Br. J. educ. Psychol., 58, 1-14, 1988

ENHANCING AND UNDERMINING INTRINSIC MOTIVATION: THE EFFECTS OF TASK-INVOLVING AND EGO-INVOLVING EVALUATION ON INTEREST AND PERFORMANCE

BY RUTH BUTLER
(Hebrew University of Jerusalem)

SUMMARY. This study was designed to test the hypothesis that intrinsic motivation will be differentially affected by task-involving and ego-involving evaluation, and that provision of both kinds of evaluation will promote ego-involvement rather than task-involvement. Twelve classes of fifth and sixth grade pupils were randomly assigned to one of three feedback conditions. Pupils received either ego-involving numerical grades or task-involving individual comments or both after performing interesting tasks, one convergent and one divergent, on each of two sessions. Interest and performance for 132 randomly selected pupils of high or low school achievement were measured at pre-test, during the manipulation and at a third session, when no further evaluation was anticipated. As hypothesised, interest and performance on both tasks at both levels of school achievement were highest after comments, both when further comments were anticipated and when they were not. Grades and grades plus comments had similar and generally undermining effects on both interest and performance, although high achievers who received grades maintained high interest and convergent thinking when further grades were anticipated. These results are discussed in terms of the contribution of this distinction between task and ego-involvement to further understanding of intrinsically motivated activity.

INTRODUCTION

RECENT research on intrinsic motivation has consistently found that rewards undermine subsequent interest for initially attractive tasks, apparently by promoting attributions of task engagement to the reward rather than to pleasure in the activity itself (see recent reviews by Lepper, 1983; Morgan, 1984). However, less attention has been directed to clarifying how initial interest can be maintained, or possibly even enhanced. Cognitive evaluation theory (Deci, 1975; Deci and Ryan, 1980) maintains that interest will not be undermined by rewards which are perceived as providing positive information about competence rather than as sources of control. However, Harackiewicz, *et al.* (1984) note that this prediction has not received consistent empirical support and suggest that this may be because the mechanism by which positive information is supposed to enhance interest remains unclear.

Similar concerns have been expressed by several researchers who argue that theories which focus primarily on the implications of exogenous perceptions of causality for task resumption have shed little light on the ways in which the experience and process of task engagement may differ under intrinsic or extrinsic conditions (Csziksentmihalyi, 1975; Maehr, 1976, Condry, 1977; deCharms, 1983). Thus it is also not clear what relationship, if any, should exist between interest, or intrinsic motivation, and performance. While early studies (Kruglanski *et al.*, 1971; Greene and Lepper, 1974) suggested that rewards may undermine creative or divergent thinking, these findings seemed compatible with general behaviour theory and had little impact on conceptualisations of intrinsic motivation. Moreover, later studies tended to find that performance was largely unrelated either to the presence or absence of rewards or to increases or decreases in subsequent interest (see review by Morgan, 1984).

Effects of Task-Involving

A somewhat clearer picture of intrinsically motivated activity seems to emerge from the approaches of deCharms (1968) and Nicholls (1983, 1984). Both authors distinguish between extrinsic, task-involved and ego-involved motivational orientations according to their characteristic foci of attention — on external goals, mastery and self-worth respectively. Intrinsic, or task-involved, motivation is characterised by the concern to improve mastery vis-à-vis task demands and/or prior performance and should be maintained as long as the task is perceived as relevant to the ongoing development and assessment of individual mastery (Nicholls, 1983, 1984). Bandura too emphasises the role of information seeking about competence in maintaining self-regulated motivation (Bandura and Schunk, 1981; Bandura, 1982).

Such positive definitions of intrinsic motivation in terms of competence development and assessment seem to have clearer implications for interest enhancement and for the relation between interest and performance than the largely negative definition of those analyses which emphasise the absence of extrinsic constraints. Thus one can predict that continuing task-involvement and interest should importantly depend on the availability of information about performance, since most activities in both experimental and applied settings do not in themselves provide a basis for the self-evaluation of competence (Bandura and Cervone, 1983). In addition, the focus of attention on mastery suggests that task-involvement will also enhance performance. Moreover, interest and performance should be maintained, at least in the short term, even when no further information is anticipated, since its provision in the past will have enhanced perceptions of the task as relevant to developing mastery and will have promoted the formation of internal standards for evaluating performance and setting goals.

This analysis seems to have clear implications for educational settings, which continually evaluate student performance. However, not all information should enhance interest, even when the task in question is attractive. Controlling information — if you engage in this activity you will be rewarded — should result in a shift from the initial task-involved orientation to the extrinsic one studied by means-end analyses. Normative information, such as that employed as an "intrinsic" motivational condition in studies guided by cognitive evaluation theory, should also promote a shift from initial task-involvement, in this case to a self-worth or ego-involved motivational orientation. Here attention is focused on demonstrating high ability or masking low ability and there is self-esteem-based pressure to achieve positive and avoid negative outcomes (Nicholls, 1984). While ego-involvement has been studied mainly in the context of achievement behaviour, rather than of continuing interest, research findings indicate that task choice, satisfaction and performance are importantly mediated by perceived ability in ego-involving settings (see review by Nicholls, 1984).

