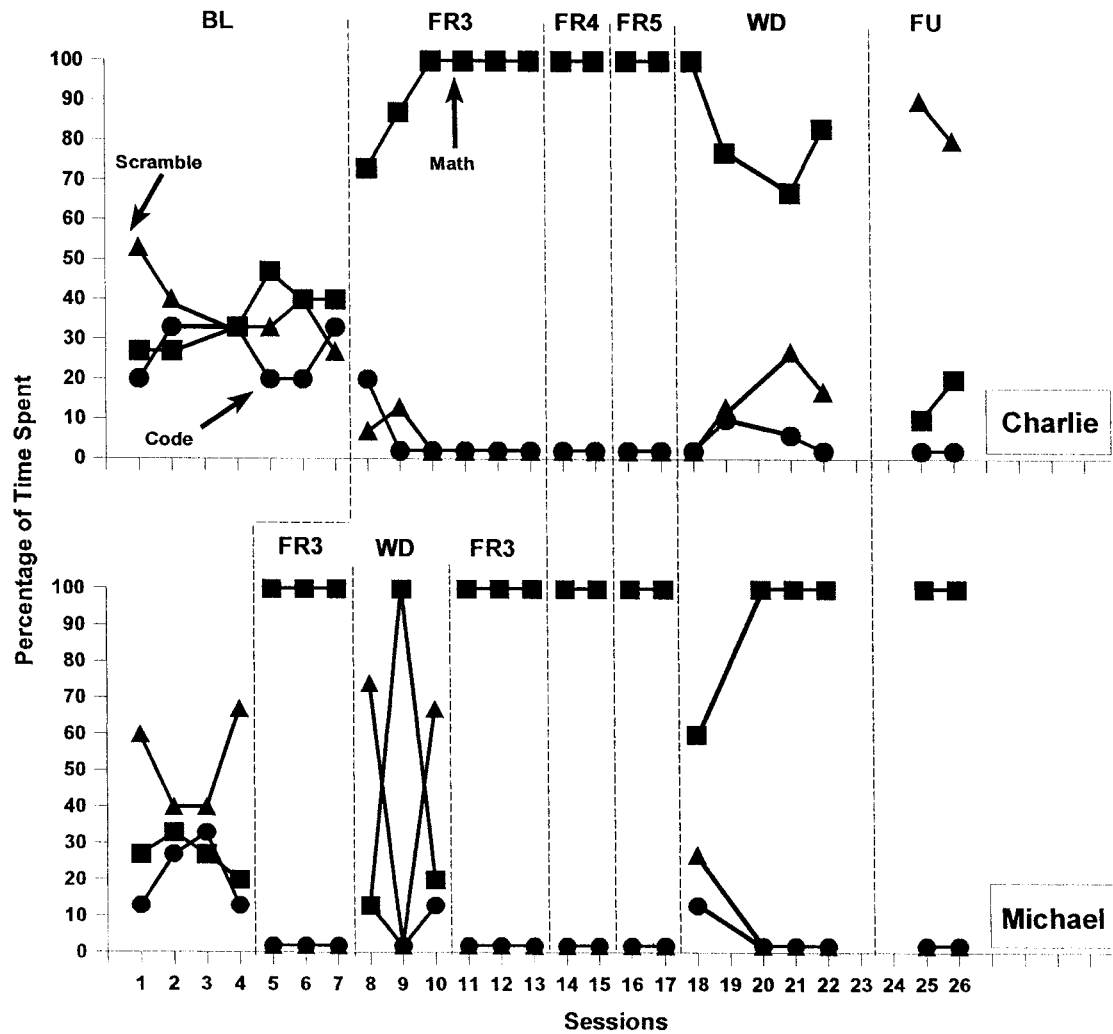


Figure 2. Mean percentages of time participants allocated to each of the three tasks. Squares represent the math task, triangles represent the scramble task, and circles represent the code task.

remained high and his ratings of liking math were the highest possible. By including data on a gradually thinning schedule of reward delivery and obtaining subjective measures, this study extends the findings of Vasta and Stirpe (1979). Collectively, the two studies do not support the strongly worded warnings about the dire effects of tangible rewards on children's interest. Nonetheless, the differing performances in follow-up underscore the importance of individualizing reward-based programs.

The generality of these findings is limited

because the contingencies were applied to problem completion irrespective of accuracy, which is not a conventional classroom arrangement. The literature, however, prescribes such contingencies for stringent tests of the decremental effects of rewards (Greene & Lepper, 1974). We also used an analogue setting, and thus the results may not generalize to the classroom. Lastly, because follow-up data were sloped for both boys, extending the follow-up period until actual trends emerged may have permitted more interpretive clarity. These limitations

notwithstanding, the combined successes achieved in this study and by Vasta and Stirpe (1979) demonstrate that more behavior-analytic research on the decremental effect of rewards in educational settings is warranted.

## REFERENCES

- Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist*, *51*, 1153–1166.
- Greene, D., & Lepper, M. R. (1974). Effects of extrinsic rewards on children's subsequent intrinsic interest. *Child Development*, *45*, 1141–1145.
- Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. New York: Houghton Mifflin.
- Vasta, R., & Stirpe, L. A. (1979). Reinforcement effects on three measures of children's interest in math. *Behavior Modification*, *3*, 223–244.

*Received October 13, 1998*

*Final acceptance April 5, 1999*

*Action Editor, James W. Halle*