Why "Outstanding" is not "Okay" Twenty Years Later -

The Relative Value of Single Word Reinforcers

George S. Lowry, 804-752-7306, glowry@rmc.edu Edward D. Showalter, 804-752-3716, eshowalt@rmc.edu Charles J. Leska, 804-752-3158, cleska@rmc.edu Randolph-Macon College, Ashland, VA 23005

ABSTRACT

Inspiring increases in worker performance continues to challenge to managers. Finding the appropriate incentives to generate intended outcomes is complicated by resource constraints. In particular, rewarding outcomes with money finds limits, both in practical and philosophical ways. Competing claims for the organization's cash limit its use as a reward, and similarly, the perceived marginal value of each additional unit does not generally equate to further increases in performance or worker satisfaction, thus challenging the utility of cash as an incentive.

An alternative is the use of verbal rewards as a substitute for financial ones. The advantages to their use include immediacy in delivery (rather than peeling off dollars each time a reward is warranted) and financial cost savings. When using words as reinforcers, two relevant questions arise surrounding the frequency of their delivery and the power behind them. In operant conditioning, these two dimensions are referred to as schedules and magnitude of reinforcement delivery.

This paper reports a replication study which tests for magnitude differences of single word reinforcers. Using a method developed in 1991 to evaluate the strength of words as reinforcers, 141 words were rated by a participant population similar to that of the 20-year old study. Analysis shows little difference in the mean ratings of the sets of words defining "high" and "low" magnitude reinforcers from 1991 and 2011. There were, however, a sufficient number of substitutions to suggest that the impact of some words as reinforcers has changed over time.

INTRODUCTION

Researchers have long studied ways to increase worker output using an operant conditioning paradigm based upon reinforcement theory. While many of these studies have examined the effects on performance of different schedules of reinforcement, few have explicitly tested the magnitude of reinforcement. Laboratory and field studies altering magnitude used money as the reinforcer (for example, Tranel, Fisher, & Fowles, 1982 or Saari & Latham, 1982). Most organizations face greater financial resource (money) demands than can be satisfied by their limited budgets. To operate within financial constraints, organizations seek and use non-monetary reinforcers, often verbal. Words offer an alternative reinforcing medium with no explicit cost to organizations and can vary in magnitude (quantity and quality). In addition, the feedback message can be tailored to address specifically targeted behaviors.

Reinforcement theory is based on the Law of Effect (Thorndike, 1911), which states rewarded responses tend to be repeated. Rewards act to increase the likelihood of the recurrence of a response, under similar circumstances, so as to produce more rewards. Thus, rewards act to reinforce, or encourage, repetition of responses.

Magnitude of reinforcement defines how much reinforcer to deliver upon satisfaction of a contingency. Magnitude is (1) amount or volume of reinforcer, (2) duration of reinforcer availability, or (3) percentage of concentration of reinforcing substance (Bonem & Crossman, 1988, p. 348).

Relevant research in the industrial and organizational literature has primarily examined the schedule effect, which is the impact of altering a reinforcement schedule while holding the amount (magnitude) of reinforcer fixed, or constant. Some research ignores the magnitude dimension altogether. Other studies attempting to show schedule effects examined the magnitude effect after-the-fact (Latham and Dossett 1978; Yukl, Latham, & Pursell 1976; and Yukl and Latham 1975). Each of these studies concluded schedule effects dominated magnitude effects. Because explicit tests of magnitude effects are lacking, these conclusions are subject to question.

RESEARCH QUESTIONS

The focus of a prior study (Lowry, 1992) was to examine the magnitude effects on performance when other reinforcing conditions (especially schedules of reinforcement) were held constant. In the prior study, a set of approximately 100 words were presented to a sample population for rating on a Likert scale. After rank-ordering the words from highest to lowest (in mean ratings which equated to highest to lowest in reinforcing value), the top and bottom 20 words were then utilized as high magnitude reinforcers (HMR) and low magnitude reinforcers (LMR). Of current interest is a series of experiments to extend that earlier work to investigate schedule and magnitude effects, however, a necessary first step is a replication of the word-discernment method to determine which single-word reinforcers may be relevant today.

To that end, two research questions arise. The first question investigates the ratings given by a current (2011) sample, drawn from a population similar to that used in 1991. Armed with the list of 1991 words and their mean ratings, of interest is any difference in mean ratings derived from a renewed set of HMR and LMR. In other words, does 20 years make a difference in the reinforcing value of dichotomized word sets?