This distinction between task-involved and ego-involved motivation suggests several predictions as to the effects of different kinds of information on both performance and interest. Individualised information which relates specifically to aspects of the task which have or have not been mastered should maintain initial task-involvement in both more and less successful students. Thus one can predict high interest and enhanced performance, even when no further information is anticipated. In contrast, the anticipation of ego-involving information, such as the normative grades prevalent in schools, should have differential effects at high and low levels of perceived ability. While grades may maintain initial levels of interest and performance in high achievers who expect these to be self-enhancing, these should be undermined in low achievers. However, the prediction that task and ego-involvement will yield similar levels of performance in high achievers who anticipate

further evaluation may be valid only for tasks requiring convergent thinking. Thus findings that divergent or creative thinking is undermined by ego-involving cues such as time limits (Wallach and Kogan, 1965) and the anticipation of evaluation (Amabile, 1979) suggest that even able subjects will perform better on such tasks when they expect comments than when they expect grades. Finally, one can predict that both interest and performance will decline even among high achievers if no further normative information is anticipated, since the task will no longer be perceived as relevant to demonstrating high ability.

This may be the psychological mechanism behind observations that pupils often seem more interested in their own and their friends' grades than in learning. Nicholls' findings (1978) on developmental transitions in the concept of ability suggest that normative evaluation may be less ego-involving in the early grades, when pupils do not perceive ability as a stable trait and tend to judge mastery relative to task demands and past performance rather than to others' outcomes. However, by fourth grade most pupils have achieved a differentiated concept of ability as a stable trait best assessed by comparison with others. At this point grades should induce a shift from task- to ego-involvement. If so, one can predict further that, just as combining initial interest with extrinsic incentives results in an extrinsic orientation, combining task and ego-involving information should induce ego-involvement, at least after fourth grade. While many teachers seem to feel that any negative effects of grades can be overcome by adding a personal comment, the present analysis suggests that this practice should affect interest and performance much as do grades alone.

Verification of these predictions as to the motivational impact of individual comments, normative grades and comments plus grades necessitated a complex design which would compare their effects on interest and performance before, during and after receipt of information, for both a convergent and a divergent task and for both high and low levels of perceived ability. In the present study, fifth and sixth grade pupils, randomly sampled from either high or low levels of school achievement, received two interesting tasks, one convergent and one divergent, on each of three sessions. No evaluation was anticipated on Session 1 (pre-test); at Session 2, pre-test tasks were returned with appropriate comments, numerical grades, or both and subjects expected to receive similar evaluation on new tasks; at Session 3 (post-test) evaluation was received for Session 2 tasks, but not anticipated for present ones. Interest and performance were measured at each session.

The experimental hypotheses can be summarised as follows:

(a) Post-test interest and performance on both tasks will be highest after receipt of comments at both levels of school achievement.

(b) High achievers will score similarly on Session 2 interest and convergent thinking in all groups, while low achievers will score highest after comments; both high and low achievers will score highest on immediate divergent thinking after comments.

(c) Subjects who received comments alone will recall these better than subjects who also receive a grade; changes in performance from pre-test to post-test will be related to the content of the comments received earlier in the Comments but not in the Grades + Comments condition.

(d) Patterns of interest and performance on Sessions 2 and 3 will be similar in the Grades and Grades + Comments conditions.

METHOD

Sample

The sample comprised 132 fifth and sixth grade Jewish Israeli pupils (68 boys, 64 girls, mean age 11·10) from 12 randomly selected classes (out of 23) in four elementary city schools serving predominantly middle-class populations. Four classes were randomly assigned to each of the three experimental conditions. Although all pupils participated in the experiment, data were analysed only for 44 pupils in each condition, randomly selected from all those whose average grade in language and mathematics in their most recent report card was in the top or bottom 25 per cent for their class. Thus there were 22 high achievers and 22 low achievers in each experimental condition.

Instruments

These consisted of three work booklets containing the experimental tasks for Sessions 1, 2 and 3 respectively. Each booklet contained two tasks, A and B. For Sessions 1 and 3 Task A consisted of a design of five concentric circles, with three or four letters of the Hebrew alphabet printed in each band. Subjects were asked to construct as many words as they could from these letters, starting with a letter in the centre and then adding a letter from one or more successive bands. Task B consisted of two examples from the divergent thinking "uses" test (Torrance and Templeton, 1963). Tasks for Session 2 were slightly different, so as to reduce boredom and practice effects. For Task A, children were asked to construct words from the letters of a long word. Task B consisted of the divergent thinking "circles" test. A pilot study established that pupils of this age found the tasks interesting and that the tasks given on Sessions 1 and 3 yielded equivalent levels of performance.

In addition, an interest questionnaire given after each session asked pupils to rate their interest and enjoyment of the tasks and the degree to which other pupils would find them interesting on seven-point scales. After Sessions 1 and 3 subjects were also asked to state how many additional tasks they would like to receive (from 1 to 7) and after Session 3 also to recall the evaluation received on Session 2.

Feedback conditions

Feedback was given after Sessions 1 and 2 as follows. Unless clearly inappropriate, each child was given similar feedback after each session to reduce the possibility that discrepancies would in themselves affect post-test measures.

(i) *Comments group.* Feedback consisted of one sentence, which related specifically to the performance of the individual child but did not include any information beyond that specified in the criteria for success given to all subjects in the general instructions before each session. For each session, comments took one of the following standardised forms, with slight variations in wording but not content:

> *Task A:* You thought of quite a few correct words; maybe it is possible to think of more short words/more long words/even more words.
>
> *Task B:* You thought of quite a few interesting uses (or "ideas" for the circles task); maybe it is possible to think of more uses (or ideas. . .)/more different uses/more unusual uses which other children may not think of.

Thus the first phrase for each task was the same, but was followed by one of three different suggestions for improvement.

(ii) *Grades group.* Final performance scores for all pupils, including those of average ability, were computed as described below. These scores were then converted into grades so as to follow a normal distribution ranging from 40 to 99.

Thus high achievers tended to receive relatively high grades and low achievers relatively low ones.

(iii) *Grades + Comments group.* Subjects were given both a grade and a comment, as described above.

Procedure

The experiment consisted of three sessions. Session 1 was conducted in one day and Sessions 2 and 3 two days later, with an interval of three hours between them. Sessions were conducted during regular school hours by one of two undergraduate students in psychology. Session 1 instructions, identical for each feedback group, explained that the experimenters had constructed some tasks and were interested in seeing how different children answered them; they hoped that the children would enjoy doing them. Instructions for Task A were then read out. These included rules regarding acceptable words and criteria for successful performance (according to the correctness, number and length of words). Pupils were asked to begin and after 10 minutes to stop and turn to Task B. Instructions for the "uses" task, adapted from Torrance and Templeton (1963), also included criteria for success (according to the number, variety and originality of responses). After 10 minutes the tasks were collected and the interest questionnaire distributed.

In Session 2, two days later, these tasks were returned. Pupils in Group 1 were told that each had been given an appropriate comment on his/her performance. Pupils in Group 2 were told that each had been given a grade which indicated how they had done relative to other pupils in the class, and those in Group 3 that each had been given a grade and a comment. All pupils were instructed to look at the tasks, to see how they had done, after which they would be given new tasks. Pupils in Group 1 were told that they would receive a comment on the new tasks too, in Group 2 that they would receive a grade and in Group 3 that they would again receive a grade and a comment. Booklet 2 was then distributed and the procedure for Tasks A and B was followed as for Session 1. The interest questionnaire was then given out. In Session 3, three hours later, booklet 2 was returned with the evaluation appropriate to each group and booklet 3 was distributed. Pupils were given a few minutes to look through booklet 2 and were then told that the experimenters would like them to try out some new tasks. They were told that these would not be evaluated or returned. The procedure for Tasks A and B and for the interest questionnaire was identical to previous sessions. On completion, pupils were engaged in a discussion about the experiment.

Scoring

Task A. Note was made of the number of short (two or three letter words) and long (four or more letter) words and of the total number of words. A final score was computed by awarding each short word one point and each long one two points.

Task B. Scoring was according to the categories defined by Torrance and Templeton (1963). Counts were made of the number of responses (fluency), categories (flexibility), elaborated responses and original responses, where originality was defined as a non-bizarre idea which appeared in no more than 10 per cent of the protocols. A final score was computed from the sum of scores in each category, which for the "uses" task were averaged over the two examples given in Sessions 1 and 3. Originality was given a weight of 2 and the other components a weight of 1. Tasks were scored by two judges working independently. Inter-judge agreement for final scores was high: $r = 0{\cdot}91$ for Session 1, $r = 0{\cdot}89$ for Session 2 and $r = 0{\cdot}93$ for Session 3.

RESULTS

Performance measures

Table 1 presents mean scores by feedback condition and achievement level for final Task A and Task B performance scores at Sessions 1, 2 and 3. A two-way analysis of variance by feedback condition and school achievement yielded no significant main or interaction effects for feedback condition on Session 1 final scores. The significant main effects for school achievement, $F(1, 125) = 13{\cdot}02$, $P < 0{\cdot}001$ for Task A and $F(1, 125) = 6{\cdot}01$, $P < 0{\cdot}05$ for Task B, confirmed that pupils whose school grades were high performed better on both tasks than pupils whose grades were low. Since preliminary analyses yielded no significant main or interaction effects for sex on any of the dependent variables, data for the sexes were combined for further analysis. Final Task A and Task B scores on Sessions 2 and 3 were then analysed with 3 × 2 (Feedback Condition × School Achievement) ANOVAS. When appropriate, specific hypotheses were then tested using orthogonal planned contrasts.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR FINAL SCORES AT EACH SESSION BY FEEDBACK CONDITION AND SCHOOL ACHIEVEMENT