A second question asked what words populate the 2011 HMR and LMR lists that may not have been present on the 1991 lists, and similarly, in what ways do the same words differ in order or rating.

METHOD

The 1991 study measured among other things the effect of high magnitude reinforcement versus low magnitude reinforcement. Each reinforcement was given verbally using a single word from a predetermined list of high and low magnitude reinforcers. This study seeks to determine if the reinforcement level of the original word lists persists twenty years later, and in addition seeks to determine if there is perhaps an improved set of words that current participants may find more relevant.

In the original 1991 study, the high magnitude reinforcers (HMR) and low magnitude reinforcers (LMR) were determined by presenting a set of approximately 100 words to a survey group of undergraduate students. The respondents indicated the level of encouragement provided by each word on a scale from 1-7. The highest rated 20 words were determined to be HMR and the lowest rated 20 LMR. The HMR and LMR groups were compared using a t-test of difference, and the two groups were found to have significantly different means (p <.000.). The scale and instructions as presented are shown in Figure 1 and the HMR and LMR from the original study are presented in Table 1. Table 1

Figure 1: Scale and instructions for determining encouragement levels

Imagine that you have just performed a task and a person who is very important to you is about to say something about your work. This person can use only one word to indicate their feeling about the task you just completed. For each of the following words, please rate their level of encouragement to you. Circle the number associated with the level of encouragement.



Table 1 (Original 1991 Study)

Low Magnitude Reinforcers			High	High Magnitude Reinforcers			
Word	Mean Rating	<u>Std Dev</u>	Word	Mean Rating	<u>Std Dev</u>		
deliberate	2.67	1.217	ideal	5.93	1.215		
cautious	3.07	1.386	awesome	5.96	1.261		
rapid	3.07	1.412	ambitious	6.00	1.122		
quick	3.15	1.380	intelligent	6.07	1.016		
swift	3.26	1.294	fantastic	6.26	1.174		
lively	3.30	1.435	impressive	6.30	0.936		
true	3.30	1.535	incredible	6.30	0.974		
deft	3.37	1.444	marvelous	6.30	0.936		
sprightly	3.41	1.225	amazing	6.30	1.082		
tireless	3.48	1.548	extraordinary	6.33	1.155		
fine	3.52	1.572	expert	6.37	0.949		
careful	3.59	1.368	masterful	6.37	0.777		
brisk	3.63	0.099	fabulous	6.37	1.159		
able	3.63	1.590	outstanding	6.41	0.782		
dutiful	3.81	1.441	perfect	6.44	0.629		
vigilant	3.81	1.564	exemplary	6.56	0.629		
prudent	3.89	1.397	excellent	6.56	0.629		
observant	3.89	1.423	superior	6.59	0.828		
capable	3.89	1.197	exceptional	6.74	0.516		
keen	3.93	1.412	brilliant	6.78	0.497		
Mean of							
Means	3.48			6.35			
t-Test of difference in high and low magnitude group means:							
Note: Rating Scale ranged from $1 - 7$.							

To replicate the results of the first study a new word list was created. The creation of the word list started with the original 40 words from the previous study. Unfortunately the full original instrument used in 1991 was not available for this study so synonyms for each word were generated using the thesaurus function in Microsoft Word. Duplicate words were eliminated leaving 141 unique words to be rated by the students using the scale and instructions from the earlier study. The words were placed in random order and presented in ten sets of approximately 10-15 words each.

Three methods were used to reinforce the reliability of the testing instrument. (1) The words were administered to the students in two media, the majority (nine of the 10 sets of words) using a computer moderated survey and one set of 15 words using a pencil and paper test. (2) One set of words was presented to the students twice at different points and in a different order during the on-line survey to see if the ratings were consistent. (3) The computer moderated survey presented the words within each set in a random order to each participant to avoid any potential ordering bias.

RESULTS

Thirty-eight undergraduate students in a Principles of Management class were given the opportunity to participate in the rating of the words. Of the 38, 35 completed the computer-mediated portion and 33 completed the pencil and paper portion. Each of the 33 paper instruments could be matched to one of the original 35 using student provided identification numbers. The sample included a balanced number of males and female participants.

The computer mediated survey captured beginning and ending times for each participant. The minimum time to rate 141 words was 5 minutes with the longest time being 57 minutes. Average time was 14 minutes with a modal time of 9 minutes. This modal time indicates a rating time on task of approximately 4-6 seconds per word.

In order to check the reliability of the instrument, two t-tests of means were administered: one between the two duplicate sets of words within the computer mediated survey, and one between the words on the computer mediated survey and the paper instrument. In both cases there was no significant difference between the means indicating a high level of test reliability within the test and across instrument types.