		Comments group		Grades group		Grades plus Comments group	
		High	Low	High	Low	High	Low
Task A							
Session 1	M	18·77	10·14	19·64	9·86	19·60	9·64
	SD	6·63	5·02	8·67	4·88	5·80	5·64
Session 2	M	25·36	17·86	24·95	12·50	16·77	6·55
	SD	9·33	8·70	9·91	7·86	5·61	4·73
Session 3	M	24·27	13·50	16·45	8·59	11·82	5·82
	SD	6·70	7·79	9·34	6·41	5·26	5·40
Task B							
Session 1	M	19·36	10·27	19·68	10·32	18·68	10·12
	SD	5·35	4·01	7·36	3·71	6·60	3·63
Session 2	M	25·59	17·64	16·09	12·59	14·36	11·95
	SD	9·39	10·16	6·63	5·94	5·17	3·71
Session 3	M	24·95	14·27	14·91	8·50	14·95	9·18
	SD	5·38	4·61	6·63	3·07	5·83	3·40

Session 2

Task A. It was hypothesised that while high achievers would score higher than low achievers in all conditions, low achievers would score highest after comments and the scores of high achievers would be similar in all feedback conditions. Although the interaction effect was not significant, a planned comparison for low achievers who received comments versus those who received grades or grades plus comments was significant, $F(1, 125) = 14{\cdot}54$, $P < 0{\cdot}001$; however, so too was that comparing the scores of low achievers after grades and grades plus comments, $F(1, 125) = 5{\cdot}84$, $P < 0{\cdot}05$. While the former comparison was not significant for high achievers, that comparing scores in the grades and grades plus comments conditions was, $F(1, 125) = 10{\cdot}85$, $P < 0{\cdot}001$. Thus both high and low achievers scored higher in the grades than in the grades plus comments condition.

Task B. As predicted, the results yielded a significant main effect for feedback condition, $F(2, 125) = 17{\cdot}67$, $P < 0{\cdot}001$ and a non-significant interaction effect. Planned contrasts confirmed that over both levels of achievement pupils who received comments scored highest on this divergent thinking task, $F(1, 125) = 33{\cdot}96$, $P < 0{\cdot}001$, while there was no significant difference in the scores of those who received grades or grades plus comments. Table 1 indicates that these trends were found at both levels of achievement. It is interesting that although the main effect for school achievement was significant, $F(1, 125) = 14{\cdot}15$, $P < 0{\cdot}001$, low achievers who received comments scored higher than high achievers in the other conditions (see Table 1).

Session 3

It was hypothesised that scores for both tasks would be highest after comments at both levels of achievement and that there would be no significant differences in scores after grades and grades plus comments.

Task A. The ANOVA yielded the predicted main effects for feedback condition $F(2, 125) = 29{\cdot}60$, $P < 0{\cdot}001$ and school achievement, $F(1, 125) = 57{\cdot}48$, $P < 0{\cdot}001$ and a non-significant interaction effect. The appropriate contrasts confirmed that scores over both levels of achievement after comments were higher than combined scores after grades or grades plus comments $F(1, 125) = 98{\cdot}68$, $P < 0{\cdot}001$. However, the comparison between scores in the latter groups was also significant, $F(1, 125) = 11{\cdot}8$, $P < 0{\cdot}001$. Thus, as in Session 2, pupils who received grades scored higher than pupils who received both a grade and a comment (See Table 1).

Task B. The ANOVA yielded the predicted main effects for feedback condition, $F(2, 125) = 44{\cdot}36$, $P < 0{\cdot}001$ and school achievement, $F(1, 125) = 75{\cdot}73$, $P < 0{\cdot}001$. The appropriate contrasts confirmed that divergent thinking scores over both levels of achievement were highest after comments, $F(1, 125) = 160{\cdot}1$, $P < 0{\cdot}001$ and did not differ significantly after grades and grades plus comments. As in Session 2, low achievers who received comments scored higher than high achievers in the other conditions (see Table 1).

Interest measures

Preliminary two-way analyses of variance yielded no significant main or interaction effects for Session 1 ratings for any of the four interest questions. Since intercorrelations between ratings for the three measures which tapped perceived interest were high: $r\ (132) = 0{\cdot}68$ to $0{\cdot}79$ for Session 1; $0{\cdot}72$ to $0{\cdot}82$ for Session 2 and $0{\cdot}71$ to $0{\cdot}81$ for Session 3, responses to these questions were averaged into a single measure. The question asking how many more tasks pupils would like to receive, which was asked only after Sessions 1 and 3, was analysed separately because of its different face content. Table 2 presents the between-session correlations for these measures within each feedback condition. As hypothesised, these were considerably higher in the Comments than in the Grades or Grades plus Comments conditions. Mean interest ratings for Sessions 2 and 3, presented as Table 3, were then analysed using two-way ANCOVAS, with the corresponding Session 1 ratings as the covariate.

Session 2

As expected, the analysis for the composite measure of perceived interest yielded a significant effect for the interaction of feedback condition with school achievement, $F(2, 125) = 27{\cdot}44$, $P < 0{\cdot}001$. The significant effect for the covariate, $F(1, 124) = 37{\cdot}21$, $P < 0{\cdot}001$ indicated that interest ratings on Sessions 1 and 2 were linearly related. Orthogonal planned contrasts performed on the adjusted cell means

confirmed that the combined interest of high achievers who received grades or grades plus comments was higher than that of low achievers in these conditions, $F(1, 125) = 13{\cdot}56$, $P < 0{\cdot}001$, while the interest of high and low achievers in the comments group did not differ significantly. Low achievers expressed most interest after comments, while high achievers expressed similar interest in all feedback conditions (see Table 3).

TABLE 2

INTERCORRELATIONS BETWEEN INTEREST MEASURES FOR EACH EXPERIMENTAL SESSION, BY FEEDBACK CONDITION

	Feedback Condition		
	Comments group	Grades group	Grades plus Comments group
Perceived Interest			
Session 1 with Session 2	0·31**	0·18	0·16
Session 1 with Session 3	0·42**	0·19	0·21*
Session 2 with Session 3	0·36**	0·24*	0·19
Extra Tasks Requested			
Session 1 with Session 3	0·43**	0·12	0·11
N	44	44	44

*$P < 0{\cdot}05$ **$P < 0{\cdot}001$

TABLE 3

MEANS AND STANDARD DEVIATIONS FOR INTEREST RATINGS AT EACH SESSION BY FEEDBACK CONDITION AND SCHOOL ACHIEVEMENT

		Comments group		Grades group		Grades plus Comments group	
		High	Low	High	Low	High	Low
Perceived Interest Composite							
Session 1	M	6·00	6·05	6·35	6·27	6·42	6·27
	SD	0·90	0·95	0·88	0·83	0·66	0·55
Session 2	M	6·33	6·23	5·62	3·52	5·71	4·03
	SD	0·46	0·65	0·56	1·02	0·49	0·97
	M[a]	6·43	6·31	5·57	3·50	5·63	4·01
Session 3	M	6·61	6·45	4·47	4·39	4·95	4·85
	SD	0·42	0·62	1·50	1·15	1·36	1·18
	M[a]	6·71	6·52	4·42	4·37	4·88	4·93
Extra Tasks Requested							
Session 1	M	5·80	5·64	6·13	5·84	6·18	5·91
	SD	1·22	1·55	0·90	1·11	1·01	1·02
Session 3	M	6·48	6·02	3·50	3·32	4·09	4·00
	SD	0·71	1·29	1·53	1·38	2·02	1·35
	M[a]	6·49	6·24	3·37	3·31	3·94	3·96

M[a] — M is adjusted for Session 1 covariant.

Session 3

As hypothesised, the analysis for perceived interest yielded a significant effect only for feedback condition, $F(2, 125) = 49{\cdot}54$, $P < 0{\cdot}001$. Planned contrasts performed on the adjusted cell means confirmed that the combined interest of high and low achievers was highest after comments, $F(1, 125) = 147{\cdot}6$, $P < 0{\cdot}001$ and that the comparison between the combined ratings of high and low achievers in the grades and grades plus comments conditions was not significant. The analysis for the number of extra tasks requested also yielded only a significant main effect, $F(2, 125) = 66{\cdot}66$, $P < 0{\cdot}001$. Again, the comparison over both levels of achievement between the comments and the other conditions was significant, $F(1, 125) = 245{\cdot}4$, $P < 0{\cdot}001$, while that between the grades and grades plus comments condition was not.

Relation between Session 1 comments and Session 3 performance

The theoretical framework developed above implied that performance on Session 3 would be more closely related to the content of the comments received after Session 1 in the Comments than in the Grades plus Comments condition. Pupils in the Comments and Grades plus Comments conditions were further divided into three groups according to which of the three possible comments they had received after each task on Session 1. Table 4 presents the percentages of subjects whose Session 3 scores on the various components of each task were higher than at Session 1.

TABLE 4

PERCENTAGES OF PUPILS WHO IMPROVED ON EACH TASK COMPONENT BY COMMENT RECEIVED ON SESSION 1

	Suggestion for Improvement			
	Task A			
	Short words		Long words	
Positive change for Session 3 — Session 1	Comment	Grades plus Comments	Comment	Grades plus Comments
Short words	87	31	42	24
Long words	27	25	95	35

	Task B					
	More ideas		More different ideas		More unusual ideas	
Positive change for Session 3 — Session 1	Comment	Grades plus Comments	Comment	Grades plus Comments	Comment	Grades plus Comments
Fluency	100	22	46	31	46	10
Flexibility	50	22	92	31	38	30
Originality	22	22	8	13	77	20

In the Comments condition one finds a clear relation between the content of the Session 1 comment and improvement on specific aspects of each task. Thus 87 per cent of the pupils told after Session 1 that they had written several correct words but could have written more short words did write more short words on Session 3, while only 27 per cent wrote more long words. Similarly, 92 per cent of subjects told that they could have written more long words did so, as compared with 42 per cent who wrote more short ones. As Table 4 indicates, the same pattern of differential improvement by comment received was found also for Task B. While these results

are hardly surprising in themselves, they are of considerable interest when compared with those for pupils who received grades plus comments. As could be expected from the data for final scores reported above, overall fewer pupils in this condition improved their scores on specific components of task performance. However, it is striking that similar percentages of subjects improved their scores on each component, regardless of which specific suggestion for improvement they had received on Session 1. Thus, for example, 35 per cent of the subjects who were told that they could have written more long words did so, but this proportion is similar to the 25 per cent who had been told after Session 1 that they could have written more short words. Table 4 reveals the same pattern of non-differential improvement also for Task B.