As in the original study the words were rank ordered by mean and the top 20 were labeled as HMRs and the lowest 20 as LMRs. These results are presented in Table 2 (Current Study).

The original HMR and LMR words all grouped in the top or bottom half of the ratings respectively indicating that these words have retained their relative level of reinforcement. The mean ratings of the 20 original HMR and 20 original LMR words were also tested against the same words in the current rating. There was again no statistical difference between the ratings over the time period indicating additional reliability of the ratings.

Table 2 (Current Study)

Low Magnitude Reinforcers			<u>High N</u>	High Magnitude Reinforcers				
Word	Mean Rating	<u>StdDev</u>	Word	Mean Rating	<u>StdDev</u>			
hurried	2.17	1.317	astonishing	6.31	0.900			
okay	2.40	1.218	flawless	6.34	1.162			
hasty	2.60	1.418	best	6.37	1.239			
wary	2.60	1.718	<u>excellent</u>	6.37	0.808			
fleet	2.71	1.467	<u>brilliant</u>	6.43	0.698			
guarded	2.86	1.611	fantastic	6.43	0.850			
<u>brisk</u>	2.88	1.338	inspiring	6.43	0.778			
bubbly	2.89	1.345	exceptional	6.46	0.919			
obedient	2.94	1.243	unbelievable	6.46	0.780			
<u>fine</u>	3.00	1.627	perfect	6.49	0.951			
<u>cautious</u>	3.06	1.662	remarkable	6.49	0.702			
alert	3.11	1.711	astounding	6.50	0.762			
compliant	3.11	1.659	breathtaking	6.51	0.742			
intentional	3.11	1.745	magnificent	6.51	0.853			
nimble	3.11	1.859	amazing	6.54	0.741			
<u>rapid</u>	3.11	1.623	marvelous	6.57	0.698			
startling	3.17	1.689	spectacular	6.60	0.736			
conscious	3.29	1.840	incredible	6.63	0.547			
<u>quick</u>	3.29	1.673	extraordinary	6.66	0.725			
unflagging	3.34	1.999	outstanding	6.69	0.631			
Mean of Mea	ns 2.94			6.49				
t-Test of difference in high and low magnitude group means								
	t= 48.89							
	p= 0.0000+							
Note: Rating Scale ranged from 1-7. Bolded & underlined words were from 1991 study.								

CONCLUSIONS

As there was no significant difference in the ratings of HMR's and LMR's between the 1991 and 2011 studies, it would be reasonable to move forward with the original word lists for a full replication of the behavioral study. The words all seem to retain both their absolute level of reinforcement (no difference between the 1991 and 2011 means) and relative level of reinforcement (a significant difference between the mean values of the high and low magnitude reinforcers.)

While using the original list would be appropriate, there is also a compelling case for using the newer list. Again there is no significant difference between the new and old lists and there exists a significant difference in the new list between high and low magnitude reinforcers. While either list should provide similar results, the participant group for the full behavioral study will come from a similar population as the reinforcer rating group as it did in 1991. It makes since then to use the updated list in the full behavioral replication study.

Beyond the simple question relevant for this study however, there may be broader managerial implications in terms of word choice that can be gleaned from this study. Managers may wish to use or avoid certain words based on the audiences perceptions of their level of reinforcement. The sample size and makeup limit the generalizability of the ratings of these words outside an undergraduate setting, but there can be little doubt that words that are often viewed as synonymous convey a substantially different level of reinforcement. Additional research is needed to determine the generalizability of these results.

Another area for future research is looking at individual differences in perceptions of level of encouragement. Preliminary analysis of the data indicates that there may be gender related differences in individual word ratings. Further investigation of this is warranted.