Recall of evaluation

This was tapped by asking pupils to recall the evaluation they had received for Task B on Session 2. Task B was chosen since the comments for Session 2 were worded somewhat differently than those for Session 1. It was emphasised that we were interested in seeing how well subjects could remember their evaluation in order to reduce the possibility that pupils might recall their evaluation but refrain from writing this down. The number of pupils who recalled their grade and/or comment accurately in each feedback condition are presented as Table 5. As hypothesised, almost all pupils who received a grade or both a grade and a comment remembered the grade. Somewhat fewer pupils who received only a comment remembered all of it, but all but two recalled at least one component. However, only 45 per cent of the pupils who received both a grade and a comment remembered even part of the comment, although they were specifically asked to recall both the grade and comment.

TABLE 5

RECALL OF TASK B SESSION 2 EVALUATION IN EACH FEEDBACK CONDITION

	Comments group	Grades group	Grades plus Comments group
Percentage of subjects recalling evaluation			
Comment alone	84 (95[a])	—	5 (11[a])
Grade alone	—	89	53
Grade plus Comment	—	—	23 (34[a])

[a] — per cent who recalled only one component of the comment.

DISCUSSION

The results confirm the importance of distinguishing between task-involvement and ego-involvement when investigating intrinsic task motivation. As hypothesised, both high and low achievers who received comments continued to express high interest both on Session 2, when they anticipated further comments, and at post-test, when they did not. Similar patterns were received also for their performance on both convergent and divergent tasks. Additional findings that interest was slightly higher at post- than at pre-test, that performance on both tasks was considerably higher at post-test and that later interest was highly correlated with initial interest provided further support for the prediction that initial task-involvement would be maintained in this condition.

In contrast, it was hypothesised that normative grades would cause a shift from the initial task-involved orientation at pre-test to an ego-involved orientation in later sessions, not only when grades alone were provided, but also when these were given in conjunction with task-involving comments. As hypothesised, in both conditions initial interest was less predictive of interest on both later sessions than after comments. In addition, the results confirmed the predictions that immediate interest and convergent thinking would be maintained for high achievers and undermined for low achievers, and that both immediate divergent thinking and subsequent interest and performance on both tasks would be undermined at both levels of achievement. Thus task-involving feedback does seem to have different effects on both interest and performance than ego-involving feedback. In addition, while many teachers seem to feel that any negative effects of grades can be ameliorated by adding a personal comment, the above results suggest that this practice too will induce an ego-involved orientation. This suggestion was further supported by the finding that pupils did indeed tend to recall the grade rather than the comment.

While the differential effects on interest of comments as compared with grades and grades plus comments are consistent with the above characterisation of these conditions as task and ego-involving respectively, it is important to consider whether they could not have been predicted also by cognitive evaluation theory. Both anecdotal evidence (Holt, 1964) and some research findings (Maehr and Stallings, 1972; Harter, 1978) suggest that grades are perceived as potent sources of control over learning. The present findings of reduced subsequent interest after grades as compared with comments is consistent with cognitive evaluation theory's prediction that controlling information will undermine interest relative to non-controlling information, and Butler and Nisan (1986) did indeed interpret their findings that grades yielded lower subsequent interest than task-related comments within this framework. Moreover, if grades are perceived primarily as salient extrinsic incentives, it is not surprising that they continued to undermine subsequent interest in the present study also when they were given in conjunction with comments.

However, in this case, the two grades conditions should also have undermined interest relative to pre-test on Session 2 and not only on Session 3, since at both points the grades should have provided sufficient justification for task engagement. Moreover, these decrements should have been more marked for low achievers who received poor grades (negative information) than for high achievers who received high ones, not only on Session 2 but also on Session 3. Finally, the comments provided, while probably less controlling than grades, seem at least as controlling as the phrase: "Good. You should keep up the good work" which Ryan (1982) found to undermine interest. In addition, they provided mixed rather than clearly positive information. Thus conceptualising these feedback conditions as task-involving and ego-involving respectively seems to provide a more comprehensive framework than the controlling-informational distinction of cognitive evaluation theory for interpreting both the differential effects of grades on immediate as compared with subsequent interest and the consistently high interest expressed after comments on all sessions. Furthermore, while normative information about outcomes may undermine subsequent interest less than tangible rewards, the implication that this does indeed seem to promote ego-involvement rather than task-involvement reinforces the reservations noted above regarding its conceptualisation as an "intrinsic" motivational condition in many studies.

It seems less relevant to ask how cognitive evaluation theory would interpret the findings for performance, since this has been concerned primarily with predicting subsequent interest. Those studies which have measured performance have tended to compare performance during the manipulation with interest at post-test. These

studies have yielded inconsistent results (Morgan, 1984), which have led some researchers (cf. Deci and Ryan, 1980) to the somewhat counter-intuitive conclusion that there is no necessary relation between performance and interest. In contrast, our results suggest that different kinds of feedback, conceptualised as task and ego-involving respectively, do indeed affect both interest and performance in similar ways. However, this becomes clear only when these are both assessed at the same points in the motivational process.

In addition, the present findings also provide some insights as to the ways in which task-involved and ego-involved motivational orientations affect the quality of performance. Pupils who received comments improved their Session 3 performance primarily on those components of each task which their Session 1 comments had suggested they improve. This trend is consistent with the hypothesis that task-involved, intrinsically motivated subjects will be receptive to and will actively utilise cues which can facilitate the formation of internal standards for guiding and evaluating performance (Bandura and Cervone, 1983). However, it could also be argued that they were simply "following instructions". Alternatively, since suggestions for improvement usually related to that aspect of the task for which initial scores were relatively low, their differential improvement may have been due to a ceiling effect alone. If so, enhanced performance after comments may have had little to do with task-involved or with any other kind of motivational orientation. However, were this the case one would expect to find similar increments in the performance of pupils who received both a grade and a comment. Instead, these pupils as a group did worse on Session 3 than on Session 1, and even those who improved seemed to do so without relation to the specific content of the comment they had received. Thus while our results suggest that Butler and Nisan's (1986) interpretation of similar performance increments after comments as compared with grades and no feedback in terms of higher concurrent interest alone was too narrow, they also suggest that motivational factors importantly determine the conditions under which specific guidelines will indeed be utilised to improve performance. However, further research is necessary both to establish the reliability of these rather complex effects and to pinpoint more clearly the relative effects of interest and information on subsequent performance.

While ego-involving evaluation affected interest rather differently than the extrinsic incentives employed in other studies, their effects on immediate performance seem quite similar. Thus anticipation of grades undermined performance on the novel, divergent Task B more than on the convergent Task B. In addition, while final Task A scores on Session 2 were similar after grades and comments, this was because the former generated more short words, but fewer long words than the latter. Similarly, fluency was less undermined by grades than were the other components of divergent thinking. Thus anticipation of grades, as of rewards, seems to encourage even high achievers to seek reward attainment in the easiest way possible, by concentrating on the quantity rather than on the quality of performance (Greene and Lepper, 1974).

The reduced interest and performance of low achievers who anticipated further grades, given either with or without a comment, is hardly surprising in view of the considerable anecdotal and research evidence that competitive, ego-involving situations have more adverse effects on the learning and motivation of low than of high ability students (Holt, 1964; Covington and Beery, 1976; Ames, 1984; Nicholls, 1984). Indeed, critics of grading tend to focus on the plight of the failing student. However, the present findings suggest that ego-involving settings may exact a not inconsiderable price among successful students too. Narrow preoccupation with grade attainment seems to affect the quality, if not the quantity, of immediate task

performance, and to undermine divergent thinking in particular. Perhaps more significantly, the maintenance of ego-involved motivation seems to depend on the continued availability of opportunities for social comparison, so that both effort and interest wane when these are withdrawn. Thus it is suggestive that while low achievers in the grades groups expressed less interest when they anticipated further grades than when they did not, high achievers expressed least interest at post-test. In contrast, the findings for comments suggest that task-involvement is beneficial not only at low, but also at high levels of achievement.

In this context, one should note that while pupils at both levels of ability received similar comments, grades were rewarded objectively so that high achievers tended to receive high grades and low achievers low ones. Thus any effects of generalised perceptions of ability were confounded with those to be expected after experiencing specific success or failure outcomes. However, it is probably more difficult to separate these effects than is generally realised. Frieze *et al.* (1983) review several findings which suggest that objectively similar information about outcomes will be interpreted differently by different subjects, possibly as a function of such achievement-related variables as causal attributions for outcomes, achievement motivation or perceived ability. Similarly, balancing designs by giving both high and low achievers either success or failure feedback also provides some subjects with information that disconfirms expectancy, as when low achievers receive success feedback or high achievers failure feedback, which seems in itself to have considerable and differential motivational impact (Harackiewicz *et al.*, 1985). The present design, while problematic, seemed to reduce at least this pitfall, since the grades received were probably consistent with subjects' experience and expectancies, as in classroom situations.

To conclude, on a theoretical level, the results of the present study suggest that some of the difficulties faced by mean-ends analyses in conceptualising and predicting how interest and performance can be maintained or enhanced, and not just undermined, may be resolved by distinguishing not only between constrained and non-constrained but also between task-involved and ego-involved task-engagement. In addition, they suggest that different motivational orientations have implications not only for subsequent interest, but also for immediate interest and performance and for subsequent performance. Moreover, they suggest that, at least for the age-group studied, combining task and ego-involving evaluation will induce an ego-involving orientation, just as does the provision of ego-involving evaluation alone. These results seem to have implications for applied settings. While applications of the literature on rewards generally discuss when these will be more or less undermining of interest and performance (deCharms, 1983; Lepper, 1983), this study implies that promoting task-involvement may also enhance the interest and performance of most students. The comments employed here were based on a simple principle, easily applicable in the classroom, of relating both to an aspect of the task performed adequately and to one which could be improved. However, further research is clearly necessary to clarify the effects over time of systematically providing such feedback in applied settings and of reducing the use of normative grades.

Correspondence and requests for reprints should be addressed to Dr. Ruth Butler, School of Education, The Hebrew University of Jerusalem, Jerusalem 91905, Israel.