SELECTED REFERENCES

- Agarwal, A. & Tripathi, K. N. (1984). Effects of time, value and contingency of external rewards on intrinsic motivation. <u>Psychologia</u>, <u>27</u>, 254-261.
- Arnold, Hugh J. (1985). Task performance, perceived competence, and attributed causes of performance as determinants of intrinsic motivation. <u>Academy of Management Journal</u>, <u>28</u>(4), 876-888.
- Berger, C. J., Cummings, L. L., & Heneman, H. G. (1975). Expectancy theory and operant conditioning predictions of performance under variable ratio and continuous schedules of reinforcement. Organizational Behavior and Human Performance, 14, 227-243.
- Bellezza, Francis S., Greenwald, Anthony G., & Banaji, Mahzarin R. (1986). Words high and low in pleasantness as rated by male and female college students. Behavior Research Methods, Instruments, & Computers, 18 (3), 299-303.
- Blackman, Derek. (1974). Operant conditioning: An experimental analysis of behavior. London: Methuen & Company, Ltd.
- Blakely, Elbert, Starin, Stephen, & Poling, Alan. (1988). Human performance under sequences of fixedratio schedules: Effects of ratio size and magnitude of reinforcement. The Psychological Record, 38, 111-119.
- Bonem, M., & Crossman, E. K. (1988). Elucidating the effects of reinforcement magnitude. Psychological Bulletin, 104 (3), 348-362.
- Bourne, Lyle E., & Battig, William F. (1966). Complex processes. In Joseph B. Sidowski (ed.), Experimental methods and instrumentation in psychology (pp. 541-576). New York: McGraw-Hill Book Company.

- Buskist, William, Oliveira-Castro, Jorge, & Bennet, Robert. (1988). Some effects of response-correlated increases in reinforcer magnitude on human behavior. Journal of the Experimental Analysis of Behavior, 49, 87-94.
- Campbell, D. T., & Stanley, Julian C. (1963). Experimental and quasi-experimental designs for research. Reprinted from Handbook of Research on Teaching. Boston: Houghton Mifflin Company.
- Cautela, Joseph R. (1972). Reinforcement survey schedule: Evaluation and current applications. Psychological Reports, 30, 683-690.
- Cohen, Bertram D., Kalish, Harry I., Thurston, John R., & Cohen, Edwin. (1954). Experimental manipulation of verbal behavior. Journal of Experimental Psychology, 47 (2), 106-110.
- Crespi, Leo P. (1944). Amount of reinforcement and level of performance. Psychological Review, 51, 341-357.
- Denenberg, Victor H. (1976). Statistics and experimental design for behavioral and biological researchers. New York: John Wiley & Sons.
- Deslauries, B. C., & Everett, P. E. (1977). Effects of intermittent and continuous token reinforcement on bus ridership. Journal of Applied Psychology, 62(4), 369-375.
- Dinsmoor, James A (1966). Operant Conditioning. In Joseph B. Sidowski (ed.), Experimental methods and instrumentation in psychology (pp. 421-450). New York: McGraw-Hill Book Company.
- Edwards, Allen L. (1960). Experimental design in psychological research. New York: Hold, Rinehart and Winston.
- Fantino, Edmund. (1982). The effects of amount and immediacy of reinforcement on self-determined performance standards. Behavior Analysis Letters. 2(6), 327-334.
- Firestone, Philip, and Douglas, Virginia I. (1977). The effects of verbal and material rewards and punishers on the performance of impulsive and reflective children. Child Study Journal, 7 (2), 71-78.
- Fukui, Itaru, & Kimura, Hiroshi. (1988). Effects of reinforcement magnitude and frequency on human choice behavior. Japanese Psychological Research. 30(3), 105-113.
- Gambrill, Eileen D. (1977). Behavior Modification: Handbook of Assessment, Intervention, and Evaluation. San Francisco: Jossey-Bass Publishers.
- Ghiselli, Edwin E. (1964). Theory of psychological measurement. New York: McGraw-Hill Book Company.
- Greenspoon, Joel. (1955). The reinforcing effect of two spoken sounds on the frequency of two responses. American Journal of Psychology, 68, 409-416.
- Guilford, J. P. (1954). Psychometric methods (2nd ed.). New York: McGraw-Hill Book Company.
- Hafner, James L. (1973). A method for quantifying the reinforcement value of verbal items. The Journal of General Psychology, 89, 115-121.