REFERENCES

AMABILE, T. M. (1979). Effects of external evaluation on artistic creativity. *J. Person. soc. Psychol.*, 37, 221-233.

AMES, C. (1984). Competitive, cooperative and individualistic goal structures: a cognitive-motivational analysis. In AMES, R., and AMES, C., (Eds.), *Research on Motivation in Education, Vol. 1*, Orlando, Flor.: Academic Press.

BANDURA, A. (1982). Self efficacy mechanism in human agency. *Am. Psychol.*, 37, 122-147.

BANDURA, A., and CERVONE, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal-systems. *J. Person. soc. Psychol.*, 45, 1017-1028.

BANDURA, A., and SCHUNK, D. H. (1981). Cultivating competence, self-efficacy and intrinsic interest through proximal self-motivation. *J. Person. soc. Psychol.*, 41, 586-598.

BUTLER, R., and NISAN, M. (1986). Effects of no feedback, task-related comments and grades on intrinsic motivation and performance. *J. educ. Psychol.*, 78, 210-217.

CONDRY, J. C. (1977). Enemies of exploration: self-initiated versus other-initiated learning. *J. Person. soc. Psychol.*, 35, 459-477.

COVINGTON, M. V., and BEERY, R. (1976). *Self Worth and School Learning*. New York: Holt, Rinehart and Winston.

CSIKSZENTMIHALYI, M. (1975). *Beyond Boredom and Anxiety*. San Francisco: Jossey-Bass.

DECHARMS, R. (1968). *Personal Causation*. New York: Academic Press.

DECHARMS, R. (1983). Intrinsic motivation, peer tutoring and cooperative learning: practical maxims. In LEVINE, J. M., and WANG, M. C. (Eds.), *Teacher and Student Perceptions*. Hillsdale, NJ: Lawrence Erlbaum.

DECI, E. L. (1975). *Intrinsic Motivation*. New York: Plenum Press.

DECI, E. L., and RYAN, R. M. (1980). The empirical exploration of intrinsic motivational processes. In BERKOVITZ, L. (Ed.), *Advances in Experimental Social Psychology, Vol. 13*. New York: Academic Press.

FRIEZE, I. H., FRANCIS, W. D., and HANCUNA, B. H. (1983). Defining success in classroom settings. In LEVINE, J. M., and WANG, M. C. (Eds.), *Teacher and Student Perceptions*. Hillsdale, NJ: Lawrence Erlbaum.

GREENE, D., and LEPPER, M. R. (1974). Effects of extrinsic rewards on children's subsequent intrinsic interest. *Child Dev.*, 45, 1141-1145.

HARACKIEWICZ, J. M., MANDERLINK, G., and SANSONE, C. (1984). Rewarding pinball wizardry: effects of evaluation and cue value on intrinsic interest. *J. Person. soc. Psychol.*, 47, 287-300.

HARACKIEWICZ, J. M., SANSONE, C., and MANDERLINK, G. (1985). Competence, achievement orientation and intrinsic motivation: a process analysis. *J. Person. soc. Psychol.*, 48, 493-508.

HARTER, S. (1978). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. *Child Dev.*, 49, 788-799.

HOLT, J. (1964). *How Children Fail*. New York: Dell.

KRUGLANSKI, A. W., FRIEDMAN, J., and ZEEVI, G. (1971). The effects of extrinsic incentives on some qualitative aspects of task performance. *J. Personality*, 39, 606-617.

LEPPER, M. R. (1983). Extrinsic reward and intrinsic motivation: implications for the classroom. In LEVINE, J. M., and WANG, M. C. (Eds.), *Teacher and Student Perceptions*. Hillsdale, NJ: Lawrence Erlbaum.

MAEHR, M. L. (1976). Continuing motivation: an analysis of a seldom considered educational outcome. *Rev. educ. Res.*, 46, 443-462.

MAEHR, M. L., and STALLINGS, W. M. (1972). Freedom from external evaluation. *Child Dev.*, 43, 177-185.

MORGAN, M. (1984). Reward-induced decrements and increments in intrinsic motivation. *Rev. educ. Res.*, 54, 5-30.

NICHOLLS, J. G. (1978). The development of the concepts of effort and ability, perception of own attainment and the understanding that difficult tasks require more ability. *Child Dev.*, 49, 800-814.

NICHOLLS, J. G. (1983). Conceptions of ability and achievement motivation: a theory and its implications for education. In PARIS, S. G., OLSON, G. M., and STEVENSON, H. W. (Eds.), *Learning and Motivation in the Classroom*. Hillsdale, NJ: Lawrence Erlbaum.

NICHOLLS, J. G. (1984). Conceptions of ability and achievement motivation. In AMES, R. E., and AMES, C. (Eds.), *Research on Motivation in Education*. Orlando, Flor.: Academic Press.

RYAN, R. M. (1982). Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *J. Person, soc. Psychol.*, 43, 450-461.

TORRANCE, A. E., and TEMPLETON, O. E. (1963). *Minnesota Tests for Creative Thinking*. Minneapolis: University of Minnesota Press.

WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children*. New York: Holt, Rinehart and Winston.

(*Manuscript received 2nd December*, 1986)