- Hafner, James L. (1973). Quantifying the reinforcement value of verbal items on college students. Journal of the Association for the Study of Perception, 8 (1), 1-7.
- Hafner, James L. & Linkenhoker, Dan. (1974). Quantifying the reinforcement value of verbal items on drug addicts that were inmates of a prison. Journal of the Association for the Study of perception, 9, (1), 17-25.
- Hilgard, Ernest R., & Bower, Gordon H. (1975). Theories of Learning. (4th ed.). Englewood Cliffs, N.J. Prentice-Hall, Inc.
- Jablonsky, S. F., & DeVries, R. (1972). Operant conditioning principles extrapolated to the theory of management. Organizational Behavior and Human Performance, 7, 340-358.
- Jaccard, James. (1983). Statistics for the behavioral sciences. Belmond, CA: Wadsworth Publishing Company.
- Kintz, B. L. & Pappas, James P. (1965). A scaling evaluation of the effectiveness of verbal reinforcers. The Journal of Psychology, 59, 359-365.
- Klatzky, Roberta L. (1980). Human memory: Structures and processes (2nd ed.). San Francisco: W. H. Freeman and Company.
- Latham, G. P., & Dossett, D. L. (1978). Designing incentive plans for unionized employees: A comparison of continuous and variable ratio reinforcement schedules. Personnel Psychology, 31, 47-61.
- Lindman, Harold R. (1974). Analysis of variance in complex experimental designs. San Francisco: W. H. Freeman and Company.
- Monge, Rolf H. & Blank, Harriett D. (1970). Is there an incentive magnitude effect on adult performance? Psychological Reports, 27, 983-992.
- Pittman, Thane S., Davey, Margaret E., Alafat, Kimberly A., Wetherill, Kathryn V. & Kramer, Nancy A. (1980). Informational versus controlling verbal rewards. Personality and Social Psychology Bulletin, 6 (2), 228-233.
- Premack, David. (1971). Catching up with common sense or two sides of a generalization:
 Reinforcement and punishment. In Glaser, Robert (Ed.), The Nature of Reinforcement (pp. 121-150), New York: Academic Press.
- Pritchard, Robert D., Hollenback, John, & DeLeo, Philip. (1980). The effects of continuous and partial schedules of reinforcement on effort, performance, and satisfaction. Organizational Behavior and Human Performance, 25, 336-353.
- Pritchard, R. D., Leonard, D. W., Von Bergen, C. W., & Kirk, R. J. (1976). The effects of varying schedules of reinforcement on human task performance. Organizational Behavior and Human Performance, 16, 205-230.
- Rachlin, Howard. (1976). Introduction to Modern Behaviorism. (2nd ed.). San Francisco: W. H. Freeman and Company.

- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80, 1-28.
- Runquist, Willard N. (1966). Verbal Behavior. In Sidowski, Joseph B. (Ed.), Experimental Methods and Instrumentation in Psychology (pp. 487-540), New York: McGraw-Hill Book Company.
- Saari, Lise M., & Latham, Gary P. (1982). Employee reactions to continuous and variable ratio reinforcement schedules involving a monetary incentive. Journal of Applied Psychology, 67(4), 506-508.
- Sidowski, J. B., & Lockard, R. B. (1966). Some preliminary considerations in research. In Joseph B. Sidowski (ed.), Experimental methods and instrumentation in psychology (pp. 3-32). New York: McGraw-Hill Book Company.
- Shaughnessy, John J. & Zechmeister, Eugene B. (1985). Research Methods in Psychology. New York: Alfred A Knopf.
- Siegel, Sidney. (1956). Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill Book Company.
- Skinner, B. F. (1953). Science and Human Behavior. New York: The Free Press.
- Tranel, Dan T, Fisher, Anne E., & Fowles, Don C. (1982). Magnitude of incentive effects on heart rate. Psychophysiology, 19(5), 514-519.
- Vallerand &, Robert J. (1987). On the motivational effects of positive verbal reinforcement on performance: Toward an inverted-u relationship. <u>Motivation and Emotion</u>, <u>11</u> (4), 367-378.
- Vallerand &, Robert J. (1983). The effect of differential amounts of positive verbal feedback on the intrinsic motivation of male hockey players. Journal of Sport Psychology, 5, 100-107.
- Yukl, G. A., & Latham, G. P. (1975). Consequences of reinforcement schedules and incentive magnitudes for employee performance: Problems encountered in a field setting. Journal of Applied Psychology, 60, 294-298.
- Yukl, G. A., Latham, G. P., & Pursell, E. D. (1976). The effectiveness of performance incentives under continuous and variable ratio schedules of reinforcement. Personnel Psychology, 29, 221-231.
- Yukl, G. A., Wexley, K. N., & Seymore, J. D. (1972). Effectiveness of pay incentives under variable ratio and continuous reinforcement schedules. Journal of Applied Psychology, 56, 19-23.
- Zeiler, Michael. (1977). Schedules of reinforcement: The controlling variables. In W.K. Honig & J.E.R. Staddon (Eds.), Handbook of Operant Behavior (pp. 201-232), Englewood Cliffs, NJ: Prentice-Hall.
- Zinser, Otto, Young, Janice Gayle, & King, Phillip E. (1982). The influence of verbal reward on intrinsic motivation in children. The Journal of General Psychology, 106, 85-91.